

Beoordelen is mensenwerk

Bevindingen over de wenselijkheid en mogelijkheid van een gezamenlijk protocol voor het beoordelen van (kern)werkstukken

van de expertgroep Protocol
in opdracht van de Vereniging Hogescholen



1 februari 2014

Inhoudsopgave

Voorwoord.....	3
DEEL 1: Wenselijkheid en mogelijkheid van een protocol.....	4
1. Aanleiding, context en opdracht	5
2. Samenstelling en werkwijze expertgroep	6
3. Conceptueel model van afstuderen.....	7
4. Analyse	11
5. Conclusies en aanbevelingen	16
Deel 2: Protocol Verbeteren en Verantwoorden van Afstuderen in het hbo.....	19
Inleiding.....	20
Protocol	22
Uitwerking en toelichting.....	23
Deel 3: Visie op onderzoekend vermogen	28
Bijlagen	37
Bijlage 1: Lijst van gesprekspartners.....	38
Bijlage 2: Verklarende woordenlijst	39
Bijlage 3: Overzicht beoordelaarseffecten	40
Bijlage 4: Toetsing en beoordeling van eindniveau bij een visitatie.....	42
Bijlage 5: Handreiking Ontwikkelen beoordelingsmodellen	46
Bijlage 6: Handreiking Kalibreersessies	51
Gebruikte bronnen	53

Voorwoord

Voor u ligt het rapport van de expertgroep protocol. De expertgroep heeft van de Vereniging Hogescholen de opdracht gekregen om de wenselijkheid en mogelijkheid te verkennen van een gezamenlijk, bottom-up opgesteld protocol of protocollen voor het beoordelen van (kern)werkstukken en te adviseren over de wijze waarop deze tot stand dienen te komen en aan welke kwaliteitseisen deze dienen te voldoen.

De expertgroep is zich bij haar werk voortdurend bewust geweest van het dilemma dat in het idee van gezamenlijke protocollen zit verscholen. Enerzijds is het verantwoord beoordelen van (kern)werkstukken van studenten zo belangrijk voor de kwaliteit en geloofwaardigheid van het onderwijs dat er sterke waarborgen moeten zijn dat dit proces goed verloopt en transparant is. Dit maakt de vraag naar standaardisering en protocollering reëel. Anderzijds is het vaak onmogelijk en ook onwenselijk het beoordelen van prestaties terug te brengen tot een soort algoritme. Beoordelen is mensenwerk. Het streven naar volledige standaardisatie en objectiviteit leidt vrij gemakkelijk tot beoordeling van observeerbare maar oppervlakkige aspecten van het presteren, die geen recht doen aan de werkelijke kwaliteit van het te beoordelen product of gedrag.

Beoordelen is in de eerste plaats het werk van vakbekwame professionals. Dat werk moet worden ondersteund met het aanreiken van criteria en werkwijzen zodat het zoveel mogelijk geldigheid heeft. Maar de docent mag daarbij niet het gevoel krijgen dat het zijn of haar beoordeling niet meer is.

De expertgroep heeft haar opdracht uitgevoerd in het licht van dit dilemma tussen dicht-timmeren en het timmermansoog. Ze heeft daarbij gewerkt vanuit het principe: gezamenlijk wat kan, lokaal wat moet. In deel 1 geeft de expertgroep antwoord op de vraag wat de wenselijkheid en mogelijkheid is om te komen tot een gezamenlijk protocol. Met dit deel voldoet de expertgroep aan de opdracht. Kernboodschap van deel 1 is dat afstuderen in het hbo veel meer is dan het schrijven van een werkstuk.

In deel 2 doet de expertgroep een voorstel voor een landelijk protocