Verankerd in leren

Vijf bouwstenen voor professioneel beoordelen in het hoger beroepsonderwijs

Dominique Sluijsmans

Zuyd Onderzoek

Verankerd in leren

Vijf bouwstenen voor professioneel beoordelen in het hoger beroepsonderwijs

Dominique Sluijsmans

Rede in verkorte vorm uitgesproken bij de aanvaarding
van het ambt van lector Professioneel Beoordelen
aan Zuyd Hogeschool op vrijdag 19 april 2013

Inhoudsopgave

Over deze rede 4
Kijken naar studenten als inspiratiebron 5
Van vakscholen naar kennisinstellingen 6
Investeren in beoordelen om de kwaliteit van het diploma te garanderen 8
Waarom een lectoraat Professioneel Beoordelen? 9
Bouwsteen 1: de professie 11
Van stabiele beroepen naar dynamische professies 11
Van Ausbildung naar Bildung 13
Van schools curriculum naar een rijke leeromgeving 15
Bouwsteen 2: een programma van professioneel beoordelen 18
Van losse toetsen naar een programma van beoordelen 18
Van sluitstuk van leren naar start voor leren 21
Van convergent naar divergent beoordelen 23
Bouwsteen 3: de professionele beoordelaar 26
Van vakdocent naar expertbeoordelaar 26
Van één objectieve beoordelaar naar meerdere subjectieve beoordelaars 29
Van informatiearm beoordelen naar informatierijk beoordelen 30
Bouwsteen 4: de professionele student 34
Van informatieverwerver naar kenniswerker 34
Van scholier naar onderzoekende professional 36
Van 'zeg me wat ik moet doen' naar 'ik weet wat ik wil doen' 39
Bouwsteen 5: een professionele beoordelingscultuur 43
Van individueel leren naar collectief leren 44
Van een afhankelijke opstelling naar het nemen van regie 46
Van ad-hockwaliteit naar duurzame kwaliteit 48
Naar een definitie van professioneel beoordelen 52
Samen praktijkonderzoek doen naar professioneel beoordelen 53
Tot slot 58
Referenties 60

Over deze rede

Geacht College van Bestuur,
Beste leden van het lectoraat Professioneel Beoordelen,
Beste collega's van Zuyd en daarbuiten,
Lieve familie en vrienden,

Sinds mijn eerste schooldag op de basisschool in 1977 tot aan mijn afstuderen in 1997 heb ik vele toetsen gemaakt. Het zullen er rond de 4.200 zijn geweest. Deze toetsen werden door individuele leerkrachten en docenten geconstrueerd en nagekeken en standaard aan het einde van een instructie- of onderwijsblok afgenomen. Toetsresultaten werden teruggekoppeld via cijfers, door de leerkracht of een lijst op een prikbord. Deze toetsen vereisten voornamelijk kennisreproductie of kennistoepassing. Het maken van deze toetsen voelde vooral als een hordeloop met het ultieme doel een voldoende te scoren of studiepunten bij elkaar te sprokkelen om je diploma te halen. Het moment van het eindexamen staat nog in mijn geheugen gegrift: een gymzaal met tientallen tafels, een paar meter van elkaar om afkijken en overleg tussen leerlingen te voorkomen.

Deze reeks van losse, onsamenhangende toetsen die een afwachtende, passieve houding bij leerlingen en studenten stimuleren, is niet langer verdedigbaar als het gaat om borgen van kwaliteit van beoordelen in het hoger beroepsonderwijs[1]. De traditionele beelden van beoordelen zoals zojuist geschetst, passen niet meer bij de dynamiek waarin we jonge mensen tot gepassioneerde professionals opleiden. Ik ben daarom blij u vandaag kennis te laten maken met het lectoraat Professioneel Beoordelen. Ik zal uitleggen wat ik met professioneel beoordelen bedoel en welke ontwikkelingen ertoe bijdragen dat we anders over beoordelen in het hbo zouden moeten nadenken. Om professioneel beoordelen gestalte te geven is inzicht in de volgende fundamentele vraag van belang: wie leidt het hbo op?

1 In het vervolg van deze rede aangeduid als hbo

Kijken naar studenten als inspiratiebron

Op dit moment staan 421.560 studenten ingeschreven bij één van de 39 door de overheid bekostigde hogescholen in Nederland[2]. Van dit aantal volgen 53.203 studenten een deeltijdopleiding en 11.584 een duale opleiding. Jaarlijks ontvangen ongeveer 70.000 studenten een diploma. Bijna 15.000 studenten worden opgeleid bij Zuyd Hogeschool. De afgelopen maanden ben ik vaak bij de faculteiten van onze hogeschool op bezoek geweest, niet alleen hier in Heerlen maar ook op de locaties in Sittard en Maastricht. Voor een kennismaking, een overleg, een professionaliseringsmiddag of gewoon voor de gezelligheid. In deze faculteiten kom ik veel studenten tegen. Zij zijn druk in overleg met hun medestudenten en docenten en gebruiken daarbij hun laptops, boeken en telefoons. Ik kom deze studenten ook tegen in bedrijven, in de horeca, in het ziekenhuis, in scholen en in zorginstellingen. Zowel binnen de onderwijsomgeving als in de echte beroepspraktijk werken zij toe naar de rol van zelfstandig beroepsbeoefenaar.

Het is boeiend deze studenten te observeren. Ik realiseer mij dat zij straks een grote verantwoordelijkheid dragen in onze maatschappij. Zij zijn de toekomstige zelfstandig ondernemers die nieuwe spannende plannen zullen bedenken. De analisten, de ontwerpers van digitale netwerken en ontwikkelaars van nieuwe stoffen. Zij zijn de designers die multimedia-applicaties zoals websites, animaties, computerspellen of videofilms zullen bedenken, vormgeven en bouwen. Zij zijn straks de deskundigen die weten hoe ze weefsels van patiënten kunnen onderzoeken. De deskundigen die door het uitvoeren van DNA-onderzoek gerechtelijke onderzoeken ondersteunen. Door hen worden wij straks verwend met nieuwe modeontwerpen, sieraden, decors, foto's en films of ideeën voor het inrichten van ons huis. Zij zijn ook degenen die zorgen voor een aangenaam verblijf in hotels en ons laten genieten van prachtig toneel en dans en ontroerende concerten. Zij helpen straks werknemers loopbaanmogelijkheden uit te stippelen en ondersteunen werkgevers bij de werving en selectie van nieuwe medewerkers. Zij zullen ook degenen zijn die aan ons de juiste zorg kunnen bieden als we dat nodig hebben en ons kunnen helpen ons leven psychisch, sociaal of fysiek gemakkelijker te maken. Zij zullen vele vrouwen bijstaan bij hun zwangerschap en bevalling. Bijdragen aan een goede jeugdhulpverlening, kinderopvang, verslavingszorg en gehandicaptenzorg. Zij zullen jonge kinderen in de school nieuwsgierig maken en hen helpen hun talenten te ontwikkelen. Deze studenten zullen ons leven op een bepaalde manier mooier, makkelijker of zinvoller maken.

Deze denkbeelden van professionals aan het werk inspireren mij om na te denken over hoe we in het hbo het fundament kunnen leggen voor hun loopbanen. Dat fundament leggen is een niet te onderschatten klus. Het betekent de verant-

2 Zie www.hbo-raad.nl

woordelijkheid nemen om studenten in veilige, inspirerende leeromgevingen op te leiden met passende curricula en bekwame docenten. Het betekent ook te kunnen verantwoorden dat het diploma dat wordt uitgereikt van grote waarde is. Als burger moet je er immers op kunnen vertrouwen dat beginnende beroepsbeoefenaren het beroep op een bepaald niveau kunnen uitvoeren: dat er geen gekke dingen gebeuren bij de diagnose van een probleem, dat een ontwerp van een product voldoet aan kwaliteit, dat kinderen onderwijs krijgen dat van hoge kwaliteit is en dat je bij ziekte wordt behandeld met de best voorhanden zijnde zorg. Je hoopt ook dat zij bevlogen hun vak uitoefenen en handelen volgens de laatst beschikbare kennis en professionele standaarden die in het specifieke beroepsdomein gelden.

Van vakscholen naar kennisinstellingen

Het goed opleiden van studenten voor hun toekomstige beroep vraagt om een passende leeromgeving waarin studenten alle mogelijkheden hebben om te leren. Het hbo voorziet in deze omgeving. Maar hoe is het hbo gegroeid tot wat het anno 2013 is? Een overzicht van Van Bemmel (2006) laat zien dat het hbo haar oorsprong al kent in de zeventiende eeuw (zie tabel 1).

Tabel 1. Belangrijke jaartallen uit de geschiedenis van het hbo*

Jaar	*Wat gebeurde er?*
1682	Start hbo-opleiding beeldende kunst te Den Haag (de eerste kunstopleiding)
1785	Oprichting eerste zeevaartopleiding op hbo-niveau (te Amsterdam)
1797	Oprichting eerste hbo-opleiding tot leraar (te Groningen: 'kweekschool')
1818	Opheffing van de gilden in 1818 - eind ambachtelijk opleidingssysteem
1847	Oprichting 'Inrichting van Handel en Nijverheid' te Amsterdam (één van de eerste economische hbo-opleidingen)
1848	Oprichting eerste technische hbo-opleiding (te Rotterdam)
1878	Oprichting opleiding ziekenverpleegster te Amsterdam (één van de eerste gezondheidszorgopleidingen in het hbo)
1899	Oprichting 'Opleidingsrichting voor Socialen Arbeid' te Amsterdam
1919	Wet op het Nijverheidsonderwijs (inclusief een deel van het hbo)
ca. 1925	Oprichting eerste landbouwopleidingen in het hbo (in Groningen en Deventer)

1963	Wet op het Voortgezet Onderwijs (gehele hbo onder overheidsbekostiging gebracht)
1975	Oprichting overlegorgaan HBO-raad
1983	Schaalvergroting, Taakverdeling en Concentratie in het hbo (STC-operatie)
1985	Nota Hoger Onderwijs: Autonomie en Kwaliteit (HOAK)
1986	Wet op het hbo/invoering naam 'hogescholen'/vorming HBO-raad nieuwe stijl
1989	Start hbo-visitaties (kwaliteitszorg)
1993	Wet op het Hoger Onderwijs en Wetenschappelijk Onderzoek (WHW)
1993	Eerste door HBO-raad afgesloten cao-hbo
1994	Hogescholen kopen 'hun' gebouwen van de overheid
2001	Convenant lectoren en kenniskringen in het hoger onderwijs
2002	Invoering Bachelor/Masterstructuur (BaMa)
2003	Start hbo-accreditaties (kwaliteitszorg)
2004	Formulering Dublin descriptoren als eindkwalificaties bachelor en masteropleidingen
2006	Voorstel van Wet op het Hoger Onderwijs en Onderzoek (WHOO)
2007	Brancheprotocol Kwaliteitszorg Onderzoek (BKO) 2009-2015
2009	Kwaliteit als opdracht
2009	Instelling validatiecommissie kwaliteitszorg onderzoek (VKO)
2009	Kennisbases lerarenopleidingen voortgezet onderwijs en primair onderwijs
2010	Gedragscode praktijkgericht onderzoek voor het hbo
2010	Inwerkingtreding Wet Versterking Besturing
2011	Inwerkingtreding nieuw accreditatiestelsel
2012	Differentiëren in drievoud. Advies van de Commissie Toekomstbestendig Hoger Onderwijs Stelsel - Commissie Veerman
2012	Rapport Vreemde ogen dwingen – Commissie Bruijn
2012	Rapport Onderwijs: De derde dimensie (Bruijn & De Vries)

* gegevens na 2006 door auteur aangevuld

Dit beknopte, lang niet uitputtende overzicht laat zien dat het hbo zich sinds het losweken van het voortgezet onderwijs in 1984 - bijna dertig jaar geleden - heeft ontwikkeld van 'losse' vakopleidingen tot kennisinstellingen die zorg dragen voor een aantal bachelor- en masteropleidingen in zeven hbo-sectoren (economie, gezondheidszorg, techniek, landbouw, kunst, pedagogisch en sociaal-agogisch). Iedere hogeschool geeft een eigen invulling aan deze sectoren. Zo kent Zuyd Hogeschool tien faculteiten[3] die samen verantwoordelijkheid dragen voor veertig bachelor- en twaalf masteropleidingen. De ontstane kennisinstellingen, georganiseerd in kleine, middelgrote en grote hogescholen, hebben zich tot doel gesteld te investeren in de kwaliteit van het onderwijs (bijvoorbeeld door het aanbieden van meer masteropleidingen), in het verhogen van de professionaliteit van de studenten en docenten (bijvoorbeeld door het (leren) uitvoeren van onderzoek), in talentontwikkeling en in kennisdeling met de beroepspraktijk (Van Bemmel, 2006). De aandacht voor onderzoek heeft een formeel karakter gekregen door de invoering van lectoraten - waarvan er vandaag de dag bijna vijfhonderd bestaan -, een brancheprotocol Kwaliteitszorg Onderzoek (HBO-raad, 2008), de instelling van de validatiecommissie kwaliteitszorg onderzoek (VKO) en het opstellen van een gedragscode praktijkgericht onderzoek voor het hbo (Commissie Gedragscode Praktijkgericht Onderzoek, 2010). Hogescholen krijgen de taak het onderzoekend vermogen bij studenten te bevorderen zodat zij later bij kunnen dragen aan vernieuwingen in de beroepspraktijk (HBO-raad, 2009). Ook het behalen van masters en doctoraten door onderwijzend personeel is een ontwikkeling die expliciet is vastgelegd in de prestatieafspraken die elke hogeschool in 2012 met het ministerie van Onderwijs, Cultuur en Wetenschappen heeft gemaakt. Het belang van vervlechting van onderwijs en onderzoek wordt herhaaldelijk benadrukt (Commissie Veerman, 2012; Kwakman, 2003; Schuiling, 2012; Van der Klink, 2011). Tot slot is in het hbo een ontwikkeling zichtbaar naar een toegenomen toezicht op hogescholen door het instellen van visitatiecommissies en later het accreditatiestelsel. Samenvattend bevindt het hbo zich in een continu veranderingsproces, waarbij het belang van kwaliteitsborging van de centrale taakstelling van het hbo - het opleiden van onderzoekende professionals - steeds groter wordt (Leijnse, Hulst & Vroomans, 2006).

Investeren in beoordelen om de kwaliteit van het diploma te garanderen

De waarde van diploma's die het hbo uitgeeft moet worden gegarandeerd. Het bewustzijn dat deze garantie voor een aanzienlijk deel wordt bepaald door de kwaliteit van beoordelen is groeiende. Het aanbod van cursussen en boeken op

3 Faculteit Bèta Sciences and Technology, faculteit Commercieel en Financieel Management, faculteit Gezondheidszorg, faculteit Hotel en Facility Management, faculteit ICT, faculteit International Business and Communication, faculteit van de Kunsten, faculteit Sociale Studies en Educatie, faculteit Verloskunde.

het gebied van beoordelen is de laatste twee decennia enorm gegroeid[4]. De noodzaak tot een kwaliteitsimpuls is onlangs omgezet in een aantal politieke acties. In september 2010 is de wet Versterking Besturing in werking getreden. Deze wet houdt onder andere in dat de examencommissie het orgaan is dat op objectieve en deskundige wijze vaststelt of een student voldoet aan de voorwaarden die de onderwijs- en examenregeling (OER) stelt ten aanzien van kennis, inzicht en vaardigheden die nodig zijn voor het verkrijgen van een diploma. Daarmee is de rol van examencommissies verschoven van het uitvoeren van procedures rondom toetsing naar het borgen van de kwaliteit van de toetsing. Een handreiking voor examencommissies vervult hierin een belangrijke functie (HBO-raad, 2011). Deze handreiking gebruiken hogescholen voor het schrijven van een handboek voor examencommissies[5]. Ook is er het nieuwe accreditatiestelsel dat in januari 2011 in werking is getreden en waarbij *toetsing en gerealiseerde eindkwalificaties* als aparte standaard is opgenomen (standaard 3 van het accreditatiestelsel beperkte opleidingsbeoordeling). Recentelijk is het rapport *Vreemde ogen dwingen* gepubliceerd onder leiding van de Commissie Bruijn (2012; zie tabel 1). Deze commissie heeft een aantal maatregelen geformuleerd waarmee externe validering van de kwaliteit van beoordelen kan worden vormgegeven. Voorbeelden van deze maatregelen zijn het gezamenlijk ontwikkelen van landelijke toetsen door de opleidingen zelf, het certificeren van examinatoren, het inrichten van een basis- en seniorkwalificatie examinering (BKE/SKE) voor docenten en het borgen van toetsdeskundigheid in examen- en visitatiecommissies. Samengevat wordt beoordelen niet beschouwd als iets dat losstaat van onderwijs, maar direct de kwaliteit van leren beïnvloedt.

Waarom een lectoraat Professioneel Beoordelen?

Het is een goede zaak dat beoordelen meer aandacht krijgt vanuit de politiek, maar we moeten de vraag blijven stellen of deze maatregelen voldoende zijn voor het opleiden van studenten tot jonge, gepassioneerde professionals. Het lectoraat Professioneel Beoordelen legt zich toe op vragen als: Welke functie heeft beoordelen in het hbo? Wat vraagt toekomstbestendig beoordelen van opleidingen? Wat is er nodig om kwaliteit van beoordelen te borgen? Op basis van wetenschappelijke en maatschappelijke argumenten geef ik invulling aan het begrip *professioneel beoordelen* door vijf bouwstenen te benoemen die belangrijk zijn voor een visie op beoordelen die past bij de dynamische context van het hbo. Door het formuleren van kernboodschappen leg ik steeds de relatie tussen de betreffende bouwsteen en professioneel beoordelen. Als de vijf bouwstenen de revue zijn gepasseerd licht ik toe hoe het lectoraat de komende jaren aan de slag gaat.

4 Zie bijvoorbeeld het boek Toetsen in het hoger onderwijs van Berkel en Bax (2002) dat al aan een derde druk toe is.

5 Zie voor een mooi voorbeeld het Handboek Examencommissies van de Hogeschool van Arnhem en Nijmegen (HAN).

Voor beoordelingsvraagstukken zal deze rede geen kant-en-klare praktische oplossingen bieden. Ook zal ik lang niet recht doen aan de volle breedte van het hbo en het kennisdomein van beoordelen. Bij een aantal van u roep ik wellicht het gevoel op van 'Weer iets nieuws?' of 'Dat doen we toch al lang?' Misschien zeg ik in mijn enthousiasme dingen waar u het grondig mee oneens bent. Ik streef met deze rede een fundamentele discussie over beoordelen in het hbo na, door kritische vragen te stellen die uitgangspunt zullen zijn voor verder onderzoek. De titel van mijn rede is in dit kader niet toevallig gekozen. Ik hoop u ermee te kunnen inspireren om beoordelen te verankeren in de visie op leren. Hoe dan ook stel ik uw kritische feedback op mijn betoog op prijs.

De professie

Hoger beroepsonderwijs betekent dat een student wordt opgeleid voor een professie en dat na een aantal jaren wordt beoordeeld of de student deze professie op een acceptabel niveau beheerst. In dit hoofdstuk leg ik de eerste bouwsteen van professioneel beoordelen: de professie. Drie ontwikkelingen met directe consequenties voor beoordelen in het hbo zijn:

1) de ontwikkeling van stabiele beroepen naar dynamische, complexe professies;
2) de ontwikkeling van beroepsvaardigheden naar bredere vaardigheden; en
3) de ontwikkeling van schoolse curricula naar rijke leeromgevingen.

Onderstaand wordt elke ontwikkeling beknopt geschetst en aangegeven wat de implicaties voor beoordelen zijn. Het hoofdstuk vat ik samen in enkele kernboodschappen.

Van stabiele beroepen naar dynamische professies

Ik stel mijn kinderen - ze zijn nu zeven en negen jaar oud - wel eens de vraag: wat wil je later worden? Deze vraag werd ons als kind op school ook vaker gesteld. Meestal noemden de jongens beroepen als brandweerman of politieman. Bij meisjes kwam het vaak niet verder dan juffrouw of verpleegster (of mama). Als ik een rondje maak in mijn vriendenkring blijkt echter dat acht van de tien van hen een beroep uitoefent dat ten tijde van onze lagere schooltijd - het huidige basisonderwijs - nog niet bestond, althans niet in die benaming. Voorbeelden van deze nieuwe beroepen zijn accountmanager, softwareontwikkelaar, radio-engineer en lector. De beroepen die wel al in onze jeugd bestonden werken tegenwoordig met vernieuwde procedures en technieken, aangepast aan de eisen van deze tijd. Vaak zijn deze vernieuwde procedures en technieken gerelateerd aan nieuwe technologieën. Zo werken automonteurs met computerprogramma's om een probleem te detecteren voordat ze naar een sleutel grijpen en zorgt de keukenspecialist ervoor dat je al direct met een computersimulatie kunt zien hoe je nieuwe keuken er uit gaat zien. De innovaties in de medische wetenschap zijn revolutionair te noemen. Denk bijvoorbeeld aan de sleutelgatchirurgie, waarbij er maar een kleine incisie nodig is en patiënten sneller kunnen herstellen.

Career watchers - ook een beroep dat nog niet bestond toen ik tien was - voorspellen op basis van vergrijzing, economische groei, technologische en milieugerelateerde ontwikkelingen dat in de toekomst beroepen als socialmediamanager, applicatie-ontwikkelaar voor Smartphones, specialist cloud computing, energieadviseur, ingenieur elektrische auto's, expert zonne-energie en zorgmanager, belangrijk zullen worden[6]. Of wat dacht u van beroepen als koersmotivator, datapiloot of holistiek manager, beroepen die worden beschreven in de Trendrede, een initiatief van Nederlandse toekomstdenkers[7].

Een blik naar het verleden laat een oneindige lijst met uitgestorven beroepen zien. Zo weten - denk ik - nog maar weinig mensen wat een bombazijnmedewerker, een pluimgraaf, moertapper of builenvuller is. Kijkend naar het heden, zien we dat er veel beroepen zijn die in onze jeugd nog niet bestonden. Kijkend naar de toekomst, zullen vele nieuwe beroepen hun intrede doen. Ongeacht welke beroepen nog van belang zullen worden kunnen we stellen dat er definitief geen sprake meer is van stabiele beroepen met een vaststaand profiel van kennis en vaardigheden. Nu is er sprake van dynamische, complexe professies. Ik spreek bewust over professie, omdat professie meer betekent dan het uitoefenen van een beroep alleen. Diegene die een professie uitvoert wordt een professional genoemd. Op de kenmerken van een professional kom ik later in deze rede terug. Een complexe professie heeft drie kenmerken. Een *eerste* kenmerk verwijst naar de hoogwaardigheid van een specifieke deskundigheid. Het eigen maken van deze deskundigheid tot hoog niveau vergt veel tijd, training en volharding. De ontwikkeling van deskundigheid verloopt bovendien niet in een lineair proces, maar gaat met vallen en opstaan (Boshuizen & Van de Wiel, in voorbereiding). Een *tweede* kenmerk betreft de ingewikkelde mix van typen kennis en vaardigheden die voor het goed uitoefenen van de specifieke deskundigheid van belang zijn. Deze specifieke deskundigheid wordt steeds meer vertaald in zogenaamde BoKS (Bodies of Knowledge and Skills). De ontwikkeling en validering van landelijke kennisbases is hier een voorbeeld van en wordt gezien als een maatregel om tot een gemeenschappelijk beeld te komen van de domeinspecifieke beroepseisen. Het gaat dan niet alleen om vakinhoudelijke kennis of materiedeskundigheid maar ook over onderzoekende vaardigheden, informatievaardigheden, reflectieve vaardigheden, communicatieve vaardigheden en meningsvorming (Kessels, 1996). Vakinhoudelijke kennis wordt voor een deel gekenmerkt door een korte halfwaardetijd (Weggeman, 2000). De halfwaardetijd is de periode waarin kennis op de arbeidsmarkt nog maar de helft waard is van wat die kennis waard was op het moment dat iemand zijn opleiding voltooide. Een voorbeeld is een softwarepakket dat door een nieuwe versie of een ander

6 Zie http://www.careerbuilder.nl/blog/2011/10/10-beroepen-van-de-toekomst/
7 Zie www.trendrede.nl

pakket wordt vervangen. Voor een net afgestudeerd elektrotechnisch ingenieur bedroeg de halfwaardetijd in 1997 vijf jaar, tien jaar eerder was dat nog tien jaar (Den Hertog, 1997). Het is van belang zorgvuldig te bewaken welke kennis duurzaam is en welke afhankelijk is van toekomstige ontwikkelingen. Een *derde* kenmerk is dat binnen een professie taken moeten worden uitgevoerd onder voortdurend wisselende condities, waardoor nieuwe doelstellingen niet uitblijven. Deze wisselende condities vragen het flexibel aanwenden van de juiste kennis en vaardigheden. De wijze waarop deze kennis en vaardigheden worden aangewend verschilt per professie.

Voor professioneel beoordelen betekent de ontwikkeling van stabiele beroepen naar complexe professies dat kennis en vaardigheden die moeten worden beoordeeld niet altijd een lange houdbaarheidsdatum hebben. Een zorgvuldige analyse van huidige opleidingsprofielen, kennisbases en eindkwalificaties is nodig om goed te kunnen bepalen welke kennis en vaardigheden zinvol zijn om te beoordelen. Deze analyse leidt bij voorkeur niet tot lineaire lijsten van competenties, kwalificaties en indicatoren vervat in papieren tijgers, zij leidt bij voorkeur wel tot een visualisatie waarin de samenhang tussen de beoogde kennis en vaardigheden zichtbaar wordt. Een visualisatie heeft de eigenschap om zich aan te kunnen passen op het moment dat nieuwe kennis en vaardigheden voor handen zijn.

Van Ausbildung naar Bildung

Het feit dat het hbo niet meer opleidt tot stabiele beroepen maar tot complexe professies vraagt continue aanpassing van gewenste kennis en vaardigheden. In het hbo zijn deze gewenste kennis en vaardigheden in beroepsprofielen vastgelegd. Beroepsprofielen beschrijven op een redelijk toekomstbestendige manier de essentie en centrale activiteiten van het beroep. De profielen zijn gevalideerd door relevante betrokkenen zoals het werkveld en vakverenigingen (Boon & Van der Klink, 2000; Smeijsters & Sporken, 2004). Vanuit het beroepsprofiel wordt door hogescholen gezamenlijk een opleidingsprofiel vastgesteld waarin de kennis en vaardigheden voor de startende beroepsbeoefenaar worden beschreven in termen van te realiseren eindkwalificaties. In deze eindkwalificaties worden de domeinspecifieke kennis en vaardigheden genoemd, meestal beschreven in taken, rollen, gedragingen, handelingen, dimensies, resultaten en kennisbases (Van Berkel, 2012). Wanneer deze kennis en vaardigheden worden beheerst is sprake van *professioneel vakmanschap* (HBO-raad, 2009). Professioneel vakmanschap uit zich in professioneel gedrag en professioneel handelen. Ik licht deze begrippen later in deze rede toe.

Naast domeinspecifieke kwalificaties zijn vijf generieke eindkwalificaties vastgelegd in de zogenaamde Dublin descriptoren: kennis en inzicht, toepassen van kennis en inzicht, communicatie, oordeelsvorming en leervaardigheden. De descriptoren beschrijven in algemene termen de eindkwalificaties van de bachelor, de master en de doctor[8]. Afgeleide beschrijvingen van de Dublin descriptoren zijn de tien algemene hbo-competenties[9]. Ook de HBO-raad heeft een standaard van de professionele bachelor gedefinieerd. Deze standaard houdt in dat een opleiding tot hbo-bachelor ervoor zorg dient te dragen dat studenten een *gedegen theoretische basis* verkrijgen, dat zij het *onderzoekend vermogen* verwerven dat hen in staat stelt om bij te kunnen dragen aan de ontwikkeling van het beroep, dat zij over voldoende *professioneel vakmanschap* beschikken, en tenslotte de *beroepsethiek en maatschappelijke oriëntatie* ontwikkelen die past bij een verantwoordelijke professional. Vanzelfsprekend is de internationale dimensie onderdeel van elk onderdeel van deze standaard (HBO-raad, 2009). Zowel in de Dublin descriptoren als in de standaard van de professionele bachelor wordt het onderzoekend handelen benadrukt. Het wordt steeds belangrijker dat studenten op een systematische wijze leren, vragen uit de beroepspraktijk te beantwoorden en dat zij ook leren onderzoeksmatig te handelen. De invulling van onderzoek en de wijze waarop studenten dit leren verschilt per hogeschool.

Uit de beschrijving van eindkwalificaties blijkt dat er in het hbo enerzijds aandacht moet zijn voor het ontwikkelen van professioneel vakmanschap (in het Duits *Ausbildung* genoemd; de **b** van h**b**o) en anderzijds ook voldoende ruimte moet worden gecreëerd voor vaardigheden als onderzoek, kritisch leren denken en persoonlijke groei (de **h** van **h**bo). Een mooi woord waarin zowel de ontwikkeling van professioneel vakmanschap als het doen van onderzoek als vorm van professionele ontwikkeling samenkomt, vind ik *Bildung*. Dit begrip is geïntroduceerd door de Duitse geleerde en diplomaat Wilhelm von Humboldt (1767-1835). Bildung kan het best worden vertaald als *zelfontplooiing*, dus niet alleen de verwerving van algemene kennis maar ook het ontwikkelen van vermogens tot moreel oordelen en kritisch denken. Bildung kan bereikt worden door langdurige en intensieve studie. Opvallend is dat dezelfde professor Bruijn die in het rapport *Vreemde ogen dwingen* nogal instrumentele maatregelen adviseerde, in november 2012 in het rapport *Onderwijs: De derde dimensie* ook pleit voor hernieuwde aandacht voor Bildung. Het rapport is het resultaat van een werkgroep die zich boog over fundamentele vragen als: Voor welke uitdagingen staat het onderwijs in de toekomst? Waarmee dienen komende generaties uitgerust te worden zodat zij kunnen participeren in alle facetten van het leven? De werkgroep is daarbij geïnspireerd door grote klassieke en hedendaagse denkers. Centrale boodschap

8 Zie www.jointquality.org
9 Zie http://www.hboproofafstuderenenbegeleiden.nl

van het rapport is dat het Nederlandse onderwijs naast de twee bestaande doelen van kennisverwerving en het bijbrengen van vaardigheden ook een hoger doel dient dat niet is geënt op economische nastreefbare doelen (Bruijn & De Vries, 2012; Reijngoud, 2011). Dit doel kan worden omschreven als een veelomvattende vorm van Bildung, gecombineerd met *21st century skills*[10] en empathische aspecten. De nadruk op empathisch onderwijs dient als tegenwicht voor mogelijk nadelige effecten van technologische groei, zoals het verleren van allerlei sociale vaardigheden.

Samengevat zijn de doelen van het hbo in de loop der jaren verbreed en verrijkt. Dat betekent dat beoordelen geënt moet zijn op dit brede scala aan doelen. Om professioneel beoordelen te kunnen vormgeven is het noodzakelijk een visie te formuleren. Deze visie beschrijft hoe een hbo-opleiding mogelijkheden creëert voor het leren van professioneel vakmanschap, onderzoekend vermogen en zelfontwikkeling. Het heeft mijn voorkeur dat deze visie wordt vertaald in beelden, vignetten of scenario's omdat deze meer draagvlak en herkenning zullen oproepen bij relevante betrokkenen, zoals studenten, docenten en professionals in de beroepspraktijk.

Van schools curriculum naar een rijke leeromgeving

In de vorige paragrafen heb ik geconstateerd dat de doelen van het hbo breder en complexer zijn dan ooit tevoren en dat kennis en vaardigheden voor een bepaalde professie niet meer eenduidig vast te stellen zijn maar continu aan verandering en bijstelling onderhevig. Dit heeft implicaties voor de inrichting van de leeromgeving van studenten. Curricula kunnen niet meer worden ontworpen met een vast aanbod van onderwijsinhouden maar zullen voortdurend moeten worden bijgesteld door onderwijs dat past bij nieuwe eisen van de beroepspraktijk. Dit biedt kansen het onderwijs binnen het hbo te transformeren van gestandaardiseerde schoolse curricula naar rijke leeromgevingen. Deze leeromgevingen worden op de eerste plaats gekenmerkt door uitdagende taken die authentieke situaties uit de beroepspraktijk representeren en kunnen alleen met de juiste toepassing van kennis worden uitgevoerd. Ze zijn bij voorkeur mede ontworpen door het werkveld en worden gekenmerkt door een eigentijdse visie op de beroepspraktijk en vormen de ruggengraat van een programma van beoordelen (Sluijsmans, Straetmans, & Van Merriënboer, 2008; Straetmans, Sluijsmans, Bolhuis & Van Merriënboer, 2003). Co-constructie van kennis staat in deze taken centraal en ook is er aandacht voor een onderzoekende houding en reflectief vermogen. Een dergelijk programma licht ik toe bij bouwsteen 2. Naast uitdagende taken is

10 Deze zijn communicatie, samenwerken, ICT-geletterdheid, creativiteit, kritisch denken, probleemoplosvaardigheden, sociale en- culturele vaardigheden (inclusief burgerschap) (Voogt & Pareja Roblin, 2010).

er in een rijke leeromgeving sprake van een palet aan didactische werkvormen, zoals onderzoekend leren, samenwerkend leren en projectmatig werken. Deze manieren van leren kunnen steeds meer worden gefaciliteerd door de mogelijkheden van informatie- en communicatietechnologie (ICT), bijvoorbeeld social software als e-mail, blogging en chat (De Vries, 2007). Door de opkomst van ICT is het mogelijk om in andere omgevingen te leren en zijn studenten niet meer beperkt door de muren van de opleiding. Een mooi voorbeeld van een virtuele leeromgeving is GIMMICS. In dit virtuele spel, dat oorspronkelijk is uitgewerkt door de afdeling Farmacie van de Rijksuniversiteit van Groningen, wordt de dagelijkse praktijk van een apotheek gesimuleerd. In GIMMICS runnen studenten in teams van vijf à zes personen gedurende vier weken dagelijks hun eigen apotheek in het fictieve Pildorp. Zij voeren casussen uit waarin een beroep wordt gedaan op de kennis en vaardigheden van apotheker, bijvoorbeeld aan de patiënt goed uitleggen wat de bijwerkingen zijn van een bepaald medicijn. De casussen verschillen in moeilijkheid en ondersteuning, waardoor het mogelijk is deze aan te passen aan het niveau van de student. Wanneer casussen goed worden uitgevoerd krijgt de apotheek er virtuele patiënten bij, anders blijven de patiënten weg. De onderdelen van GIMMICS waarop de gameleiding een beoordeling geeft - en de criteria die daarbij worden gehanteerd - worden niet op voorhand aan de deelnemende teams bekend gemaakt. Deze wijze van beoordelen sluit naadloos aan bij de praktijk van alledag: is een patiënt/cliënt tevreden dan komt hij terug, is hij niet tevreden dan zal hij klagen of wegblijven. Beoordeling van het vak vindt plaats op basis van aanwezigheid (indien het verzuim tijdens de game meer dan twee dagdelen betreft volgt een vervangende opdracht), inzet tijdens GIMMICS en een reflectieverslag.

Rijke leeromgevingen maken het voor studenten mogelijk te leren in een variëteit van contexten. Ze leren niet meer uitsluitend door het doorlopen van een formeel, in het rooster vastgelegd curriculum met lessen en toetsweken, maar bewegen zich organisch tussen verschillende omgevingen als de werkplek, sociale netwerken, verenigingen en bijbanen. Dit vraagt ook om een andere manier van werken met elkaar: studenten dragen zelf meer verantwoordelijkheid in het kiezen van taken en de wijze waarop ze hun kennis en vaardigheden willen aantonen. Voor professioneel beoordelen betekent een rijke leeromgeving dat studenten zowel individueel als in samenwerking betekenisvolle taken uitvoeren in een variëteit van formele, informele en non-formele contexten. ICT kan hierbij een belangrijke rol vervullen[11].

11 Hoewel ik het gebruik van ICT bij beoordelen vanwege mijn gebrek aan expertise op dit gebied niet expliciet verder uiteenzet in deze rede, is het van belang bewust te zijn van deze ontwikkelingen, ook als het gaat om beoordelen. Judith van Hooijdonk, Marcel Schmitz, Chris Kuijpers en Chris Kockelkorn zijn collega's binnen Zuyd die zich met passie inzetten ICT meer te integreren in leren en beoordelen.

In dit hoofdstuk heb ik kort de historie van het hbo beschreven door drie ontwikkelingen te schetsen:

1) de ontwikkeling van stabiele beroepen naar dynamische, complexe professies;
2) de ontwikkeling van Ausbilding naar Bildung; en
3) de ontwikkeling van schoolse curricula naar rijke leeromgevingen.

Een vervolgvraag is hoe beoordelen kan bijdragen aan het behalen van de drie doelen van het hbo, namelijk professioneel vakmanschap, onderzoekend vermogen en zelfontwikkeling. Een professioneel programma van beoordelen is de tweede bouwsteen, die in het volgende hoofdstuk aan bod komt.

Kernboodschappen bouwsteen 1, de professie:

- Een zorgvuldige analyse van huidige opleidingsprofielen, kennisbases en eindkwalificaties is nodig om tot de juiste keuzes te komen welke kennis en vaardigheden zinvol zijn te beoordelen. Deze analyse leidt bij voorkeur tot een visualisatie waarin de samenhang tussen de beoogde kennis en vaardigheden zichtbaar wordt en die snel kan worden aangepast op het moment dat nieuwe kennis en vaardigheden voorhanden zijn.
- Om professioneel beoordelen te kunnen vormgeven is een visie nodig die beschrijft hoe een hbo-opleiding professioneel vakmanschap, onderzoekend leren en zelfontwikkeling faciliteert. Bij voorkeur wordt deze visie vertaald in beelden, vignetten of scenario's.
- Het is van belang dat de wijze van beoordelen aansluit bij de mogelijkheden die rijke leeromgevingen bieden. Voor professioneel beoordelen betekent dit dat studenten zowel individueel als in samenwerking met anderen hun kennis en vaardigheden ontwikkelen en aantonen door het uitvoeren van betekenisvolle taken in een variëteit van formele, informele en non-formele contexten. ICT kan hierbij een belangrijke rol vervullen.

Een programma van professioneel beoordelen

In het vorige hoofdstuk heb ik de professie benoemd als eerste bouwsteen voor professioneel beoordelen. Daarbij heb ik het ontwikkelen van professioneel vakmanschap, het stimuleren van onderzoekend vermogen en zelfontwikkeling als de drie belangrijkste doelen van het hbo benoemd. Om aan het einde van de opleiding te kunnen vaststellen of de student voldoende bekwaam is, is een aanzienlijke bewijslast nodig waarmee deze bekwaamheid kan worden aangetoond. In dit hoofdstuk zet ik uiteen wat nodig is om tot een onderbouwde uitspraak over hbo-bekwaamheid te komen door drie ontwikkelingen te schetsen:

1) de ontwikkeling van losse toetsen naar een samenhangend programma van beoordelen;
2) de ontwikkeling van beoordelen als sluitstuk van leren naar beoordelen als start voor leren; en
3) de ontwikkeling van convergent naar divergent beoordelen.

Het hoofdstuk wordt met samenvattende kernboodschappen afgesloten.

Van losse toetsen naar een programma van beoordelen

Om professioneel vakmanschap, onderzoekend vermogen en zelfontwikkeling van studenten te kunnen beoordelen is één keer beoordelen niet genoeg. Er moet sprake zijn van een *programma van beoordelen*: een zorgvuldig doordacht, evenwichtig programma van beoordelingsmomenten dat informatie oplevert over het kennen en kunnen van de student (Van der Vleuten et al., 2012). Het is van belang om bij het ontwerp van een programma van beoordelen rekening te houden met de expertiseontwikkeling van studenten. Vanuit cognitief-psychologisch perspectief wordt de ontwikkeling van studenten gekenmerkt door de overgang van een conceptueel rijke en rationele kennisbasis (ontwikkeld tijdens de opleiding) naar een non-analytische bekwaamheid om situaties in de beroepspraktijk te herkennen en te lijf te gaan. Uit onderzoek blijkt dat deze bekwaamheid afhankelijk is van specifieke situaties en niet als vanzelfsprekend transfereert naar andere situaties. Dus bekwaam handelen in de ene situatie blijkt een

slechte voorspeller van bekwaam handelen in een nieuwe situatie (Van der Vleuten et al., 2012). De consequenties voor een programma van beoordelen zijn tweeledig. Ten eerste zullen beoordelingen aan het begin van een opleiding op andere kennis en vaardigheden zijn gericht dan aan het einde van de opleiding. Ten tweede kan één beoordeling nooit voldoende zijn om bepaalde kennis of vaardigheden te beoordelen (Dijkstra, Van der Vleuten & Schuwirth, 2010) maar is er een groot arrangement aan beoordelingen nodig om tot een valide en betrouwbare uitspraak over eindkwalificaties te komen. Hierbij geldt dat het geheel meer is dan de som der delen. Deze beoordelingen zou ik *beoordelingstaken* willen noemen. Beoordelingstaken zijn professionele taken die representatief zijn voor de professie waar studenten actief aan werken, ook wel performance assessments genoemd (Govaerts, 2011). Een essentieel kenmerk van deze taken is dat ze de beoordelaar en de student zinvolle informatie opleveren over de mate van bekwaamheid en wat er nodig is deze bekwaamheid verder te ontwikkelen (Sluijsmans, 2008).

Beoordelingstaken met duidelijke beoordelingscriteria als basis kunnen verschillende vormen aannemen en worden uitgevoerd in veel gevarieerde situaties (Janssen-Noordman & Van Merriënboer, 2002). Beoordelingstaken worden afgeleid van beroeps- en opleidingsprofielen en kunnen betrekking hebben op een actueel probleem dat zich in een professie voordoet. Omdat deze taken de ruggengraat zijn van een programma van beoordelen en een afspiegeling zijn van verwachte eindkwalificaties zullen deze taken met de grootste zorgvuldigheid moeten worden ontworpen en gevalideerd in discussie met experts uit het werkveld (Straetmans et al., 2003). Omdat het in vage beschrijvingen van competenties en indicatoren aan context ontbreekt werkt het vaak goed beoordelingstaken te formuleren op basis van kritische beroepssituaties. Elke beoordelingstaak is gericht op één of meerdere gedragsdoelen. Deze doelen beschrijven het gewenste resultaat, de condities waaronder kennis en vaardigheden moeten worden gedemonstreerd en een set van beoordelingscriteria (Van Merriënboer & Kirschner, 2007). In lijn met de doelen van het hbo kunnen beoordelingscriteria betrekking hebben op professioneel vakmanschap (kan de student de kennis adequaat toepassen in een specifieke professionele situatie?), op onderzoekend vermogen (stelt de student zich nieuwsgierig op in de beroepspraktijk?) en op zelfontwikkeling (stelt de student zich onderzoekend naar zichzelf op?).

Kort door de bocht zijn er twee soorten beoordelingstaken. De eerste soort betreft op voorhand ontworpen beoordelingstaken waarmee studenten aantonen of ze de kennisbasis van een professie beheersen en deze kennis ook weten in te zetten in hun professionele handelen. Deze taken richten zich op het *weten*, het

weten hoe en *het tonen* (Miller, 1990). Voorbeelden van beoordelingstaken die zich hierop richten zijn casussen, werkveldopdrachten, simulaties of projecten. Deze taken kunnen verschillen in de mate van complexiteit en de mate waarin wordt verwacht dat de taken zelfstandig moeten worden uitgevoerd. Studenten kunnen individueel of in teams aan deze taken werken of samen met docenten en de beroepspraktijk. De tweede soort beoordelingstaken zijn de taken die worden uitgevoerd in de beroepspraktijk. Deze taken kunnen niet a priori worden ontworpen. Ze zijn onvoorspelbaar en kunnen van de student meer of minder zelfstandigheid, expertise en inzet vereisen. Meestal zijn dit de taken die worden uitgevoerd in de stage en die zich richten op het *doen* (Miller, 1990).

Voor beide soorten van beoordelingstaken moet de student beschikken over routinematige kennis en vaardigheden die een hoge mate van automatisering vereisen. Denk bijvoorbeeld aan het beheersen van de tafels om te kunnen rekenen, het kunnen typen om notulen te kunnen uitschrijven of het kennen van het periodiek systeem om scheikundige formules te kunnen ontleden. Het heeft geen nut deze kennis en vaardigheden apart te beoordelen op vooraf bepaalde momenten, omdat ze vooral moeten worden geleerd door veel 'drill en practice'. Omdat routinetaken te allen tijde automatisering tot doel hebben, oefenen studenten deze vooral zelfstandig of in het bijzijn van een docent zodat directe correctieve feedback kan worden gegeven.

De uitvoering van beoordelingstaken levert informatie op, bijvoorbeeld een verslag, een presentatie of een video. De informatie die door het uitvoeren van alle beoordelingstaken wordt verzameld, wordt gebruikt om te beoordelen en om goede, verantwoorde beslissingen te nemen. Dit kunnen kleine beslissingen zijn zoals het geven van een voldoende voor een opdracht of grote beslissingen zoals het verlenen van een diploma of getuigschrift. Afhankelijk van de zwaarte van de beslissing zijn de consequenties voor de student groter. Bij het niet halen van een vak zijn er meestal nog herkansingen, maar een onvoldoende voor een scriptie betekent geen diploma. Een stelregel is dat hoe meer er op het spel staat, hoe meer er in de kwaliteit en interpretatie van deze informatie moet worden geïnvesteerd (Van der Vleuten & Schuwirth, 2011). Dus om te bepalen of een student een hbo-diploma heeft verdiend zal veel informatie nodig zijn van veel beoordelingstaken. Ook is er een zorgvuldige analyse nodig om te kunnen bepalen welke beoordelingstaken essentieel zijn voor bepaalde beslissingen en welke van minder belang zijn. Opleidingen hebben een programma van beoordelen bij voorkeur niet in lijvige toetsplannen beschreven maar vertaald in een visualisatie (bijvoorbeeld diagrammen of grafieken) waarin de samenhang en verhouding tussen deze beoordelingstaken zichtbaar wordt en die als basis kan dienen voor interne en externe verantwoording. Het is van belang te voorkomen

dat een student in het vierde jaar nog niet alle credits uit de propedeuse heeft behaald. Hoe alle informatie moet worden geaggregeerd tot een betrouwbare beslissing over hbo-bekwaamheid is nog een spannende vraag en vereist meer onderzoek. Om resultaten van beoordelingstaken betekenisvol aan elkaar te verbinden is er een algoritme nodig. Een interessante aanzet hiertoe is de Protocol Portfolio Scoring[12].

Samengevat bestaat een programma van professioneel beoordelen uit een groot aantal beoordelingstaken die samen bewijslast leveren over hbo-bekwaamheid. Het opzetten van een dergelijk programma is niet zozeer een toetskundig vraagstuk maar vooral een ontwerpvraagstuk (Van der Vleuten & Schuwirth, 2005) waarbij een goede aansluiting tussen leren, onderwijs en beoordelen een essentieel uitgangspunt is (Biggs, 1996). Beoordelen heeft echter niet alleen de functie om bekwaamheid vast te stellen. Beoordelen moet vooral de student helpen verder te leren.

Van sluitstuk van leren naar start voor leren

Wanneer beoordelingen als geïsoleerde momenten in een curriculum worden ingebouwd zonder voldoende aansluiting met leren en instructie, heeft beoordelen vooral in de beleving van studenten de betekenis van afsluiten en de eindigheid van leren. Het leren houdt op als de voldoende binnen is. Professioneel beoordelen betekent echter dat beoordelen niet alleen bekwaamheid vaststelt, maar vooral handvatten aanreikt zodat studenten zich steeds verder kunnen ontwikkelen naar een onderzoekende professional. Een programma van beoordelen biedt kansen om beoordelen, naast certificering, *te verankeren in leren*. Omdat er in een programma van beoordelen beoordelingstaken worden uitgevoerd in een variatie aan representatieve beroepscontexten en -situaties is het mogelijk groei over taken heen vast te stellen (Straetmans, 2006; Straetmans, Sluijsmans, Bolhuis & Van Merriënboer, 2003; Sluijsmans, Straetmans & Van Merriënboer, 2008). Zo is het niet alleen mogelijk vast te stellen of een student bepaalde criteria in verschillende situaties beheerst, het zogenaamde *criteriumgericht* beoordelen, maar ook of een student groeit ten opzicht van een vorige taak, het zogenaamde *ipsatief* beoordelen (ipse=zelf in het Latijn; Sluijsmans, 2008). Bij een ipsatief beoordelingsmodel is het individuele leerpad van de student uitgangspunt voor de beoordeling.

Om resultaten van beoordelingstaken te kunnen gebruiken als een start van een leerproces in plaats van een afsluiting zijn een paar zaken belangrijk. Een eerste aandachtspunt is dat de beoordelingstaak de student confronteert met een

12 Zie voor een gedetailleerde beschrijving hiervan Sluijsmans et al., 2008; Straetmans et al., 2003.

authentieke taak die een situatie uit de professie representeert. De validiteit van een beoordelingstaak wordt vooral bepaald door datgene wat de taak uitlokt bij de student (de *stimulus*) en niet zozeer door de manier waarop het resultaat wordt beoordeeld (de *reponse*) (Van der Vleuten, 1996). Wanneer de taak authentiek is ervaart de student deze als nuttig en belangrijk en dat leidt weer tot een hogere motivatie (Segers, 2004). Een voorbeeld van een authentieke taak is de onderstaande simulatietaak uit de opleiding Fysiotherapie van Zuyd Hogeschool.

> *Stella Jongemans is een atlete van 22 jaar. Zij komt bij jou met de volgende hulpvraag: pijnvrij weer kunnen hardlopen. Dit is niet mogelijk omdat ze na twintig minuten trainen knieklachten krijgt rechts, die zich lokaliseren aan de ventrale zijde van de knie. 's Morgens als ze trap af moet lopen gaat dit ook met dezelfde pijn gepaard. Ze vertelt dat als ze lang heeft moeten zitten in de schoolbanken en weer moet opstaan, dit moeilijk gaat en wederom dezelfde stekende pijnklachten in de knie geeft. Het is de eerste keer dat ze deze knieklachten heeft. De knie wordt niet dik en warm en er is geen sprake van instabiliteitsklachten en geen slotverschijnselen en voor zover ze kan herinneren is er geen trauma geweest. De klachten zijn nu zes weken aanwezig en ze zijn sluimerend begonnen. Patiënte is goed gezond en slikt geen medicatie. Ze zit goed in haar vel en er zijn geen problemen in de werksfeer.*

In deze taak, die representatief is voor de beroepspraktijk, wordt de student aangesproken om zijn beschikbare kennis en vaardigheden in te zetten en waar nodig op te zoeken of te oefenen. Omdat de taak complex is wordt de student gestimuleerd de taak onderzoekend te benaderen. De taak kan zowel vanuit een criteriumgericht perspectief worden beoordeeld (voldoet de student aan de criteria?) als ook ipsatief (voert de student deze taak beter uit dan de vorige taak?). Omdat taken zoals geschetst in het voorbeeld een natuurgetrouwe afspiegeling moeten zijn van een taak uit de professie, is het van belang bij het ontwerp van taken kwaliteitschecks in te bouwen. Professionals uit de beroepspraktijk zijn bijvoorbeeld van belang bij het valideren van taken. Samengevat is het veel belangrijker te investeren in het ontwerp van goede taken dan in bijvoorbeeld beoordelingslijsten en checklists, hoewel we nu vaak het omgekeerde zien.

Een tweede aandachtspunt in het vormgeven van beoordelen als start voor leren is dat een programma van beoordelen al vroeg in de opleiding de student kansen biedt relevante bekwaamheden te oefenen. Het is zinvol te starten met uitdagende beoordelingstaken die gericht zijn op belangrijke vaardigheden, zoals onderzoekend handelen. Dit kan bijvoorbeeld door mini-versies van praktijkopdrachten te laten uitvoeren die representatief zijn voor de afstudeeropdracht (Kempen, Keizer & Van den Berg, 2012). Op deze manier wordt voorkomen dat studenten veel te laat in hun opleiding worden geconfronteerd met nieuwe

kennis en vaardigheden die veel oefening en feedback vereisen. Doel is niet dat studenten deze taken al perfect uitvoeren maar juist studenten de ruimte geven gecontroleerd fouten te maken. Het maken van fouten is noodzakelijk om te groeien en bekwaam te worden. Het is dus belangrijk om na de beoordelingstaak genoeg nieuwe taken aan te bieden om fouten of misstappen te kunnen bijstellen.

Een derde aandachtspunt is dat naast de gangbare psychometrische criteria als betrouwbaarheid en validiteit andere criteria nodig zijn om vast te stellen of beoordelingstaken aan de maat zijn (Schuwirth & Van der Vleuten, 2005). De laatste jaren is dan ook (inter)nationaal geïnvesteerd in deze uitbreiding, waarbij de effecten van beoordelen op het leren, de zogenaamde consequentiële validiteit, een prominentere plek krijgen. Drie voorbeelden hiervan zijn de twaalf kwaliteitscriteria van Baartman, Bastiaens, Kirschner en Van der Vleuten (2006), de zeven principes van Assessment 2020 (Boud and associates, 2010) en de tien kenmerken van duurzaam beoordelen (Sluijsmans, 2010). Tot slot wordt de leerwaarde van beoordelen alleen bevorderd door de student rijke informatie te bieden over waarom de taak al dan niet goed is uitgevoerd. Om de resultaten van een beoordelingstaak zinvol te laten zijn voor leren zijn analytische lijsten met ontelbare beoordelingscriteria niet afdoende en zelfs demotiverend voor zowel de beoordelaar als de student. Door deze fragmentatie is de essentie van een professie vaak niet meer zichtbaar en wordt fragmentarisch beoordelen alleen maar in de hand gewerkt (Korthagen, 2004). Bovendien wordt leerrijke informatie door deze lijsten verarmd door vinkjes en kruisjes. In een professioneel programma is geen sprake van dit tot één punt lopend, convergente beoordelen, maar wordt geïnvesteerd in divergent beoordelen dat voorziet in mogelijkheden om te leren.

Van convergent naar divergent beoordelen

Als ik leerkrachten op basisscholen spreek, zijn velen van hen niet zo geïnteresseerd in een resultaat van de Cito-toets. Het resultaat bevestigt het beeld dat zij al van een leerling hebben en als het resultaat afwijkt wijten ze deze afwijking aan de toets of het moment. De lange tijd die zij hebben gehad om hun leerlingen te observeren is voor hen van veel grotere waarde dan de toetscore. Toch zijn de gangbare manieren om studenten te informeren over resultaten van beoordelingstaken het geven van cijfers, dichotome kwalificaties (voldoende, onvoldoende) of andere informatiearme vormen, zoals afgevinkte criterialijsten. Onderzoek toont aan dat deze communicatievormen de zwakste vormen zijn van feedback en niet of nauwelijks aanzetten tot verder leren (Hattie & Timperley, 2007). Ze bieden ook weinig houvast voor het interpreteren van de waarde van een dergelijk resultaat. Zo kan een zes op de ene beoordelingstaak evenveel waard zijn als een acht in de andere beoordelingstaak. Vergelijk het met het krijgen van de uitslag van een bloedonderzoek. Als blijkt dat je cholesterol hoger is dan 8,0

mmol/l, weet je dat deze sterk verhoogd is maar nog niet a) waardoor dat komt en b) hoe je deze waarde kunt verlagen.

Professioneel beoordelen betekent dat resultaten niet meer worden uitgedrukt in cijfers maar worden gecommuniceerd op manieren die passen bij professioneel handelen. Dit betekent dat na de uitvoering van beoordelingstaken, maar ook voorafgaand aan de uitvoering van beoordelingstaken, veel aandacht is voor informatierijke leeractiviteiten waarin sprake is van intensieve interactie tussen beoordelaar en student. Resultaten van een recent uitgevoerde reviewstudie laten zien dat er acht methoden zijn die divergente informatie opleveren voor de student (Sluijsmans, Joosten & Van der Vleuten, 2013). Deze zijn het geven van feedback, het effectief vragen stellen, het organiseren van dialogen, reflectieve lessen, zelfbeoordelingen, peerbeoordelingen, het werken met beoordelingsrubrieken en het formatief gebruik van summatieve beoordelingen. Daarnaast is een aantal methoden gevonden dat niet op grote schaal beschreven is[13]. Bij sommigen is sprake van procesactiviteiten waarbij de docent en/of de student via handelingen het leerproces sturen, terwijl anderen meer direct ingezet kunnen worden als instrument, zoals de beoordelingsrubrieken. Om een instructieproces te doorlopen is een combinatie van methoden nodig. De genoemde methoden vinden plaats op een continuüm van informeel naar formeel en van on-the-fly naar ingebedde leeractiviteiten (Shavelson et al., 2008). Met on-the-fly en informeel leren worden die leeractiviteiten bedoeld die onverwacht en ongepland gebeuren, bijvoorbeeld als een docent tijdens het begeleiden van studenten ter plekke feedback geeft op wat hij hoort. Geplande interactie is bewust. De docent plant en zoekt naar manieren om de kloof tussen wat studenten al beheersen en waar ze naar toe gaan te ontdekken. Dit kan bijvoorbeeld door het formuleren van een aantal vragen. Deze vragen worden op een bepaald moment tijdens het onderwijs aan de studenten gesteld en door discussie krijgt de docent zicht op waar studenten staan en wat ze nog moeten leren. Ingebedde leeractiviteiten zijn 'klaar voor gebruik' en zijn informatief voor het identificeren van de kloof tussen huidige en gewenste leersituatie. Methoden voor divergent beoordelen maken het mogelijk om aan te sluiten bij verschillende instructiemomenten: Wat is de gewenste situatie? Waar staat de student nu? Hoe kan de student de gewenste situatie bereiken? (Black & Wiliam, 2009). Ze dragen bij aan het verhelderen, begrijpen en delen van doelen en criteria voor succes, het bewerkstelligen van effectieve groepsdiscussies en activiteiten die bewijs leveren voor leren, het activeren van studenten als belangrijke informatiebronnen voor elkaar en het activeren van studenten in het stimuleren van eigenaarschap over het eigen leren (Wiliam, 2011).

13 Het voert te ver deze methoden hier te beschrijven. Ik verwijs naar de reviewstudie voor een uitgebreide toelichting op deze methoden.

In dit hoofdstuk heb ik een professioneel programma van beoordelen gedefinieerd als een verzameling samenhangende beoordelingstaken dat bewijs genereert voor het behalen van hbo-bekwaamheid. Deze beoordelingstaken worden niet beschouwd als sluitstukken maar als startpunten voor leren. Een goed zicht op de gewenste eindkwalificaties (bouwsteen 1) en op het ontwerp van een professioneel programma van beoordelen (bouwsteen 2) zijn noodzakelijke - maar niet afdoende - condities om de kwaliteit van beoordelen te kunnen garanderen. De kwaliteit en effectiviteit van beoordelen voor leren wordt voor een aanzienlijk deel bepaald door de kwaliteit van de uitvoerders. In het volgende hoofdstuk wordt een derde bouwsteen van professioneel beoordelen gelegd: de beoordelaar.

Kernboodschappen bouwsteen 2,
een professioneel programma van beoordelen:

- Om vast te kunnen stellen of eindkwalificaties zijn gerealiseerd is een analyse van beoordelen op programmaniveau belangrijk, waarbij duidelijk moet worden hoe de verschillende beoordelingstaken samen bewijslast leveren voor een beslissing over hbo-bekwaamheid.
- Informatierijke feedback op beoordelingstaken biedt studenten kansen om verder te leren. Een docent heeft een arsenaal aan methoden tot zijn beschikking om beoordelingstaken in te zetten als startstuk voor leren.
- Divergent beoordelen betekent dat resultaten niet worden vertaald in cijfers maar in kwalitatieve en narratieve informatie.

De professionele beoordelaar

Een aantal van u, waarschijnlijk van ongeveer dezelfde leeftijd als ik, kent nog wel het programma *Klassewerk*, een kinderprogramma van de KRO dat werd uitgezonden van 1970 tot 1989 en de laatste elf jaar werd gepresenteerd door Hans van Willigenburg. Per aflevering gingen twee scholen met elkaar de strijd aan door op een aantal terreinen opdrachten uit te voeren. Eén van die opdrachten heette Muziekwerk. In dit onderdeel bespeelde één (of meerdere) leerling(en) een muziekinstrument samen met pianist Tonny Eyk, waarna door dezelfde Tonny Eyk een cijfer werd gegeven (doorgaans een 8). Buiten het feit dat ik goede herinneringen koester aan dit programma en dit mij weer kind laat voelen, is het mijn eerste herinnering van de rol van beoordelaar op tv. Vandaag de dag zijn er talloze programma's die experts uit een specifiek domein plaatsen in de rol van beoordelaar, bijvoorbeeld in programma's als Topchef, Cup Cake Cup, Maestro en So You Think You Can Dance. De kwaliteit van de beoordelaar wordt pas recent in het hbo erkend als belangrijke factor in het borgen van onderwijskwaliteit. Ook in bredere zin wordt onderschreven dat de kwaliteit van het onderwijs staat of valt met de kwaliteit van de docent en de waardering die er voor deze docent is (Commissie Veerman, 2011). In dit hoofdstuk beschrijf ik de rol van beoordelaar aan de hand van drie relevante ontwikkelingen:

1) de ontwikkeling van de vakdocent naar expertbeoordelaar;
2) de ontwikkeling van één objectieve beoordelaar naar meerdere subjectieve beoordelaars; en
3) de ontwikkeling van informatiearm beoordelen naar beoordelen als rijke kans voor leren.

Ook dit hoofdstuk sluit ik af met enkele kernboodschappen.

Van vakdocent naar expertbeoordelaar

In de laatste decennia zijn de verantwoordelijkheden van hbo-docenten sterk uitgebreid (Plake, 1993). Was de docent eerst nog vooral de expert op een bepaald vakgebied die studenten daarin onderwees, nu heeft de docent ook rollen als coach, mentor, stage- en scriptiebegeleider. Dit maakt dat de functie van docent

niet meer kan worden vervat in een aantal taken en bevoegdheden, maar dat sprake is van een portfolio van rollen (Schuiling, 2007). Een rol die recent expliciete aandacht krijgt is de rol van beoordelaar. Docenten zijn doorgaans niet in deze rol opgeleid en voelen zich daardoor lang niet altijd bekwaam in het ontwerpen, uitvoeren en analyseren van beoordelingen (Straetmans, 2006). Beoordelen kost veel tijd en wordt door docenten niet altijd als hun meest uitdagende taak ervaren (Brown, 2004; Gibbs & Simpson, 2004). Sommige docenten voelen zich onzeker wanneer zij moeten beoordelen, zeker als het gaat om belangrijke beoordelingstaken, zoals scripties. Knelpunten die docenten bijvoorbeeld noemen bij het beoordelen van afstudeeropdrachten zijn het ontbreken van duidelijke beoordelingscriteria, het zelf niet deskundig zijn in het doen van onderzoek en het ontbreken van praktijkervaring waardoor feedback geven op de situatie van studenten in de beroepspraktijk lastig wordt (Kempen, Keizer & Van den Berg, 2012).

Het is essentieel in de kwaliteit van de beoordelaar te investeren, omdat de kwaliteit van beoordelen voor een aanzienlijk deel afhangt van de kwaliteit van de beoordelaar (Van der Vleuten et al., 2010). In Nederland bestaat er nog geen gevalideerde kennisbasis voor de rol van beoordelaar in het hbo. Wel zijn er naar aanleiding van het advies van Bruijn (2012) initiatieven ontplooid tot het formuleren van een basiskwalificatie examinering (BKE) en nemen ook hogescholen het heft in handen te komen tot een profiel van de hbo-docent als beoordelaar. Om dit beeld te verkrijgen kunnen aanknopingspunten worden gevonden in bijvoorbeeld studies van Brookhart (2009) en Stiggins (2009). In een aantal landen is kennis en kunde over beoordelen al opgenomen in registratieprocedures (AFT, NCME & NEA, 1990; South African Qualification Authority, 2001). De NSW Institute of Teachers (2005) heeft een aantal professionele standaarden geformuleerd in de vorm van een rubriek. Hoe dan ook is beoordelen een complexe taak waar niet lichtzinnig over mag worden gedacht. Goed kunnen beoordelen is moeilijk en er is veel training en ervaring voor nodig om te groeien van onervaren beoordelaar naar expertbeoordelaar.

Op basis van een recente literatuurstudie die ik samen met Desirée Joosten, Sabine van Eldik en Linda Jakobs heb uitgevoerd in het kader van het SURF-project *Bewust en bekwaam toetsen* zou ik onderscheid willen maken in twee typen expertbeoordelaars. Ik noem ze voor het gemak *type I en type II expertbeoordelaars*. Type I expertbeoordelaars kenmerken zich in diepgaande, instrumentele en procedurele kennis over beoordelen. Het gaat hier om beoordelaars die kennis hebben van verschillende beoordelingsvormen, die weten hoe je goede vragen construeert, die begrijpen hoe je toetsmatrijzen opstelt en psychometrische analyses kunt uitvoeren om de betrouwbaarheid te bepalen. Deze kennis wordt

overwegend ingezet bij beoordelingen die voor een groot deel kunnen worden gestandaardiseerd. Deze beoordelaars zijn niet per definitie vakinhoudelijk deskundigen, maar kunnen bijvoorbeeld collega's ondersteunen in het ontwerp, de uitvoering en de analyse van beoordelingstaken.

Type II expertbeoordelaars zijn beoordelaars die een hoge mate van professioneel vakmanschap bezitten en in staat zijn complex professioneel gedrag van studenten te analyseren, te interpreteren en te beoordelen. Deze beoordelaars bezitten zowel vakinhoudelijke expertise als procesexpertise (Kempen, Keizer & Van den Berg, 2012). Deze beoordelaars zijn geen experts in instrumentele kennis over beoordelen zoals de type I expertbeoordelaar (zo zullen ze niet weten waar de *p*-waarde van een toetsitem voor staat), maar zij onderscheiden zich van de type I beoordelaar door het feit dat zij een doorleefd beeld hebben van wat in een bepaalde professie aan kennis en vaardigheden wordt verwacht en nodig is. Van Driel en Verloop (1998) omschrijven deze kennis van de professie als praktijkkennis: een geaccumuleerd en geïntegreerd geheel van persoonlijke kennis, opvattingen en waarden over de wijze van beroepsuitoefening. Een professional bouwt praktijkkennis op door persoonlijke en professionele ervaringen. Bennett (2011) en Bell & Cowie (2001) benadrukken tevens het belang van deze praktijkkennis van beoordelaars, aangezien deze kennis nodig is om de goede vragen te stellen, prestaties te analyseren, feedback te geven of instructie aan te passen. Voorbeelden van type II beoordelaars zijn specialisten uit de beroepspraktijk die jarenlange ervaring hebben in het werken met studenten. Voor het beoordelen van onderzoekend vermogen betekent dit dat deze beoordelaars beschikken over een hoge dosis ervaring in onderzoekend handelen in de professie. Type II expertbeoordelaars zijn de wijze uilen met het scherpzinnige oog. Onderzoek van Govaerts et al. (2011) toont aan dat deze expertbeoordelaars verschillen van onervaren beoordelaars in het feit dat zij in hun oordeelsvorming meer aandacht besteden aan situatiespecifieke en taakgerelateerde aspecten en meer in staat zijn tot een betekenisvolle interpretatie van de wijze waarop een student een beoordelingstaak uitvoert. Novieten beschrijven letterlijk geobserveerd gedrag. Bovendien beschikken expertbeoordelaars over verfijnde, goed gestructureerde mentale modellen en zijn zij in staat holistisch te oordelen, terwijl novieten vanuit een analytisch perspectief beoordelen (Govaerts, Schuwirth, Van der Vleuten & Muijtjens, 2011).

In het hbo staan complexe doelen als professioneel vakmanschap, onderzoekend vermogen en zelfontwikkeling centraal. Om te kunnen vaststellen of deze doelen worden behaald zijn vele beoordelingstaken nodig, georganiseerd in een programma van beoordelen. Om te komen tot goede beoordelingstaken zijn type I beoordelaars nodig die begrijpen welke stappen in een dergelijk ontwerpproces

moeten worden doorlopen en hoe resultaten moeten worden geanalyseerd (de eerste soort van taken, zie bouwsteen 2). Complexe beoordelingstaken in niet-gestandaardiseerde situaties (de tweede soort van taken, zie bouwsteen 2) vragen echter om bekwame beoordelaars die in staat gedrag te analyseren, te interpreteren en te vertalen naar een informatierijk oordeel. Dit voltrekt zich in een zowel cognitief en affectief proces in sociale interactie tussen beoordelaar en student. Ik zou ervoor willen pleiten beoordelaars in het hbo vooral te professionaliseren richting type II expertbeoordelaars en in navolging van de studies van Govaerts (2011) meer onderzoek te verrichten naar de wijze waarop deze beoordelaars het professioneel gedrag van studenten beoordelen.

Van één objectieve beoordelaar naar meerdere subjectieve beoordelaars

Vorig jaar hebben vast velen van u gekeken naar de Olympische Spelen die in Londen werden gehouden. Tijdens deze spelen werd bij een aantal sporten gebruik gemaakt van jury's, zoals bij het turnen - niemand zal snel de prestatie van Epke Zonderland vergeten - , ritmische gymnastiek en de dressuurproef. Bij de dressuurproef is sprake van een panel van zeven juryleden dat de proef beoordeelt op criteria als techniek en artistiek gehalte. De juryleden zijn allemaal erkend als deskundigen op het gebied van de dressuur, noem het type II beoordelaars. Elk jurylid beoordeelt een dressuurproef op deze criteria op een schaal van 1 tot 10. Vervolgens worden de individuele beoordelingen met elkaar vergeleken en is er discussie wanneer op de criteria meer dan vijf punten verschil is tussen beoordelaars. Dit voorbeeld uit de sport vind ik boeiend. Het zet aan tot nadenken over hoe beoordelingsprocedures uit andere disciplines waarin belangrijke beslissingen op het spel staan, kunnen worden vertaald naar het onderwijs. Analoog aan het voorbeeld is het in het hbo belangrijk bij beoordelingstaken die gericht zijn op complex professioneel gedrag ook meerdere beoordelaars in te zetten om tot een betrouwbaar informatierijk oordeel te komen. Doel is dan niet te komen tot objectiviteit, maar tot betrouwbaarheid en intersubjectiviteit (Schuwirth, 2007). Adriaan de Groot legde al in 1966 in zijn boek *Vijven en Zessen* haarfijn uit hoe feilbaar de beoordelaar is in het komen tot beslissingen. Door gebruik te maken van een pool van expertbeoordelaars, zoals dat bijvoorbeeld ook wordt gedaan bij de kunsten, wordt de kwaliteit van een oordeel geborgd (Baer & McKool, 2009).

Het prevaleren van meerdere beoordelaars boven één beoordelaar is belangrijk vanwege drie argumenten. Een eerste argument is dat, ondanks gedetailleerde beoordelingslijsten met uitgewerkte criteria, een beoordelaar zich schuldig maakt aan zogenaamde beoordelaarseffecten (Govaerts et al., 2012; Van der Vleuten et al., 2010). Voorbeelden van beoordelaarseffecten zijn het *halo-effect* (een beoordelaar geeft het werk van een vlijtige student onbewust een hogere

beoordeling dan het even goede werk van een minder vlijtige en oplettende student), het *significsch effect* (bij de beoordeling van een taak let de ene beoordelaar vooral op houding, terwijl de andere beoordelaar meer naar de inhoud en de structuur kijkt), het *volgorde-effect* (na een reeks zwakke resultaten wordt een middelmatige prestatie ten onrechte overgewaardeerd) en het *vermoeidheidseffect* (na de zoveelste beoordeling nemen vermoeidheid en irritatie toe en wordt het oordeel lager). Meerdere beoordelaars zorgen voor een evenwichtige verdeling in beoordelaarseffecten. Uitzonderingen blijven echter bestaan. Denk bijvoorbeeld aan de eerder aangehaalde dressuurproef op de Olympische Spelen, waar de Nederlandse amazone Adelinde Cornelissen in de dressuurfinale een prachtige foutloze kür aflegde, maar een mindere score kreeg dan de Britse thuisrijdster Charlotte Dujardin, die er met het goud vandoor ging. Het verwijt aan de jury (door Nederlandse fans uiteraard) was chauvinisme. Een tweede argument voor het inzetten van meerdere beoordelaars is dat het de reikwijdte aan expertise vergroot. Elke beoordelaar, afkomstig uit de opleiding of de beroepspraktijk, brengt eigen kennis en ervaringen in waar de studenten zinvolle informatie uit kunnen halen (Delandshere & Petrovsky, 1998). Zo kan bij het beoordelen van een gesprek tussen hulpverlener en cliënt de ene beoordelaar goede feedback bieden over interpersoonlijke zaken, terwijl een andere beoordelaar meer feedback kan geven op vakinhoudelijke aspecten van de taak. Een derde argument is dat het inzetten van meerdere beoordelaars holistisch en kwalitatief oordelen mogelijk maakt. Beoordelaars hebben doorgaans een hekel aan gedetailleerde beoordelingsformats en checklists omdat deze niet hun expertise representeren en hun vrijheid in beoordelen ontnemen (Van der Vleuten, 1996). Wanneer de beoordelingstaak helder is gedefinieerd en de kaders waarbinnen wordt beoordeeld duidelijk zijn, bijvoorbeeld een simulatie, zijn checklists alleen maar zinvol als er procedurele handelingen moeten worden afgetikt (heeft de student bijvoorbeeld de cliënt een hand gegeven?). Voor deze aspecten van taakuitvoering is echter geen expertbeoordelaar nodig. Recent onderzoek van Govaerts et al. (2012) toont echter aan dat rijke verbale feedback van expertbeoordelaars tijdens taakuitvoering niet automatisch ook leidt tot rijke feedback als deze schriftelijk moet worden beschreven. Dit zou ervoor kunnen pleiten dat we beoordelaars verbale feedback laten geven en niet dwingen tot het invullen van allerlei formulieren.

Van informatiearm beoordelen naar informatierijk beoordelen

Als ik u de keuze zou voorleggen óf te worden beoordeeld door een onervaren beoordelaar met een analytische beoordelingslijst óf te worden beoordeeld door een ervaren beoordelaar met een blanco papier, zouden de meesten van u waarschijnlijk kiezen voor de tweede. We vinden het belangrijk door iemand te worden beoordeeld die we bekwaam en intelligent vinden (Vonk, 2009). In de vorige paragrafen heb ik uiteengezet wat de kenmerken zijn van een expertbe-

oordelaar en waarom het zinvol is meerdere beoordelaars in te zetten bij belangrijke beoordelingstaken. Door divergente methoden in te zetten wordt beoordelen een continu proces en houdt het nooit op zolang studenten en beoordelaars - in de hoedanigheid van docenten, experts uit de beroepspraktijk en medestudenten - in dialoog aan het leren zijn (Ruiz-Primo, 2011).

Informatiearme oordelen waar studenten niet veel meer mee kunnen dan vaststellen of ze hun credits hebben behaald passen niet bij professioneel beoordelen. Professioneel beoordelen gaat om het creëren van kansen om te leren. Om te illustreren wat ik hiermee bedoel laat ik u graag kennis maken met de masterclasses van Thomas Hampson[14]. Thomas Hampson is een Amerikaanse baritonzanger die regelmatig masterclasses verzorgd voor studenten aan het conservatorium. Deze masterclasses vind ik fascinerend om verschillende redenen. Je ziet een expert aan het werk die niet alleen beschikt over veel expertise over in dit geval de inhoud van het muziekstuk en de wijze waarop het stuk moet worden gezongen, maar ook over beoordelaarsexpertise. Hij is in staat de student duidelijk te maken wat er gebeurt en hoe dat anders kan door het zelf voor te doen, voorbeelden te noemen of fysieke aanwijzingen te geven, bijvoorbeeld over de beweging of de zanghouding. Hampson is in staat om relevante informatie te benoemen en te integreren tot betekenisvolle feedback voor de student. Dit zijn kenmerken van een expertbeoordelaar (Govaerts, Van de Wiel & Van der Vleuten, 2012). Een andere observatie is dat er in de masterclass geen sprake is van papieren gedetailleerde beoordelingslijsten maar van continue, verbale feedback waarbij direct wordt gecheckt of deze door de student wordt begrepen. Bovenal zie je hier een gemotiveerde student die geniet van het leren van de expert. Maar er is ook een proces van vallen en opstaan zichtbaar, iets dat we ook nog kennen toen we leerden fietsen en zwemmen. Je ziet de student feedback incasseren om vervolgens om te zetten naar aangepast handelen.

Het voorbeeld van de masterclass heb ik gebruikt om te illustreren hoe leren en beoordelen kunnen samensmelten en hoe kwaliteit van dit proces wordt geborgd door de beoordelaar. Eerder heb ik al een aantal methoden genoemd waarin leren en beoordelen worden vervlochten. Vaak zijn docenten in het hbo zich niet bewust dat deze methoden ook informatie opleveren over de kennis en vaardigheden van studenten. De wijze waarop docenten vorm geven aan de rol van beoordelaar wordt in hoge mate beïnvloed door de percepties en overtuigingen die zij hebben over leren en beoordelen (Birenbaum, 2011; Brookhart et al., 2008; Mui So & Hoi Lee, 2011; Sadler, 1989; Tierney, 2006). Zo zal een docent met de perceptie dat toetsing is bedoeld voor certificering deze perceptie ook laten

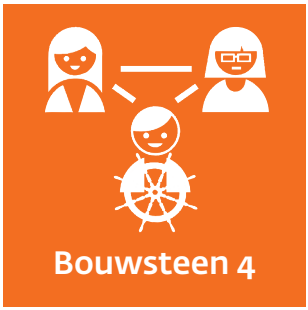
14 Zie http://www.pickstaiger.org/video/thomas-hampson-vocal-master-class-excerpts

doorwerken in de invulling van rol van beoordelaar, door vooral met beoordelen bezig te zijn aan het einde van een onderwijsperiode. Percepties en overtuigingen over leren en beoordelen worden vaak niet gearticuleerd en blijven verborgen in het hoofd van de docent (*tacit knowledge*; Sadler, 1989). Uit de studie van Tierney (2006) blijkt dat het veranderen van de rol van beoordelaar ook een verandering in concepties vereist. Een onderzoekende houding naar de eigen percepties en overtuigingen en het uitwisselen met collega's uit de opleiding en de beroepspraktijk blijkt voorwaardelijk voor een structurele verandering in de wijze waarop een docent de rol van beoordelaar inkleurt (Sluijsmans et al., 2013).

In dit hoofdstuk heb ik de beoordelaar beschreven als derde bouwsteen van professioneel beoordelen. De kwaliteit van beoordelen kan niet worden gegarandeerd als er in de opleiding niet voldoende wordt geïnvesteerd in bekwame beoordelaars. Professionalisering dient niet alleen gericht te zijn op de instrumentele kant van beoordelen, maar vooral op het gedrag van de beoordelaar. Door de wijze van beoordelen komt het professioneel gedrag van de beoordelaar tot uiting. Omdat professioneel beoordelen plaatsvindt in continue interactie tussen beoordelaars en student, heeft ook de student een belangrijke rol. De student benoem ik in het volgende hoofdstuk als de vierde bouwsteen van professioneel beoordelen.

Kernboodschappen bouwsteen 3,
de professionele beoordelaar:

- Professioneel beoordelen vraagt een goede mix van type I en type II expertbeoordelaars. Zij hebben kennis van het technisch proces van beoordelen maar ook een doorleefd beeld van wat in een bepaalde professie aan kennis en vaardigheden wordt verwacht en nodig is.
- Objectief beoordelen bestaat niet, betrouwbaar beoordelen wel. Om betrouwbaarheid van beoordelen te garanderen zijn veel subjectieve, holistische oordelen nodig van verschillende beoordelaars. Deze beoordelaars kunnen docenten zijn, werkplekbegeleiders, experts of studenten.
- Expertbeoordelaars beoordelen niet alleen op vaststaande momenten met vormen als cijfers, maar beoordelen continu door studenten te voorzien van feedback, door dialogen met ze aan te gaan, door ze vragen te stellen en ze te betrekken bij discussies over beoordelingstaken. Door deze activiteiten zijn beoordelaars vooral gericht op verder leren en het behouden van motivatie.



De professionele student

Als je studenten zou bevragen hoe zij tegen beoordelen aan kijken, zullen weinig van hen zeggen dat zij er dolenthousiast van worden. De meeste studenten zullen beoordelen associëren met cijfers en presteren. Dat is jammer, omdat beoordelen veel meer voor hun leren zou kunnen opleveren. In dit hoofdstuk onderbouw ik dat de student een belangrijke rol vervult in professioneel beoordelen. Drie ontwikkelingen zijn op zijn plaats:

1) de student die zich ontwikkelt van kennisverwerver naar kenniswerker;
2) de student die zich ontwikkelt van scholier naar professional; en
3) de student die zich ontwikkelt van iemand die afwacht wat hij of zij moet doen naar iemand die zelf verantwoordelijkheid neemt voor het eigen leren.

Ik besluit het hoofdstuk met enkele kernboodschappen.

Van informatieverwerver naar kenniswerker

Om een bekwame professional te worden moet de student zich de kennisbasis eigen maken die voor het uitvoeren van professionele taken nodig is. Een softwareontwikkelaar moet kennis hebben van programmeertalen, een laborant moet het scheikundig periodiek systeem tot in de finesse beheersen en een leerkracht moet kennis hebben van de ontwikkelingsfasen van het jonge kind. Het belang van kennis wordt de laatste jaren weer meer benadrukt als gevolg van de kritiek dat onderwijsvernieuwingen het belang van kennis zouden hebben verwaarloosd (Onderwijsraad, 2007). Een kennisbasis is een systematische beschrijving van bewezen kennis uit de professionele praktijk. Het is een web van samenhangende kenniselementen dat betekenis krijgt door deze kenniselementen aan elkaar te verbinden en te integreren (Verharen, Van Bommel & Vosselman, 2011). Er is de laatste jaren veel energie gestoken in het formuleren en valideren van kennisbases, bijvoorbeeld in de domeinen educatie en social work. Kennisbases vormen in het hbo de basis voor het ontwerp van onderwijs en (landelijke, digitale) kennistoetsen. Een voorbeeld is het project *10voordeleraar*[15], waar op basis van de kennisbases van de lerarenopleidingen kennistoetsen

15 Zie www.10voordeleraar.nl

worden ontwikkeld. Redactieteams werken in dit project intensief samen aan het vullen van een toetsbank met honderden kwalitatief goede toetsvragen, op basis waarvan toetsen kunnen worden samengesteld. Deze toetsen worden vanaf het volgend studiejaar digitaal afgenomen in het derde jaar van de opleiding en zullen zwaarwegend zijn, bijvoorbeeld in de beslissing of een student kan worden toegelaten tot de afstudeerfase. Hoewel ik met het ontwerpen van dergelijke toetsen op zichzelf geen probleem heb, ben ik wel kritisch over de wijze waarop deze toetsen worden geïmplementeerd. Ik zie vaak dat kennistoetsen worden aangeboden in een toetslokaal, waar overleg tussen studenten uit den boze is. Studenten hebben vooral voor dit moment hard gestudeerd en persen dan alle informatie die ze hebben opgeslagen er weer uit. Het blijkt dat beoordelingen die op deze wijze worden georganiseerd strategieën als reproduceren stimuleren (Crooks, 1988; Scouller, 1998; Scouller & Prosser, 1994). Deze reproductiegerichte aanpak, waarin de student zich voornamelijk richt op het in het hoofd stampen van informatie, leidt nooit tot beklijving van deze kennis in de context van een professionele taak. In deze wijze van beoordelen schuilt het gevaar van het bekende toetsgedrag (*wash-back effect*; Alderson & Wall, 1993). De boeken verdwijnen in de kast als de voldoende binnen is. Als studenten kennis tijdelijk opslaan maar deze niet eigen maken, wordt schijnsucces gemeten (Meherns, Popham & Ryan, 1998).

In lijn met de doelen van het hbo – professioneel vakmanschap, onderzoekend vermogen en zelfontwikkeling – zou ik toe willen werken naar een visie waarin kennis niet wordt beschouwd als het verwerven van informatie die te reproduceren is op daarvoor vastgestelde momenten, maar naar het aanleren van vaardigheden met kennis te werken in het licht van de toekomstige professie. Een dergelijke visie zou ook beter passen bij deze tijd. Het proberen kennis geheim te houden in gesloten examenenveloppen en beveiligde toetsbanken staat naar mijn gevoel haaks op de wereld waarin studenten leven. Studenten hebben dankzij hun laptops en telefoons tegenwoordig 24 uur per dag, 7 dagen in de week, toegang tot het wereldwijde web. We neigen in paniek te raken als studenten een digitale toetsbank weten te kraken of worden boos als een Cito-toets op straat komt te liggen. De vraag is of we dat steeds willen voorkomen of accepteren dat dat steeds meer zal gebeuren (overigens vind ik het kunnen kraken van een toetsbank een mooi bewijs van creatief vermogen). Het lijkt mij belangrijker dat studenten leren de informatie waar zij continu toegang toe hebben op waarde te schatten door deze zorgvuldig te analyseren, te beoordelen en te relateren aan het oplossen van een specifieke professionele taak. Dat zij, gezien de eerder genoemde korte halfwaardetijd van kennis, leren hoe zij hun kennis up-to-date kunnen houden.

Omdat informatie overal en altijd beschikbaar is en ook gevoelig is voor verandering en aanpassing gezien de snelle ontwikkelingen, wordt het belangrijker studenten op te leiden naar *kenniswerkers*. Een kenniswerker is iemand die effectief kan omgaan met informatie, leervermogen heeft, initiatief neemt, goed contact kan leggen met betrokkenen in de beroepspraktijk, veel autonomie nodig heeft om te leren en te werken, intrinsiek gemotiveerd is, naar aanvullende kennis zoekt als deze nodig is en nieuwsgierig is naar nieuwe ontwikkelingen (Van Rooij, Verdonschot, Dewulf & Smits, 2008). Kenniswerkers laten de volgende attitudes zien (Mijland, 2011): een levenlang willen leren, leren in dialoog, het ontwikkelen van een T-profiel (zowel breed oriënteren als specialiseren), delen van kennis, werken als een onderzoeker, durf om te kiezen en het nastreven van een passie. Er kunnen vervolgens drie typen kenniswerkers worden onderscheiden (Van Rooij et al., 2008): de *probleemoplosser c.q. verbeteraar* (de student die praktijkvragen weet op te lossen of te verbeteren), de *innovator* (de student die in het laboratorium of werkplaats nieuwe ideeën ontwikkelt die de professie tot vooruitgang gaan helpen) of de *specialist* (de student met veel kennis op een specifiek domein).

Voor professioneel beoordelen betekent het opleiden van kenniswerkers dat er in een programma van beoordelen niet zozeer beoordelingstaken worden aangeboden in de vorm van kennistoetsen gericht op reproductie, maar dat er beoordelingstaken worden aangeboden die gericht zijn op het aantonen van het kunnen werken met kennis. Voorbeelden van dergelijke taken zijn:

- het maken van conceptmaps van kennisgebieden en relaties tussen deze kennisgebieden;
- het aan anderen kunnen uitleggen hoe zij nieuwe kennis kunnen koppelen aan hun voorkennis;
- het aan anderen kunnen uitleggen welke kennis nodig is voor de analyse van een praktijkvraag;
- het door voorbeelden kunnen uitleggen waarom bepaalde kennis is verouderd en waarom deze niet meer past in de huidige praktijk.

Van scholier naar onderzoekende professional

Elk jaar stromen ongeveer dertigduizend mbo'ers, veertigduizend havisten en tienduizend vwo'ers het hbo in. Deze jonge mensen, vaak nog geen twintig jaar oud, zijn met bepaalde concepties van leren en beoordelen opgevoed en betreden het hbo met deze concepties. Hardnekkige beelden zijn dat beoordelen losstaat van het onderwijs en beoordelen iets is dat de docent doet door het geven van cijfers. Studenten in het hbo zullen al snel merken dat er andere zaken van ze wordt verwacht dan in hun eerder genoten schooltijd. Het gaat nu om het werken aan professionele taken, vaak in samenwerking met hun medestudenten,

waarbij een hoge mate van zelfstandigheid wordt verwacht. Bij het uitvoeren van die taken wordt een beroep gedaan op een toenemend onderzoekend vermogen, door taken voor te leggen die probleemanalyse en een specifieke toepassing van de professionele kennisbasis vereisen om tot een adequate oplossing te komen. De studenten ontwikkelen zich in het hbo - en ik heb het dan vooral over de voltijdstudenten - van scholieren naar jonge onderzoekende professionals. Professionaliteit is een woord dat de laatste jaren steeds meer te horen is. We zeggen wel eens: 'Dat vind ik niet professioneel' of 'Het ontbreekt deze persoon aan professionaliteit'. Meestal herkennen we snel het ontbreken van professionaliteit. Recente voorbeelden zijn het veelvuldig gebruik van doping door wielrenners (met Lance Armstrong als grootste desillusie), de bankiers met hun dikke bonussen en Ernst Jansen Steur, oud-neuroloog van Medisch Spectrum Twente, die jarenlang foute diagnoses heeft gesteld. Wat is een onderzoekende professional dan wel?

Een eerste kenmerk van een onderzoekende professional is het kunnen tonen van *professioneel gedrag* en *professioneel handelen* (Maister, 2000). Onder professioneel gedrag wordt observeerbaar gedrag verstaan waarin de *normen en waarden* en de juiste inzet van de kennisbasis van de professie zichtbaar zijn. Professioneel gedrag komt tot uitdrukking in woord, gedrag en uiterlijk. De mate waarin professioneel gedrag kan worden getoond is in hoge mate afhankelijk van de complexiteit en de condities van de situatie waarin de student zich bevindt (Crossley et al., 2002). Een exacte definitie van professional is afhankelijk van de professie, omdat in elk domein andere waarden, normen en gedrag wordt gewenst (Aguilar, Stupans & Scutter, 2011). Onder professioneel gedrag van een arts valt bijvoorbeeld de toepassing van de humaan-biologische kennis, de gedragswetenschappelijke kennis, de medisch-technische kennis en de sociaal-wetenschappelijke aspecten van de inrichting van het gezondheidszorgstelsel die dienen als basis voor de studie geneeskunde. Onder medisch- technische kennis vallen ook de klinische vaardigheden als anamnesetechniek en alle aspecten van de fysische diagnostiek en zaken als mondelinge en schriftelijke uitdrukkingsvaardigheden (Projectteam Consilium Abeundi, 2002). Professioneel handelen is handelen dat gebaseerd is op een systematisch geordende, overdraagbare kennisbasis waarbij de normen en waarden van de beroepsgroep tot uiting komen (Lunenberg, Dengerink & Korthagen, 2013).

Een tweede kenmerk van een onderzoekende professional is een voortdurende *nieuwsgierigheid* naar nieuwe kennis over de professie en de *motivatie* kennis en vaardigheden verder eigen te maken, samen met andere professionals. Dit uit zich in de wijze waarop een student vragen stelt over een bepaalde praktijkvraag, de methode die hij kiest om de praktijkvraag nader te onderzoeken en het

vermogen oplossingen van betekenis te laten zijn voor de beroepspraktijk. Verder is het kenmerkend dat een professional voor een aanzienlijk deel zelf bepaalt hoe het werk inhoudelijk het best kan worden ingevuld, waarbij nieuwe problemen en nieuwe situaties niet uit de weg worden gegaan.

Tot slot is een professional in staat *zichzelf te sturen* en zich te blijven ontwikkelen, ook na het afsluiten van een formeel opleidingstraject. In de laatste paragraaf van dit hoofdstuk ga ik dieper in op de ontwikkeling van studenten naar zelfsturende professionals en de consequenties voor beoordelen.

Een mooi voorbeeld van een onderzoekende professional en kenniswerker is Freek van Delft. Freek heeft in 2012 bij Zuyd Hogeschool de scriptieprijs gewonnen. Deze prijs wordt jaarlijks uitgereikt aan de Zuyd-student die op een originele wijze een vraag uit de beroepspraktijk heeft onderzocht. Freek heeft in samenwerking met het Delftse bedrijf MediShield en de Technische Universiteit Delft met zijn onderzoek een belangrijke bijdrage geleverd aan de ontwikkeling van de Seton-tang. Met een Seton-tang kun je twee uiteinden van een holle slang naadloos aan elkaar smelten. Het ringetje dat daardoor ontstaat kan gebruikt worden om na een fisteloperatie de wond open te houden. Nu gebruiken chirurgen nog draadjes die ze aan elkaar knopen. Dat leidt tot veel ongemak, ontstekingen en complicaties. Met de ringetjes genezen mensen sneller. Patent voor de tang is aangevraagd. Ik vind Freek een onderzoekende professional en een innovatieve kenniswerker omdat hij een positieve, structurele bijdrage heeft weten te leveren aan een technische innovatie in de medische zorg. Dat heeft hij in grote mate van zelfstandigheid gedaan en daarbij heeft hij professioneel gehandeld in termen van goede communicatie met betrokkenen en rekening gehouden met de regels die in het vakgebied gelden. Ook heeft hij gedurfd buiten de veilige omgeving van de opleiding met een interdisciplinair team te werken.

Voor professioneel beoordelen betekent de ontwikkeling van scholier tot onderzoekende professional dat er in de beoordeling voldoende aandacht is voor aspecten die een onderzoekende professional kenmerken. De beoordeling van professionaliteit vindt bij voorkeur niet plaats door het afvinken van attitudeaspecten - daarmee zou je het concept professioneel gedrag tekort doen - maar vindt plaats in een samenwerking tussen studenten en professionals uit de beroepspraktijk, tussen docenten en studenten onderling. Met andere woorden, als je wilt dat studenten zich professioneel gedragen, behandel ze dan ook als professionals (Maister, 2000). Belangrijke vormen om te bewijzen dat een student zich gedraagt als een onderzoekende professional zijn onder andere het laten presenteren van casussen waarin de student bijvoorbeeld demonstreert hoe hij een bepaald probleem systematisch heeft aangepakt of het organiseren

van 360° feedback (Brett & Atwater, 2001). Ook het voeren van functionerings- en beoordelingsgesprekken zijn geschikte vormen van beoordelen. Het is van belang dat studenten niet zozeer worden beoordeeld met methoden die zij kennen uit hun schoolse leven maar vooral al kennismaken met methoden die ook in hun professie worden gebruikt.

Van 'zeg me wat ik moet doen' naar 'ik weet wat ik wil doen'

De bekende pedagoog Martinus Langeveld betoogde al in de jaren dertig en veertig van de vorige eeuw dat het doel van opvoeding is (thuis en op school) dat de student uiteindelijk op een zelfverantwoordelijke manier beslissingen kan nemen, zowel ten aanzien van zichzelf als ten aanzien van de maatschappij. Opvoeding is gebaseerd op de fundamentele hulpeloosheid van het kind. De rol van de opvoeder is die van het plaatsvervangend geweten zolang het kind nog niet zelf in staat is verantwoordelijkheid te dragen voor zijn handelingen en doet dat gewetensvol, liefdevol, onbaatzuchtig en oprecht. Ultieme doel van de opvoeding is een kind tot *zelfverantwoordelijke zelfbepaling* te laten komen. Ook Kees Boeke, die in 1926 de Werkplaats Kindergemeenschap oprichtte, was een pionier als het gaat over de verantwoordelijkheid van het kind voor het eigen leren.

Beoordelingen in het hbo zijn op dit moment in een aantal opleidingen nog overwegend prescriptief van aard. De opleiding bepaalt wat, hoe en wanneer wordt beoordeeld - het programma van beoordelen staat vast - en heeft daarmee de volledige controle over het onderwijsleerproces van de student. Zoals we weten stuurt de wijze van beoordelen vaak het studeergedrag van studenten (Frederiksen, 1984; Natriello, 1987). Een extern gestuurde wijze van beoordelen leidt tot een extrinsiek gestuurde wijze van leren en daarmee logischerwijs tot demotivatie. Beoordelen dat wordt benaderd als iets dat los staat van onderwijs belemmert studenten in hun *flow*. Flow - als belangrijk concept in de positieve psychologie - wordt gekenmerkt door totale focus, volledige betrokkenheid en succesvolle uitvoering (Csiksentmihalyi, 1990), en kan ook wat weg hebben van spelend leren zoals dat in diverse publicaties wordt gezien als belangrijke conditie voor leren (De Vries, 2007; Martens, 2007). Onderzoek laat zien dat vooral de beoordelingen die gericht zijn op cijfers negatief zijn voor leren omdat ze studenten belemmeren om verantwoordelijkheid te nemen zich te motiveren en te reguleren (Hattie & Timperley, 2007; Ryan & Deci, 2000). Beoordelen op deze wijze leidt tot *aangeleerde afhankelijkheid* (Yorke, 2003): studenten leren dat de docent bepaalt wat wordt verwacht en dat zij dat niet zelf hoeven te doen.

Erkennen dat beoordelen de motivatie en zelfontwikkeling van studenten in hoge mate beïnvloedt is cruciaal voor professioneel beoordelen (Black & Wiliam, 1998).

Dit betekent dat studenten ook door de wijze van beoordelen moeten worden gestimuleerd zelf verantwoordelijkheid voor hun leren te nemen. De laatste jaren zijn vele initiatieven ondernomen om studenten meer zeggenschap te geven door hen bijvoorbeeld zelfbeoordelingen en reflecties te laten uitvoeren. In veel opleidingen zie je echter dat dit bij studenten kan leiden tot reflectiemoeheid en afhankelijk gedrag. Reflecteren wordt een routine en studenten schrijven op wat de docent in het reflectieverslag terug wil zien (Sluijsmans, Brand-Gruwel, Van Merriënboer & Bastiaens, 2003). Onderzoek laat zien dat er nog maar weinig empirisch bewijs is dat aantoont dat studenten een goede beoordeling van zichzelf kunnen geven (Regehr & Eva, 2006; Ward, Gruppen & Regehr, 2002). Als het gaat om zelfbeoordeling overschatten studenten meestal hun prestaties. Zo denkt bijvoorbeeld ook ieder van ons dat we beter autorijden dan de gemiddelde Nederlander, hoewel dit statistisch gezien nooit zou kunnen kloppen (dit is bekend als het *better than average effect*; Myers, 1980). Dat studenten hun kennen en kunnen overschatten heeft te maken met een gebrek aan kennis over de beoordelingscriteria (Bjork, 1999; Dunning, 2005; Krueger & Dunning, 1999). Hoe kun je immers jezelf beoordelen op kennis waar je de reikwijdte niet van weet? Krueger en Dunning (1999) verwoorden dit als *unskilled and unaware of it*: niet alleen presteren studenten niet goed, ze missen bovendien de vaardigheden zich dit te realiseren. Dit is de reden dat novieten niet snel in paniek raken. Dit in tegenstelling tot de sterke studenten die veel meer besef hebben van wat er nog allemaal geleerd kan worden en zichzelf als gevolg daarvan onderschatten (Sluijsmans, 2002; Sluijsmans, Dochy & Moerkerke, 1999). Een andere reden waarom zelfbeoordelingen niet zo betrouwbaar zijn, is dat gedrag maar voor een klein deel bewust wordt aangestuurd. Studenten laten opschrijven of vertellen hoe zij zelf vinden dat ze het doen biedt alleen zicht op de processen waar zij zich bewust van zijn. Dat geeft een te onvolledig en dus onbetrouwbaar beeld.

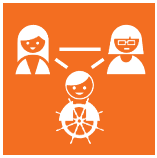
Hoe kunnen studenten dan wel worden betrokken bij beoordelen dat leidt tot zelfverantwoordelijkheid en een minder afhankelijke houding? Omdat zelfbeoordeling lang niet altijd tot accuraat bewijs leidt of een bewijs is van zelfkennis, zou ik een pleidooi willen voeren om zelfbeoordelingen sterk te reduceren (of zelfs uit een programma van beoordelen te bannen) en vooral *kijk-naar-buiten* activiteiten te organiseren onder het motto: wil je weten hoe je het doet, vraag het een ander! (Vonk, 2009). Een professionele student neemt verantwoordelijkheid in het actief zoeken naar feedback in plaats van te wachten op feedback van iemand van wie hij het misschien niet eens wilt (Boud, 2000; Kremer-Hayon & Tillema, 1999; Van Merriënboer & Sluijsmans, 2008). Onderzoek laat bijvoorbeeld positieve effecten zien van peer beoordeling op motivatie en leren, zeker wanneer er voldoende ondersteuning wordt geboden bij het leren feedback geven en ontvangen (Sluijsmans, 2002; Sluijsmans, Brand-Gruwel, Van Merriënboer & Martens,

2004; Sluijsmans, Brand-Gruwel & Van Merriënboer, 2002). Ook blijkt dat studenten graag feedback ontvangen van kritische beoordelaars omdat deze doorgaans als intelligent en deskundig worden gezien (Vonk, 2009). Een tweede manier is studenten keuzevrijheid te geven door te bepalen wanneer ze waarop willen worden beoordeeld. Dit betekent dat studenten niet alleen zelf beoordelingstaken zorgvuldig plannen en uitvoeren, maar ook bepalen op welke criteria zij feedback willen. Plannen en feedback opzoeken worden als de belangrijkste componenten in zelfsturend leren beschouwd (Butler & Winne, 1995; Ertmer & Newby, 1996; Kremer-Hayon & Tillema, 1999; Schunk, 1990; Van Merriënboer & Sluijsmans, 2008; Zimmerman, 1990). Controle over het eigen leren heeft in een aantal onderzoeken in het sportdomein al zijn meerwaarde bewezen (Wulf & Toole, 1999; Chiviacowsky & Wulf, 2002; 2005; Wulf et al., 2005). Een derde manier om studenten meer verantwoordelijkheid te geven is hen te motiveren een optimale leeromgeving voor zichzelf en hun medestudenten te creëren. Een voorbeeld is het samenstellen van zelfsturende studieteams waarin studenten relevante beoordelingstaken selecteren en bestuderen vanuit een reflectieve en samenwerkende onderzoekshouding (Kremer-Hayon & Tillema, 1999). Dit leidt vervolgens tot een hogere motivatie en tot meer leren dan wanneer het beoordelen extern wordt aangestuurd (Wulf, Raupach & Pfeiffer, 2005). De voorbeelden om studenten te betrekken bij beoordelen sluiten goed aan bij motivatietheorieën zoals de *zelfdeterminatietheorie* (Ryan & Deci, 2000; Pink, 2009), waarin competentie, autonomie en sociale verbondenheid de basisbehoeften zijn. Competentie betekent het gevoel van zinvol bezig zijn, autonomie verwijst naar de afwezigheid van voortdurende externe sturing en sociale verbondenheid doelt op het creëren van een vertrouwde leeromgeving.

In dit hoofdstuk heb ik uitgelegd wat een professionele student is en hoe studenten actiever kunnen participeren in beoordelen. Omdat de student de enige is die zichzelf meeneemt en zichzelf voortdurend kan zien, moet hij in staat zijn leren te reguleren. Met andere woorden: hij moet leren slim te leren. Ook zonder de veiligheid van de opleiding en de docenten moet de student in staat zijn verder te leren. Voor studenten betekent dit dat zij niet langer meer achteroverleunend het onderwijs consumeren maar zich veel meer opstellen als *prosumers* die mede vormgeven aan het leren en beoordelen door op eigen initiatief informatie over hun prestaties op te zoeken (Martens, 2007).

Kernboodschappen bouwsteen 4,
de professionele student:

- De professionele student in het hbo borgt zijn kennisbasis niet meer door het reproduceren van kennis op een voor hem vastgesteld moment, maar bewijst zich als kennis-werker.
- De professionele student in het hbo gedraagt zich professioneel, is nieuwsgierig, heeft een goed onderzoe-kend vermogen en beschikt over de vaardigheden zichzelf te sturen, zowel in de opleidingscontext als in de beroeps-praktijk.
- De professionele student in het hbo leert steeds meer verantwoordelijkheid te nemen voor zijn eigen leren in termen van het plannen, uitvoeren en evalueren van beroepstaken en het opzoeken van feedback. Autonomie, competentie en doelgerichtheid zijn belangrijke voor-waarden voor behoud van motivatie.



Een professionele beoordelingscultuur

In de vorige hoofdstukken zijn respectievelijk de professie, een programma van professioneel beoordelen, de beoordelaar en de student beschreven als belangrijke bouwstenen voor professioneel beoordelen. Ik heb geprobeerd duidelijk te maken dat een programma van beoordelen zodanig moet zijn samengesteld dat er enerzijds voldoende bewijs moet worden verzameld om de hbo-doelen - professioneel vakmanschap, onderzoekend vermogen en zelfontwikkeling - te kunnen behalen en anderzijds er veel kansen voor leren moeten worden gecreëerd. De beoordelaar en de student zijn de belangrijkste actoren in dit programma. Door het voeren van rijke dialogen zal het leren van de student worden versterkt met als gevolg dat de student als onderzoekende professional het hbo verlaat en zich gemotiveerd verder bekwaamt in de professie. Het voorafgaande kan echter alleen maar worden gerealiseerd als er sprake is van een leer- en werkomgeving die het mogelijk maakt om op deze manier beoordelen gestalte te geven.

Als vijfde en laatste bouwsteen geef ik invulling aan het begrip professionele beoordelingscultuur: de sociale, professionele omgeving, inclusief de daar geldende visie en kernwaarden, waarbinnen professioneel beoordelen gestalte krijgt. Kenmerken van een dergelijke cultuur zijn het bieden van een rijke leeromgeving waarin studenten, docenten en de beroepspraktijk samenwerken aan uitdagende beoordelingstaken, het kunnen inzetten van expertbeoordelaars en studenten ruimte bieden hun eigen leren mede vorm te geven. Psychologische veiligheid, openheid, transparantie, eerlijkheid en professionele integriteit zijn kernwaarden van een professionele beoordelingscultuur die is gericht op het ontwikkelen van zelfstandige, gemotiveerde professionals (Van Gennip, 2012). Een nadere definitie van een professionele beoordelingscultuur geef ik door ook hier drie ontwikkelingen te schetsen:

1) van individueel naar collectief leren;
2) van een afhankelijke opstelling naar het nemen van regie; en
3) van ad-hockwaliteit naar duurzame kwaliteit van beoordelen.

Ik sluit het hoofdstuk ook hier af met enkele kernachtige boodschappen.

Van individueel leren naar collectief leren

Het belang van goed beoordelen is evident. Goede beoordelingen zetten studenten aan tot verder leren en zullen de juiste informatie opleveren om tot beslissingen te komen over bekwaamheid. Het perfecte programma van beoordelen bestaat echter niet, net zoals perfecte beoordelaars of perfecte studenten niet bestaan. Er zullen door gebrek aan tijd, geld of de juiste kennis en vaardigheden compromissen moeten worden gesloten in de manieren van beoordelen. Als het gaat om het verhogen van kennis en vaardigheden over beoordelen zien we de laatste jaren dat in het hbo docenten en betrokkenen uit de beroepspraktijk veel ruimte wordt geboden zich te professionaliseren in beoordelen. Er zijn studiedagen en conferenties over beoordelen, er zijn binnen hogescholen expertgroepen die opleidingen adviseren en ook toetsbureaus schieten als paddenstoelen uit de grond. Deze wijzen van professionalisering zullen echter te weinig effectief en efficiënt zijn om tot een gezonde beoordelingscultuur te komen omdat deze vormen van professionalisering meer gericht zijn op *individueel* gedrag dan op het verbeteren van samenwerking en kennisdeling op *collectief* niveau (Crow & Pounder, 2000; Fishman, Best & Tal, 2003). Zolang docenten op individueel niveau werken aan hun kennisontwikkeling over beoordelen wordt te weinig geïnvesteerd in een gemeenschappelijk begrip over de functie van beoordelen op het niveau van de opleidingen en hogescholen. Zeker gezien de impact die beoordelen heeft op het leren van studenten is een gemeenschappelijk gedragen visie op beoordelen essentieel (Klimoski & Mohammed, 1994).

Om tot een gemeenschappelijk gedragen visie op beoordelen te komen is het van belang om in plaats van alleen individuele activiteiten te stimuleren ook te investeren in een professionele leergemeenschap (PL). Birenbaum et al. (2011) definiëren een PL als een '*professional community where teachers work collaboratively to reflect on their practice, examine evidence about the relationship between practice and student outcomes, and make changes that improve teaching and learning for the particular students in their classes.*' (p. 36). Opleidingen die werken als een PL reguleren hun eigen leren en voelen zich minder afhankelijk van druk van buiten, waardoor zij in staat zijn continue te verbeteren. Ze zijn lerende organisaties (Senge, 1990). In een PL is sprake van een gedeelde visie op leren en beoordelen, waarbij ruimte wordt genomen om de overtuigingen en percepties over leren en beoordelen met elkaar te delen. Andere kenmerken van een PL zijn een duidelijke cultuur in termen van waarden en normen, een goede infrastructuur gericht op ondersteuning voor onderwijsevaluatie, kennismanagement en professionele ontwikkeling, motivatie om doelen te behalen en professioneel leren in termen van betekenisvolle interacties en onderzoekend leren (Birenbaum et al., 2011). De resultaten van Birenbaum et al. laten zien dat een PL - de context waarin professioneel beoordelen wordt ingebed - ingericht wordt op basis van oprechte betrokkenheid van

studenten bij beoordelen en een continue dialoog tussen docent en student. Maar ook een continue dialoog tussen docenten en management leidt tot meer vertrouwen, respect en meerwaarde voor leren en de kwaliteit van onderwijs. De beoordelingscultuur weerspiegelt de opleidingscultuur.

Ik noem graag twee voorbeelden die ik inspirerend vind als het gaat om het vormgeven van een professionele leergemeenschap. Het eerste voorbeeld betreft het Finse *VERME* project[16]. In dit peergroupmentoringproject komen leerkrachten met verschillende expertiseniveaus bij elkaar om hun ervaringen te delen over professionele taken die ze in de praktijk tegenkomen. Het kan gaan om ordeproblemen in de klas, maar ook over het beoordelen van leerlingen. Kenmerken van een omgeving waarin leerkrachten van elkaar leren zijn psychologische veiligheid en respect voor elkaars standpunten. Ze delen en leren van elkaar door afwisseling van formele en informele momenten waardoor professionele ontwikkeling mogelijk is (Kwakman & Van den Berg, 2004). Een tweede voorbeeld is het initiatief *Operatie in Beeld* van het Oogziekenhuis in Rotterdam (De Korne, Hiddema, Bleeker & Klazinga, 2009). Dit ziekenhuis legt sinds 2008 operaties vast op video waarna een team ze analyseert om te zien of veiligheidsprocedures worden uitgevoerd, bijvoorbeeld de *time out procedure* (een laatste gezamenlijke controle van juistheid van patiënt, oog, operatie en benodigd materiaal voorafgaand aan de operatie). Ook wordt geëvalueerd hoe de teamleden onderling communiceren en in hoeverre de afgesproken checks worden uitgevoerd. Operatie in Beeld is een vorm van professionele peer beoordeling, iets dat we wel al veel zien bij studenten (Sluijsmans, 2002) maar dat nog weinig plaatsvindt op teamniveau tussen docenten (Van der Klink, 2011).

De twee voorbeelden waarin peer feedback een belangrijk element is kunnen ook in de context van het hbo binnen opleidingen worden georganiseerd, bijvoorbeeld onder begeleiding van expertisenetwerken en lectoraten binnen de hogeschool. De feedback die docenten over hun wijze van beoordelen ontvangen beperkt zich nu doorgaans tot de informatie uit de studentenenquêtes. Deze informatie biedt echter te weinig handvatten voor een kritische reflectie op het eigen handelen (Van der Klink, 2011). In een vergelijkbare opzet als de peergroup mentoring kunnen docenten elkaar voorzien van feedback op het ontwerpen van beoordelingstaken of de wijze waarop ze narratieve feedback geven aan hun studenten, maar ook het werkveld en studenten vragen om feedback. Er kan ook aandacht worden besteed aan fundamentelere zaken, zoals de vraag waarom beoordelen belangrijk is voor studenten. Om tot een duurzame verandering te komen zijn de sessies niet eenmalig maar terugkerend. Het is mogelijk om just-in-time vast te stellen aan welke kennis behoefte is en op welke manier deze kennis

16 Zie http://www.youtube.com/watch?v=glWVX7veux4 voor een impressie.

het beste kan worden ontwikkeld. Het organiseren van peer feedback tussen hbo-docenten kan worden gezien als een eigenschap van een professionele beoordelingscultuur (Raemdonck & Strijbos, 2013). De uitkomsten van deze sessie kunnen worden ontsloten voor anderen door gebruik te maken van een virtueel platform.

Van een afhankelijke opstelling naar het nemen van regie

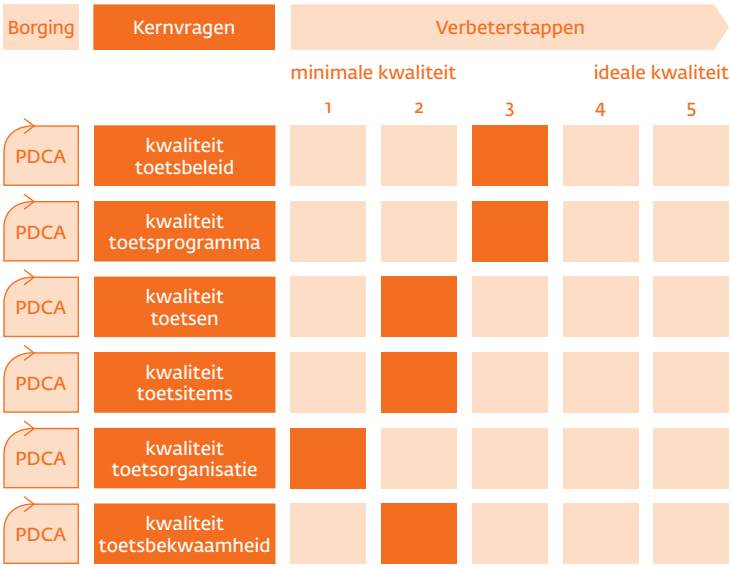
De laatste jaren heb ik veel opleidingen bezocht. Bij de opleidingen die een accreditatie voor de deur hebben staan heerst altijd wel een bepaalde spanning. De sfeer die ik dan proef heeft wat weg van een pittig examen: men is zich goed aan het voorbereiden en oefent panelgesprekken om op het moment suprême zo goed mogelijk voor de dag te komen. Enerzijds snap ik die sfeer goed; een accreditatie is het moment waarop een opleiding laat zien wat de kwaliteit is. Examencommissies moeten zich bij de accreditatie kunnen verantwoorden over de kwaliteit van beoordelen (Joosten & Sluijsmans, 2012; Smits, 2011). Opleidingen die onvoldoende scoren op toetsing en gerealiseerde eindkwalificaties (standaard 3) kunnen een herstelperiode van een jaar toegewezen krijgen. Anderzijds schrik ik van de hectiek die een accreditatie oproept. Naar mijn idee is deze hectiek niet bevorderlijk voor de rust binnen een opleiding. Het is geen goede zaak als kwaliteitsborging teveel wordt geassocieerd met eisen, procedures en formats van een externe actor (Mertens, 2011).

Het is belangrijk dat opleidingen zich minder afhankelijk opstellen richting een externe instantie als de NVAO, maar zich uitgedaagd voelen regie te nemen in het ontwikkelen van een sterk kwaliteitszorgsysteem ten aanzien van beoordelen. In een professionele beoordelingscultuur is sprake van het vermogen verantwoording af te leggen over de wijze waarop kwaliteit van beoordelen wordt geborgd. Analoog aan de HBO-raad (2009) definieer ik kwaliteitszorg als het systematische en dynamische geheel van beleid, organisatie, procedures, processen en middelen, gericht op het permanente handhaven en verhogen van de kwaliteit van beoordelen binnen de hogeschool door middel van cyclisch evalueren en verbeteren. De NVAO moedigt het aan dat opleidingen regie nemen. Ook accreditatiepanels maken zich schuldig aan beoordelaarseffecten en subjectieve oordelen. Des te belangrijker is het vanuit de opleiding een stevig verhaal te hebben over kwaliteit en kwaliteitsborging. De accreditatie-eisen zoals vastgesteld door de NVAO (2010) bieden een goede toetssteen, maar zullen pas betekenis krijgen in de context van de professie. Voor beoordelen betekent dit dat een opleiding zich moet kunnen verantwoorden over vragen als: Hebben we als opleiding een duidelijke visie op beoordelen die in lijn is met onze visie op de professie? Streven we als opleiding nog de juiste doelen na met onze studenten? Leren we hen nog de juiste dingen? Zetten we als opleiding de juiste middelen,

kennis en vaardigheden in om tot goede beoordelingen te komen die leerwaarde hebben en leiden tot goede beslissingen over hbo-bekwaamheid? Zorgen we ervoor dat niet alleen docenten zich bekwaam voelen - en zijn - in beoordelen, maar stimuleren we ook dat studenten steeds meer verantwoordelijkheid nemen in leren en beoordelen? Weten we als opleiding waar onze kracht ligt en waar niet? Hoe zorgen we dat we de kwaliteit die we niet in huis hebben toch borgen? Zorgen we voor voldoende peer review, leren we als hogescholen van elkaar en wisselen we kennis en ervaringen uit? Durven we onze eigen beoordelingspraktijk onder de loep te nemen, als individuele docent, maar ook als team?

Om regie te kunnen nemen en bovenstaande vragen te kunnen beantwoorden zijn handvatten nodig die opleidingen kunnen gebruiken om tot een kwaliteitszorgsysteem van beoordelen te komen. In de recente publicatie *Kwaliteit van toetsing onder de loep* wordt een aantal van deze handvatten besproken (Sluijsmans, Peeters, Jakobs & Weijzen, 2012). In deze publicatie beschrijven we een methodiek waarmee betrokkenen in het hbo (management, examencommissies, ontwikkelaars, docenten, studenten en werkveld) de kwaliteit van beoordelen kunnen analyseren, verbeteren en borgen. Op basis van wetenschappelijke en praktische bevindingen over kwaliteit en analysemodellen voor kwaliteit hebben wij gewerkt aan een prototype van een methodiek om tot kwaliteit te komen. Basis voor de methodiek is de kwaliteitspiramide van eigentijds toetsen en beoordelen die oorspronkelijk is ontwikkeld door Joosten-ten Brinke (2011) en aangepast is door Jaspers en Van Zijl (2011). Ik verwijs naar het artikel voor een uitgebreide operationalisatie hiervan. De methodiek biedt vervolgens handvatten om kwaliteit van beoordelen te analyseren op zes niveaus (zie figuur 1: kwaliteit van toetsbekwaamheid, toetsorganisatie, toetsitems, toetsen, toetsprogramma en toetsbeleid) en te onderzoeken hoe het met de kwaliteit is gesteld en wat daarin de beoogde verbeterslagen zijn. We hebben hiervoor het ontwerp van het auditinstrument van de Stichting Duurzaam Onderwijs (2010) gebruikt. Het auditinstrument van SDO heeft de volgende kenmerken die ook voor de door ons ontwikkelde methodiek zinvol zijn: (1) door zelfevaluatie komen opleidingen tot een positie- en ambitiebepaling, (2) er wordt uitgegaan van vijf fasen van ontwikkeling, (3) er wordt gebruik gemaakt van één set van criteria die alle betrokkenen hanteren en (4) er is een relatie met de Plan-Do-Check-Act-cyclus van Deming. Het instrument is gebaseerd op een model voor kwaliteitsmanagement, ontwikkeld door de European Foundation for Quality Management (EFQM), en uitgebreid door het Instituut Nederlandse Kwaliteit (INK) en het EFQM-INK-model (Oakland, 1993). Startpunt voor dat model is dat een organisatie als een hbo-opleiding zich ten aanzien van een aantal criteria kan bevinden in een bepaalde ontwikkelingsfase. Voor de positie- en ambitiebepaling van onze methodiek gebruiken we vijf fasen, waarbij fase 1 uitgaat van een 'minimale

kwaliteit van toetsing' en fase 5 van een 'ideale kwaliteit van toetsing' (zie figuur 1).



Figuur 1. Kwaliteitsborging van beoordelen door positie- en ambitiebepaling (Sluijsmans et al., 2012, p. 2).

Maar wat betekent minimale kwaliteit en ideale kwaliteit als het gaat om beoordelen? Hiervoor is een nadere definitie van kwaliteit nodig.

Van ad-hockwaliteit naar duurzame kwaliteit

Als ik twee portretten van koningin Beatrix zou laten zien, één gemaakt door een amateur en één gemaakt door een professioneel kunstenaar, zou u waarschijnlijk het portret van de professionele kunstenaar aanwijzen als zijnde van hogere kwaliteit. Het is niet toevallig dat de kwaliteit van de professional hoger ligt; hier gaat jarenlange expertiseontwikkeling aan vooraf en veel vallen en opstaan. Analoog aan het voorbeeld van de portretten is het een complexe klus om tot een hoge kwaliteit van beoordelen te komen. Dit proces kost veel tijd en begeleiding. De aandacht voor kwaliteit van beoordelen in het hbo is er nog maar net, maar mijn observatie is dat opleidingen deze kwaliteit snel op orde willen brengen uit vrees voor zwaarwegende consequenties. Het frustreert me nogal eens wanneer de indruk wordt gewekt dat kwaliteit van beoordelen snel te repareren is, bijvoorbeeld door vlug een beoordelingsformulier aan te passen voor de accredi-

tatie. Kwaliteit van beoordelen vergt veel meer dan een instrumentele, activiteitengestuurde ad-hocaanpak. Er is zorgvuldige aandacht nodig voor het vraagstuk over hoe kwaliteitsproblemen met beoordelen kunnen worden opgelost (Kempen, Keizer & Van den Berg, 2012). Kwaliteit van beoordelen is iets dat tussen de oren van alle betrokkenen moet zitten en dit is een proces van jaren.

Een eerste belangrijke stap is betrokkenen bewust te maken van het belang van goed beoordelen en hoe percepties op leren en beoordelen de kwaliteit van beoordelen bepalen. Dit kan alleen wanneer er sprake is van een onderzoekende houding naar de eigen beoordelingspraktijk. In het project *De toetsing getoetst* hopen we die bewustwording te stimuleren door betrokkenen in gesprek te laten gaan over de vraag waarin goede kwaliteit van beoordelen verschilt van minder goede kwaliteit. We hebben hiervoor rubrieken ontworpen waarin per rij uit figuur 1 in vijf fasen (de vijf blokken per rij) wordt aangegeven wat minimale tot ideale kwaliteit is. In het volgende licht ik dit toe voor het niveau toetsen.

In de *eerste fase* (minimale kwaliteit; zie figuur 1) is er sprake van ad-hocbeleid dat *activiteitengericht* is. Voor toetsitems (zie figuur 1) betekent dit dat de toetsen ad hoc worden ontworpen, meestal vlak voor de toetsafname. De toetsen worden ontworpen door een individuele docent. Als de toets in de ogen van ontwerpers goed is dan vertrouwt men op de kwaliteit ervan. De toets wordt incidenteel geëvalueerd door de docent zelf.

In de *tweede fase* van kwaliteit is sprake van *procesgerichtheid*. Voor het niveau van toetsen betekent dit dat er operationele afspraken zijn over toetsontwikkeling. De toetsen worden ontworpen door een groep docenten (die samen verantwoordelijk zijn voor de onderwijseenheid/beroepstaak). De toetsen zijn inhoudelijk afgeleid van de wensen van een groep docenten die betrokken is bij de betreffende onderwijseenheid/beroepstaak en de toetsen worden voorgelegd aan deskundigen binnen de opleiding (bijvoorbeeld expertisegroep toetsing) en van feedback voorzien. De toetsen worden geëvalueerd als onderdeel van kwaliteitszorg: doelen vaststellen, toets ontwerpen, afnemen, evalueren en verbeteren. Er wordt een vergelijking gemaakt tussen de toetsresultaten en de gestelde doelen.

In de *derde fase* van kwaliteit is sprake van *systeemgerichtheid*. Er zijn afspraken over toetsontwerp en -ontwikkeling die voortkomen uit de visie op toetsen en het toetsbeleid van de opleiding. De toets wordt ontworpen door (representanten van) de opleiding. Op opleidingsniveau zijn kaders en werkprocessen ontwikkeld voor toetsontwerp en toetsontwikkeling. De toetsen zijn inhoudelijk afgeleid van de eindkwalificaties van de opleiding. De opleiding heeft kwaliteitsstandaarden ontwikkeld waarmee de kwaliteit van de toets door diverse

betrokkenen binnen de opleiding vastgesteld wordt en toetsen worden geëvalueerd als onderdeel van de kwaliteitscyclus. De toetsen worden meerdere malen geëvalueerd en er worden trendanalyses gemaakt.

In de *vierde* fase van kwaliteit is sprake van *ketengerichtheid*. Er wordt actief gecommuniceerd over de aanpak van het toetsontwerp en -ontwikkeling in relatie tot de visie op toetsen en het toetsbeleid van de opleiding. De werkprocessen worden erkend en gebruikt bij het ontwerp en de ontwikkeling van toetsen. Het werkveld wordt betrokken bij het ontwerpen en ontwikkelen van toetsen. De kwaliteit van toetsen wordt vastgesteld op basis van studentevaluaties, evaluaties van de examencommissie/toetscommissie en evaluaties van het werkveld. De toets wordt systematisch geëvalueerd samen met 'klanten', zoals studenten en het werkveld, als onderdeel van de kwaliteitscyclus. De toetsresultaten worden vergeleken met die van collega-opleidingen.

Tot slot is in de *vijfde fase* (ideale kwaliteit; zie figuur 1) sprake van *maatschappijgerichtheid*. Communicatie over de aanpak van het ontwerpen en ontwikkelen van toetsen vindt tegelijkertijd plaats met het proces van visie- en beleidsontwikkeling over toetsen. Hierbij wordt geanticipeerd op actuele (maatschappelijke) ontwikkelingen en het belang van een leven lang leren. Het werkveld, studenten en andere relevante stakeholders worden betrokken bij het ontwerpen en ontwikkelen van toetsen. Om de kwaliteit van de toetsen vast te stellen worden deze voorgelegd aan (excellente) collega-opleidingen en aan (externe) deskundigen. De toetsen worden systematisch geëvalueerd als onderdeel van de kwaliteitscyclus met alle belanghebbenden, zoals de examencommissie, accreditatiecommissie, studenten en het werkveld. De toetsresultaten worden vergeleken met die van excellente collega-opleidingen.

Analoog aan het voorbeeld van toetsen kunnen de cellen van figuur 1 worden ingekleurd. Een invulling van kwaliteit van toetsbekwaamheid is te vinden in Sluijsmans et al. (2012). Met verder onderzoek zal worden bepaald of de methodiek betrouwbaar is (is de methodiek nauwkeurig genoeg, geeft de uitkomst een betrouwbare afspiegeling van de daadwerkelijke toetskwaliteit?), valide is (dekt de methodiek het domein van toetskwaliteit, worden er geen zaken over het hoofd gezien?) en transparant is (is de methodiek inzichtelijk genoeg voor de gebruiker, worden de instructies en bronnen zo begrepen als ze zijn bedoeld?). Voor dit doel wordt het prototype van de methodiek voorgelegd aan betrokkenen van een aantal opleidingen (leden toets- en examencommissies, management). In focusgroepinterviews zullen hen vragen worden voorgelegd als: Past deze opzet bij de vraag hoe kwaliteit van beoordelen en de verbeterstappen kunnen worden bepaald? Welke begrippen zijn niet duidelijk? Welke aanvullende

informatie heb je nodig om de methodiek op een goede manier te begrijpen en te benutten? Wie kunnen worden betrokken bij de positie- en ambitiebepaling? Het blijft van belang er voor te waken dat een goede schriftelijke vastlegging niet voldoende is voor kwaliteit, maar dat het kwaliteitsdenken wordt verankerd in het denken en handelen van alle betrokkenen bij beoordelen (Sluijsmans et al., 2012).

In dit hoofdstuk heb ik een professionele beoordelingscultuur benoemd als een omgeving waarin betrokkenen op basis van een visie en kernwaarden werken aan kwaliteit van beoordelen. Dit doen zij door met elkaar kennis te ontwikkelen en elkaar als kritische vrienden in te zetten, waarbij het gebruik van peer feedback een belangrijke activiteit is. Een ander kenmerk van een professionele beoordelingscultuur is de proactieve houding in het ontwikkelen van een kwaliteitszorgsysteem voor beoordelen. Dit kwaliteitszorgsysteem is bij voorkeur zo ingericht dat het niet alleen is gericht op het vaststellen van hoe de kwaliteit nu is, maar ook mogelijkheden biedt te komen tot duurzame, maatschappijgerichte kwaliteit. Wanneer een opleiding de kwaliteitszorg voor beoordelen op orde heeft, beïnvloedt deze de voorbereiding en uitvoering van een accreditatie positief. Een professionele beoordelingscultuur is het fundament voor het bereiken van professioneel beoordelen. Als deze cultuur er niet is, dan is het realiseren van professioneel beoordelen een illusie.

Kernboodschappen bouwsteen 5, een professionele beoordelingscultuur:

- Een professionele beoordelingscultuur wordt gekenmerkt door een professionele leergemeenschap waarin betrokkenen met elkaar kennis en ervaringen uitwisselen over beoordelen en zo de kwaliteit van beoordelen te versterken.
- In een professionele beoordelingscultuur is sprake van een doordacht kwaliteitszorgsysteem dat is ontworpen door opleidingen. In dit kwaliteitszorgsysteem zijn eisen van accreditatiecommissies vertaald naar de eigen opleidingscontext.
- Een professionele beoordelingscultuur wordt gekenmerkt door een bewustzijn van de mate van huidige kwaliteit van beoordelen en de ambities. Daarbij is geen sprake meer van ad-hocbeleid maar wordt toegewerkt naar een maatschappijgerichte invulling van kwaliteit van beoordelen.

Naar een definitie van professioneel beoordelen

In de voorafgaande hoofdstukken heb ik vijf bouwstenen geïntroduceerd die het mogelijk maken professioneel beoordelen in het hbo vorm te geven. Een definitie van professioneel beoordelen kan worden geformuleerd door het samenbrengen van de symbolen van de vijf bouwstenen (zie figuur 2).

Figuur 2. Vijf bouwstenen voor professioneel beoordelen in het hbo

Figuur 2 laat zich vervolgens vertalen in de volgende definitie van professioneel beoordelen:

'Professioneel beoordelen betekent dat de student in de weg naar startbekwame beroepsbeoefenaar in een *gebalanceerd en samenhangend programma van beoordelen* een groot aantal beoordelingstaken uitvoert. Deze taken zijn ontworpen door *bekwame beoordelaars*. De beoordelingstaken leveren informatie c.q. bewijs op over het kennen en kunnen. *Bekwame beoordelaars* verzamelen en interpreteren de informatie c.q. het bewijs op een zodanige wijze dat deze interpretatie niet alleen de *studenten* handvatten biedt die hen motiveren en inspireren beter te worden in de *professie*, maar ook leiden tot het nemen van betrouwbare, valide en transparante beslissingen over *het professioneel vakmanschap, het onderzoekend vermogen en de mate van zelfontwikkeling*. Het hele proces voltrekt zich in een leer- en werkomgeving waarin alle betrokkenen professioneel handelen en zich professioneel gedragen, in een leer- en werkomgeving die representatief is voor de professie en in een leer- en werkomgeving die is gericht op *kwaliteitsbewustzijn en voortdurende kwaliteitsverbetering*.'

Samen praktijkonderzoek doen naar professioneel beoordelen

Met welke vragen gaat het lectoraat aan de slag?

De vijf bouwstenen voor professioneel beoordelen in het hoger onderwijs bieden houvast bij het definiëren van praktijkgerichte onderzoeksvragen. In tabel 2 (zie pagina 54) wordt per bouwsteen een aantal vraagstukken genoemd. Dit overzicht is niet definitief of uitputtend maar geeft een eerste richting.

Hoe gaat het lectoraat met deze vragen aan de slag?

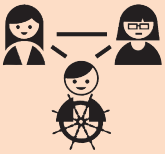
De vragen uit tabel 2 zullen worden geoperationaliseerd in onderzoeksvragen die voor het overgrote deel zullen worden beantwoord door het uitvoeren van praktijkonderzoek. Het lectoraat beoogt met dit onderzoek drie doelen:

1) het ontwikkelen van kennis en vaardigheden door docenten die deelnemen aan lectoraatsactiviteiten zoals de kenniskring;
2) het op gang brengen van kennisuitwisseling en -constructie binnen en buiten de hogeschool; en
3) een kwaliteitsimpuls geven aan beoordelen in het hbo.

Het doen van praktijkonderzoek wordt in het lectoraat beschouwd als een professionele leerstrategie waarmee docenten hun eigen dagelijkse praktijk kunnen verbeteren (Bolhuis & Kools, 2012). Er zijn vele vormen van praktijkonderzoek, waarvan de exacte invulling nog wordt bepaald door het type onderzoeksvraag. Belangrijk criterium bij het uitvoeren van het onderzoek is dat er sprake is van een systematische onderzoeksaanpak, om zo te bewaken dat de methode van onderzoek past bij de vraag die wordt onderzocht. Een ander criterium is dat de gegevens die worden verzameld om tot antwoorden te komen voldoende mogelijkheden bieden voor vertaling naar de onderwijspraktijk. Drie mogelijke vormen van onderzoek die leiden tot rijke gegevens zijn het doen van narratief onderzoek, ontwerponderzoek en beleidsonderzoek.

Narratief onderzoek is een vorm van onderzoek waarbij betrokkenen (docenten, studenten, werkveld) worden uitgenodigd hun praktijkkennis en -ervaringen mondeling en/of schriftelijk te expliciteren door bijvoorbeeld het bijhouden van een logboek, het beschrijven van anekdotes of het vertellen van verhalen over beoordelingservaringen (De Vries, 2007). Door deze manier van gegevens

Tabel 2.
Overzicht van vraagstukken van het lectoraat professioneel beoordelen

Bouwsteen	Vraagstukken
	- Wat zijn de kenmerken van een toekomstbestendig opleidingsprofiel? - Hoe kan de houdbaarheidsdatum van kennis van een professie worden vastgesteld en beoordeeld? - Wat betekent onderzoekende professional in verschillende beroepsdomeinen?
	- Hoe kan een evenwichtig programma van professioneel beoordelen in het hbo worden samengesteld op basis van een specifieke professie? - Hoe kan in een programma van beoordelen rekening worden gehouden met de expertiseontwikkeling van studenten? - Hoe kan informatie van een aantal beoordelingen worden geaggregeerd?
	- Wat zijn de kenmerken van professionele beoordelaars in het hbo, welk gedrag laten zij zien? - Welke beoordelaars zijn nodig om tot betrouwbare uitspraken over leren te komen? - Hoe kan beoordelen tijdens het leren vorm krijgen?
	- Hoe kunnen studenten worden opgeleid tot kennis-werkers en hoe kan worden beoordeeld of zij groeien als kenniswerkers? - Wat zijn de kenmerken van de professionele student in het hbo, welke gedragingen laten zij zien? - Hoe kan zelfsturing bij studenten in het hbo worden bevorderd?
	- Wat zijn de kenmerken van een kwaliteitszorgsysteem voor professioneel beoordelen? - Welke condities zijn nodig om tot een kwaliteitszorg-systeem voor professioneel beoordelen te komen? - Hoe kan structureel gebruik van een kwaliteitszorg-systeem voor professioneel beoordelen worden gefaciliteerd?

verzamelen is het mogelijk zicht te krijgen op drijfveren, motivaties, percepties en kennis van betrokkenen ten aanzien van beoordelen. *Ontwerponderzoek* is gericht op het ontwerpen, ontwikkelen, implementeren en evalueren van beoordelingsinterventies in specifieke opleidingssituaties (Van den Akker, Gravemeijer, McKenney & Nieveen, 2006). Het kan dan bijvoorbeeld gaan om het ontwerp van een professionaliseringstraject (professioneel beoordelen voor hbo-docenten), een ontwerp van een specifieke beoordelingstaak of het ontwerp van een evaluatie-instrument voor kwaliteit van beoordelen. Het doel van ontwerponderzoek is de bestaande beoordelingspraktijk (duurzaam) te verbeteren en kennis te ontwikkelen. Bij voorkeur wordt ontwerponderzoek uitgevoerd met docenten uit de praktijk waarbij samenwerking en kritische reflectie centraal staan, om zo te komen tot een ontwerp dat meerwaarde heeft voor het opleiden van studenten. *Beleidsonderzoek* is gericht op het monitoren en evalueren van reeds ingezet of nog te plannen beleid. Het is toepassingsgericht onderzoek dat voedend is voor instanties die beleid voor de samenleving ontwikkelen, uitvoeren of evalueren (Van Hoesel, 2003). Voor het lectoraat betekent dit bijvoorbeeld onderzoek naar de uitvoering en effectiviteit van toetsbeleid, zowel op het niveau van opleidingen als het niveau van de hogeschool.

Met wie werkt het lectoraat samen en hoe ziet deze samenwerking er uit?
Het praktijkonderzoek wordt uitgevoerd met collega's binnen en buiten Zuyd. De belangrijkste personen met wie ik in het lectoraat aan de slag ga zijn de leden van de kenniskring Professioneel Beoordelen. Deze kenniskring bestaat uit zeven bevlogen, enthousiaste docenten met grote affiniteit met beoordelen en het doen van onderzoek. Met veel plezier stel ik ze vandaag aan u voor: Anneriet Florack-van den Akker (De Nieuwste Pabo), Alexandra Jacobs (faculteit Sociale Studies en Educatie/opleiding Social Work), Xandra Janssen-Brandt (Academie Verloskunde Maastricht), Dorien Gerards (faculteit Gezondheidszorg/opleiding Fysiotherapie), Carin Gulikers (faculteit ICT), Elly Vermunt (faculteit Financieel en Commercieel Management) en Karin Welten (faculteit Hotel en Facility Management). Cees van der Vleuten van de Universiteit Maastricht (Graduate School of Health Professions Education) neemt als bijzonder lid deel aan de kenniskring. We zijn blij zijn kennis over beoordelen in te zetten in het onderzoek. De kenniskring is officieel 1 januari 2013 van start gegaan en komt driewekelijks een ochtend bij elkaar. Vijf werkprincipes zijn in de kenniskring leidend (conform Kwakman, 2010): maatwerk in het type onderzoek dat wordt uitgevoerd, houvast bieden in het (leren) doen van onderzoek, mogelijkheden creëren voor perspectiefwisseling (bijvoorbeeld een presentatie geven), het geven en ontvangen van constructieve feedback en ruimte geven aan rolontwikkeling in het ontwikkelen van kennis over onderzoek en beoordelen, zowel als het gaat om producten (bijvoorbeeld de ontwikkeling van een beoordelingsinstrument) als processen (wat is er nodig om

bij mijn collega's een onderzoekende houding naar de eigen beoordelingspraktijk te realiseren?). Op dit moment zijn we druk doende met vraagarticulatie en het maken van een eerste schets van ieders onderzoek. De vraagstellingen worden ingegeven door de professionele praktijk en persoonlijke interesse en hebben tot doel kennis, inzichten en producten te genereren die bijdragen aan het oplossen van de problemen in de beroepspraktijk (HBO-raad, 2010).

Naast de kenniskring werkt het lectoraat nauw samen met Marcel van der Klink, lector Professionalisering van het Onderwijs. Als programmamanagement van het innovatieprogramma *Zuyd Innoveert* zullen wij ons inzetten om binnen Zuyd het innoverend vermogen van docenten en faculteiten te stimuleren. Door het uitvoeren van microprojecten door individuele docenten en macroprojecten door teams in de faculteiten wordt de komende jaren hard gewerkt aan structurele positieve onderwijsveranderingen op thema's als studierendement, leren met nieuwe technologie, leren in en voor de professie en beoordelen.

Het lectoraat werkt ook samen met een aantal faculteiten. Activiteiten zijn gericht op het analyseren en versterken van programma's van beoordelen en het professionaliseren van docenten in bijvoorbeeld de rol van beoordelaar of het beoordelen van afstudeerwerkstukken. Het lectoraat biedt daarnaast feedback aan faculteiten op bijvoorbeeld visiedocumenten of concrete beoordelingsinstrumenten. De komende tijd wordt toegewerkt naar een structurele samenwerking tussen het lectoraat en de netwerken binnen de hogeschool die zich richten op toetsen, zoals het netwerk Toetsexpertise en het netwerk Examencommissie.

Er zal tevens nauw worden samengewerkt met de dienst Onderwijs en Onderzoek. Deze dienst wordt geleid door Rico Camp. Met Annerien Pin en Elly Theunissen onderzoekt het lectoraat of de bouwstenen van professioneel beoordelen kunnen worden verankerd in het toetsbeleid van Zuyd Hogeschool. Gezien de taakstelling van de dienst Onderwijs en Onderzoek - kwaliteitszorg en beleidsonderzoek - is het zinvol andere samenwerkingsmogelijkheden te verkennen die de relatie tussen het lectoraat en de dienst versterken.

Ook buiten Zuyd is het lectoraat actief in de samenwerking met collega-lectoraten, andere hogescholen en universiteiten. Het komende jaar continueert het lectoraat haar nauwe samenwerking met lector Desirée Joosten-ten Brinke van het lectoraat Eigentijds toetsen en beoordelen van Fontys Lerarenopleiding Tilburg. Activiteiten die in deze samenwerking centraal staan zijn het SURF-project *Bewust en bekwaam toetsen* en het schrijven van een aantal publicaties over de rol van beoordelaars en de implementatie van formatief beoordelen. Naast het SURF-project wordt in opdracht van Stichting Voortgezet Onderwijs

Parkstad Limburg (SVOPL) een scholingstraject voor docenten voorbereid. Ook in de samenwerking met Jos Castelijns, lector Eigentijds Beoordelen in het Onderwijs bij Hogeschool de Kempel, is een gemeenschappelijke drijfveer gevonden voor beoordelen om te leren. Met de Hogeschool Arnhem en Nijmegen, in het bijzonder met Annemieke Peeters en Linda Jakobs, zet het lectoraat de samenwerking voort in het kader van het project *De toetsing getoetst* (Sluijsmans et al., 2012). Tenslotte levert het lectoraat in samenwerking met de HBO-raad een bijdrage aan de vormgeving van een basis- en seniorkwalificatie examinering en de ontwikkeling van de landelijke toetsen in het kader van 1*ovoordeleraar*.

Tot slot

Goed beoordelen is belangrijk. De berichtgevingen van de laatste jaren duiden er op dat het slecht zou zijn gesteld met de kwaliteit van beoordelen in het hbo. Dat is maar ten dele waar. Ik vind dat in het hbo al een enorme slag is gemaakt met het vormgeven van beoordelen op een wijze die past bij professioneel opleiden. Er liggen ook nog vele kansen om beoordelen nog meer te laten aansluiten bij de doelen van het hbo.

Het verankeren van beoordelen in leren is de kernboodschap van deze rede. Pas als sprake is van integratie van leren en beoordelen benutten we alle kansen hbo-studenten op te leiden tot enthousiaste, verantwoordelijke en onderzoekende professionals. Enig realisme is daarbij op zijn plaats; voor het vormgeven van professioneel beoordelen bestaat geen trukendoos en implementatie van professioneel beoordelen kan niet worden gegarandeerd met intreeredes en beleidsstukken. Om tot een visie op beoordelen te komen die past bij het hbo is het doen van praktijkonderzoek en het leren van en met elkaar als professionele leergemeenschap nodig. Het vormgeven van een rijk programma van beoordelen passend bij de professie, investeren in professionele ontwikkeling van docenten als beoordelaars, studenten stimuleren verantwoordelijkheid te durven en willen nemen voor hun leren en de regie nemen over kwaliteitszorg, zie ik als belangrijke uitdagingen voor de komende jaren. Werken aan een visie en implementatie van professioneel beoordelen kan echter niet plaatsvinden zonder ruimte voor creatief en innovatief denken, met voldoende ruimte voor autonomie van professionals. Bouwstenen zijn niet alleen bedoeld om te bouwen maar ook om mee te spelen.

Graag sluit ik af met het beeld van de student die straks met passie en inzet zijn professie uitoefent en gemotiveerd is zichzelf te blijven ontwikkelen. Ik hoop dat we studenten in hun ontwikkeling naar jonge professionals kunnen laten ervaren dat beoordelen ook motiverend en leerzaam kan zijn. Dat begint met te onderkennen dat beoordelen verweven is met onderwijs en al lang niet meer louter het sluitstuk is van leren. Dat er naast de gangbare vormen van beoordelen veel andere creatieve manieren zijn om de goede informatie over studenten te verzamelen. Dat beoordelen meer is dan meten. Dat investeren in beoordelen,

investeren in leren is. Ik kijk er naar uit met collega's binnen en buiten Zuyd professioneel beoordelen een gezicht te geven.

Rest mij een kort woord van dank. Op de eerste plaats dank aan alle collega's van Zuyd, voor het warme welkom en de collegialiteit. Veel dank aan allen met wie ik binnen en buiten het lectoraat samenwerk en met wie ik veel plezier beleef. Een groot woord van dank aan het College van Bestuur, in het bijzonder aan Kitty Kwakman voor de prachtige kans om bij Zuyd mijn missie - beoordelen verankeren in leren - voort te zetten. De condities waaronder ik dit kan doen bieden alle mogelijkheden hiervoor. Ik voel mij thuis.

Tot slot dank ik u voor uw warme blijk van belangstelling.

Dominique Sluijsmans

Heerlen
19 april 2013

Referenties

AFT, NCME, & NEA (1990). *Standards for Teacher Competence in Educational Assessment of Students*. American Federation of Teachers, National Council on Measurement in Education & National Education Association.

Aguilar, A. E., Stupans, L., & Scutter S. (2011). Assessing students' professionalism: considering professionalism's diverging definitions. *Education for Health, 24*(3), 599.

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.

Baartman, L. K. J., Bastiaens, Th. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2006). The wheel of competency assessment: Presenting quality criteria for competency assessment programmes. *Studies in Educational Evaluation, 32,* 153-177.

Baer, J., & McKool, S. S. (2009). Assessing creativity using the consensual assessment technique. In S. Schreiner (Ed.), *Handbook of Research on Assessment Technologies, Methods, and Applications in Higher Education* (pp. 65-77). IGI Global.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*(5), 536–553.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*, 5-25.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347-364.

Birenbaum, M., Kimron, H., & Shilton, H. (2011). Nested contexts that shape assessment for learning: School-based professional learning community and classroom culture. *Studies in Educational Evaluation, 37*(1), 35-48.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 453-459). Cambridge, MA: MIT Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*, 7-74.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.

Bolhuis, S., & Kools, Q. (2012). *Praktijkonderzoek als professionele leerstrategie in onderwijs en opleiding*. Tilburg: Fontys Lerarenopleiding Tilburg.

Boon, J., & Van der Klink, M. (2000). De waarde van beroepsprofielen in het Hoger Onderwijs. *Tijdschrift voor Hoger Onderwijs*, *18*(1), 25-36.

Boshuizen, E., & Van de Wiel, M. (in voorbereiding). *Expertise development through schooling and work*.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, *22*, 151-167.

Boud, D. and Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council.

Brett, J. F., & Atwater, L. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology, 86*, 930-942.

Brown, G. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy and Practice*, *11*, 301-318.

Bruijn, J. A., & De Vries, S. (2012). *Onderwijs: de derde dimensie*. Geschrift 117. Den Haag: Teldersstichting.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*, 245-281.

Chiviacowsky, S., & Wulf, G. (2005). Self-controlled feedback is effective if it is based on the learner's performance. *Research Quarterly for Exercise and Sport, 76*, 42-48.

Commissie Bruijn (2012). *Vreemde ogen dwingen. Eindrapport Commissie externe validering examenkwaliteit hoger beroepsonderwijs*. Den Haag: HBO-raad.

Commissie Gedragscode Praktijkgericht Onderzoek (2010). *Gedragscode praktijkgericht onderzoek voor het hbo. Gedragscode voor het voorbereiden en uitvoeren van praktijkgericht onderzoek binnen het Hoger Beroepsonderwijs in Nederland*. Den Haag: HBO-raad.

Commissie Veerman (2012). *Differentiëren in drievoud. Advies van de Commissie Toekomstbestendig Hoger Onderwijs Stelsel*. Den Haag: HBO-raad.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*, 438-481.

Crow, G. M., & Pounder, D. G. (2000). Interdisciplinary teacher teams: Context, design, and process. *Educational Administration Quarterly, 36*, 216-254.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York, US: Harper & Row.

Delandshere, G., & Petrosky, A. (1998). Assessment of complex performances: Limitations of key measurement assumptions. *Educational Researcher, 27*(2), 14-25.

De Groot, A. (1966). *Vijven en zessen: Cijfers en beslissingen*. Groningen: Wolters-Noordhoff.

De Korne, D., Hiddema, F., Bleeker, F., & Klazinga, N. (2009). Operatie in beeld. *Medisch Contact*, *64*(2), 72-77.

De Vries, B. (2007). *Vrijheid, blijheid? Over vraaggestuurd leren en flexibele scholen*. Nijmegen: Hogeschool Arnhem en Nijmegen.

Den Hertog, F., & Huizenga, E. (1997). *De kennisfactor: concurreren als kennisonderneming*. Den Haag: Kluwer.

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.

Dijkstra, J., Van der Vleuten, C. P. M., & Schuwirth, L. W. T. (2010). A new framework for designing programmes of assessment. *Advances in Health Science Education, 15*, 379-393.

Ertmer, P. A., & Newby, T. J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science, 24*, 1-24.

Fishman, B. J., Marxa, R. W., Best, S., & Tal, R. T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education, 19*, 643-658.

Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. A*merican Psychologist, 3*, 193-202.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports student learning. *Learning and Teaching in Higher Education, 1*, 3-31.

Govaerts, M. J. B. (2011). *Climbing the pyramid. Towards understanding performance assessment*. Maastricht: Maastricht University.

Govaerts, M. J. B., Schuwirth, L. W. T, Van der Vleuten, C. P. M., & Muijtjens, A. M. M. (2011). Workplace-based assessment: effects of rater expertise. *Advances in Health Sciences Education, 16*, 151-165.

Govaerts, M. J. B., Van de Wiel, M., & Van der Vleuten, C. (2012). Quality of feedback following performance assessments: Does assessor expertise matter? *European Journal of Training and Development, 37*(1), 6-6.

Govaerts, M. J. B., Van de Wiel, M. W. J., Schuwirth, L. W. T., Van der Vleuten, C.P.M., & Muijtjens, A. M. M. (2012). Workplace-based assessment: raters' performance theories and constructs. *Advances in Health Sciences Education*, 1-22.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81-112.

HBO-raad (2008). *Kwaliteitszorgstelsel ten aanzien van het onderzoek aan hogescholen 2009-2015. Basisdocument. Nadere uitwerking van het brancheprotocol kwaliteitszorg onderzoek (BKO) in een kwaliteitszorgstelsel.* Den Haag: HBO-raad.

HBO-raad (2009). *Kwaliteit als opdracht*. Den Haag: HBO-raad.

HBO-raad (2010). *Naar een duurzaam onderzoeksklimaat. Ambities en succesfactoren voor het onderzoek aan hogescholen*. Den Haag: HBO-raad.

HBO-raad (2011). *Geslaagd! Handreiking examencommissies*. Den Haag: HBO-raad.

Janssen-Noordman, A. M. B., & Van Merriënboer, J. J. G. (2002). *Innovatief onderwijs ontwerpen: Via leertaken naar complexe vaardigheden*. Groningen: Wolters-Noordhoff.

Jaspers, M., & Van Zijl, E. (2011). *Kwaliteit van toetsing in het Hoger Onderwijs: naar een versterkte rol van examencommissies*. Tilburg: Fontys Hogescholen, Dienst Onderwijs en Onderzoek.

Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen*. Lectorale rede. Tilburg: Fontys Hogescholen.

Joosten, D., & Sluijsmans, D. M. A. (2012). Tijd voor toetskwaliteit: het borgen van toetsdeskundigheid van examencommissies. *TH&MA, 19*(4), 16-21.

Kempen, P., Keizer, J., & Van den Berg, B. (2012). Hbo-waardig afstuderen: hoe te regelen en te verankeren? *Th&ma, 4*, 38-43.

Kessels, J. W. M. (1996). *Het corporate curriculum*. Inaugurele rede. Leiden: Universiteit Leiden.

Klimoski, R, & Mohammed S. (1994). Team mental model: Construct or metaphor? *Journal of Management, 20*, 403-437.

Kremer-Hayon, L., & Tillema, H. H. (1999). Self-regulated learning in the context of teacher education. *Teaching and Teacher Education, 15*, 507-522.

Krueger, J., & Dunning, D. (1999). Unskilled, and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 82*, 180-188.

Kwakman, K. (2003). *Anders leren, beter werken*. Lectorale rede. Nijmegen: Hogeschool van Arnhem en Nijmegen.

Kwakman, K. (2010). Leren onderzoek doen binnen de hogeschool. Reisverslag naar een onderzoeksdidactiek. In G. Schuiling, H. Vermaak & H. Tours (Eds), *Leren in organisaties* (pp. 197-212). Deventer: Kluwer.

Kwakman, K., & Van den Berg, E. (2004). Professionele ontwikkeling als kennisontwikkeling door leraren: naar een betere interactie tussen praktijk en theorie. *VELON Tijdschrift voor Lerarenopleiders, 25*(3), 6-12.

Leijnse, F., Hulst, J., & Vroomans, L. (2006). Passie en precisie. Over de verandererende functie van hogescholen. *Th&ma, 5*, 47-54.

Lunenberg, M., Dengerink, J., & Korthagen, F. (2013). *Het beroep van lerarenopleider: Professionele rollen, professioneel handelen en professionele ontwikkeling van lerarenopleiders*. Amsterdam: Vrije Universiteit.

Maister, D. (2000). *True professionalism*. New York: Touchstone.

Martens, R. L. (2007). *Positive learning met multimedia* (Inaugural address). Heerlen/Hilversum: Open Universiteit Nederland/Teleac/NOT.

Meherns, W. A., Popham, W. J., & Ryan, J. M. (1998). How to prepare students for performance assessments. *Educational Measurement: Issues and Practice, 17*, 18-22.

Mertens, F. (2011). Hoger Onderwijs Autonomie en Kwaliteit. *Nederlands Tijdschrift voor Onderwijsrecht en Onderwijsbeleid, 1*, 61-66.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Acadamic Medicine, 65*, S63–S67.
Mijland, E. (2011). *Nu leren voor morgen. 25 gesprekken over eigentijds onderwijs*. Apeldoorn: Garant.
Mui So, W., & Hoi Lee, T. (2011). Influence of teachers' perceptions of teaching and learning on the implementation of Assessment for Learning in inquiry study. *Assessment In Education: Principles, Policy & Practice, 18*(4), 417-432.
Myers, D. G. (1980). *The inflated self*. New York: Seabury Press.
Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist, 22*, 155-175.
Nederlands-Vlaamse Accreditatieorganisatie (2010). *Beoordelingskaders ten behoeve van het accre-ditatiestelsel*. Den Haag: NVAO.
NSW Institute of Teachers (2005). *Professional teaching standards*. Sydney: NSW Institute of Teachers.
Oakland, J. S. (1993). *Total Quality Management: the route to improving performance*. Butterworth-Heinemann.
Onderwijsraad (2007). *Versteviging van kennis in het onderwijs II*. Den Haag: Onderwijsraad.
Pink, D. (2009). *Drive. The surprising truth about what motivates us*. New York, US: Riverhead Books.
Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher, 6*(1), 21-27.
Popham, W. J. (2001). Teaching to the test. *Educational Leadership, 58*(6), 16-20.
Projectteam Consilium Abeundi (2002). *Professioneel gedrag: Onderwijs, toetsing, begeleiding en regelgeving*. Utrecht: Projectteam Consilium Abeundi.
Raemdonck, & Strijbos, J. W. (2013). Feedback perceptions and attribution by secretarial employees: Effects of feedback-content and sender characteristics. *European Journal of Training and Development, 37*(1), 24-48.
Regehr, G., & Eva, K. (2006). Self-assessment, self-direction, and the self-regulating professional. *Clinical Orthopaedics and Related Research, 449*, 34-38.
Reijngoud, T. (2012). *Weten is meer dan weten*. Hilversum: Uitgeverij Lias.
Ruiz-Primo, M. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation, 37*(1), 15-24.
Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68-78.
Sadler, R. D. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Schuiling, G. (2007). *Werken in een portfolio van rollen*. Nijmegen: Hogeschool Arnhem en Nijmegen.
Schuiling, G. (2012). Vijf uitdagingen voor de volgende tien jaar. Aanzet voor een organisatieontwikkelingsagenda van hogescholen. *Th&ma*, 5, 11-16.
Schuwirth, L. W. T. (2007). *Meten de maat genomen*. Maastricht: Maastricht University.
Schuwirth, L. W. T., & Van der Vleuten, C. P. M. (2005). A plea for new psychometric models in educational assessment. *Medical Education*, 40, 296-300.
Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35, 453-472.
Scouller, K. M., & Prosser, M. (1994). Students' experiences in studying for multiple choice question examinations. *Studies in Higher Education*, 19, 267-279.
Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York: Doubleday.
Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E., Ruiz-Primo, M., & Yue, Y. (2008). On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers. *Applied Measurement in Education*, 21(4), 295-314.
Sluijsmans, D. M. A. (2002). *Student involvement in assessment: The training of peer assessment skills*. Unpublished doctoral dissertation, Open Universiteit Nederland, Heerlen, the Netherlands.
Sluijsmans, D. M. A. (2008). *Betrokken bij beoordelen*. Lectorale rede. Nijmegen: Hogeschool Arnhem en Nijmegen.
Sluijsmans, D. M. A. (2010). Beoordelen met het oog op de toekomst: Tien kenmerken van duurzaam beoordelen. *Examens*, 7, 5-9.
Sluijsmans, D. M. A., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Martens, R. L. (2004). Training teachers in peer-assessment skills: Effects on performance and perceptions. *Innovations in Education and Teaching International*, 41, 59-78.
Sluijsmans, D. M. A., Dochy, F., & Moerkerke, G. (1999). Creating a learning environment by using self-, peer- and co-assessment. *Learning Environments Research*, 1, 293-319.
Sluijsmans, D. M. A., Joosten-ten Brinke, D., Van der Vleuten, C. P. M. (2013). *Toetsen met leerwaarde. Een reviewstudie naar de effectieve kenmerken van formatief toetsen*. Den Haag: NWO.
Sluijsmans, D. M. A., Peeters, A., Jakobs, L., & Weijzen, S. (2012). Kwaliteit van toetsing onder de loep. *Onderwijsinnovatie*, 4, 17-25.
Sluijsmans, D. M. A., Straetmans, G., & Van Merriënboer, J. (2008). Integrating authentic assessment with competency based learning: The protocol portfolio scoring. *Journal of Vocational Education and Training*, 159-172.

Sluijsmans, D. M. A., Van Eldik, S., Joosten-ten Brinke, D., & Jakobs, L. (in voorbereiding). *Bewust en bekwaam toetsen. Een eerste invulling van een kennisbasis toetsen voor lerarenopleiders*.

Smeijsters, H., & Sporken, S. (2004). *Van taak tot competentie: 'leren leren' voor het hoger beroepsonderwijs*. Houten: Bohn Stafleu van Loghum.

Smits, R. (2011). Examencommissies HBO worden belangrijker! *Examens, tijdschrift voor de toetspraktijk, 2*, 16-19.

South African Qualification Authority (2001). *Criteria and guidelines for the registration of assessors. Policy document*. South Africa: Pretoria.

Stichting Duurzaam Hoger Onderwijs (2010). *AISHE: Auditing instrument for sustainability in higher education*. Duurzaam Hoger Onderwijs.

Stiggins, R. J. (2009). Essential formative assessment competencies for teachers and school leaders. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment*. New York: Routledge.

Straetmans, G. J. J. M. (2006). *Bekwaam boordelen en beslissen*. Lectorale rede. Deventer: Saxion Hogescholen.

Straetmans, G., Sluijsmans, D. M. A., Bolhuis, B., & Van Merriënboer, J. J. G. (2003). Integratie van instructie en assessment in competentiegericht onderwijs. *Tijdschrift voor Hoger Onderwijs, 3*, 171-197.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education, 13*(3), 239–264.

Van Bemmel, A. (2006). *Hogescholen en hbo in historisch perspectief*. Den Haag: HBO-raad.

Van Berkel, A. (2012). Kritische reflectie op competentietoetsen in het hbo. *Onderwijsinnovatie, 2*, 17-26.

Van Berkel, H., & Bax, A. (2002). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.) (2006). *Educational design research*. London: Routledge.

Van der Klink, M. (2011). *Bekwaam innoveren voor een toekomstbestendig hoger beroepsonderwijs*. Lectorale rede. Heerlen: Zuyd Hogeschool.

Van Driel, J. & Verloop, N. (1998). 'Pedagogical content knowledge': een verbindend element in de kennisbasis van docenten. *Pedagogische Studiën, 75*, 225-237.

Van Gennip, N. (2012). *Assessing together: Peer assessment from an interpersonal perspective*. Leiden: Universiteit Leiden.

Van der Vleuten, C. P. M. (1996). The assessment of professional competence: Developments, Research and Practical Implications. *Advances in Health Science Education, 1*(1), 41-67.

Van der Vleuten, C. P. M., & Schuwirth, L. W. T. (2005). Assessment of professional competence: from methods to programmes. *Medical Education, 39*, 309-317.

Van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J, Tigelaar, D., Baartman, L. K. J., Van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher, 34*, 205-214.

Van der Vleuten, C. P. M, Schuwirth, L. W. T, Scheele, F., Driessen, E. W., & Hodges, B. (2010). The assessment of professional competence: building blocks for theory development. *Best Practice & Research Clinical Obstetrics and Gynaecology, 24*, 703-719.

Van Hoesel, P. (2003). *Beleidsonderzoek als professie, een lang gekoesterd ideaal*. Inaugurele rede. Zoetermeer: EIM.

Van Merriënboer, J. J. G., & Sluijsmans, D. M. A. (2009). Toward a synthesis between cognitive load theory and self-directed learning. *Educational Psychology Review*, 21(1), 55-66.

Van Rooij, M., Verdonschot, S., Dewulf, L., & Smits, B. (2008). *Maak werk van kenniswerk! Onderzoek naar kenniswerk(ers) in opdracht van AGORIA*. Kessels & Smit.

Verharen, L., van Bommel, M., & Vosselman, M. (2011). *De sociaal-agogische kennisbasis*. Houten: Bohn Stafleu Van Loghum.

Vonk, R. (2009). *Menselijke gebreken voor gevorderden*. Schiedam: Scriptum.

Voogt, J., & Pareja Roblin, N. (2010). *21st Century Skills. Discussienota in opdracht van Kennisnet*. Enschede: Universiteit Twente.

Ward, M., Gruppen, L., & Regehr, G. (2002). Measuring self-assessment: Current state of the art. *Advances in Health Science Education, 7*, 63-80.

Weggeman, M. (2000). *Kennismanagement: de praktijk*. Schiedam: Scriptum.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*, 3-14.

Wulf, G., Raupach, M., & Pfeiffer, F. (2005). Self-controlled observational practice enhances learning. *Research Quarterly for Exercise and Sport, 76*, 107-111.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477-501.

Zimmermann, B. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3-17.

Zuyd Onderzoek
Lectoraat Professioneel Beoordelen
Postbus 550
6400 AN Heerlen
www.zuyd.nl
info@zuyd.nl

Colofon
Tekst Dominique Sluijsmans
Eindredactie Dienst Marketing en Communicatie
Basisontwerp Zuiderlicht, Maastricht
Vormgeving Lucker Design, Stan Heijnen
Luc Sluijsmans (bouwstenen)
Druk Schrijen-Lippertz, Voerendaal

© Dominique Sluijsmans 2013
Alle rechten voorbehouden. Behalve de uitdrukkelijk bij wet bepaalde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen of openbaar gemaakt zonder de voorafgaande schriftelijke toestemming van de auteur of Zuyd Onderzoek.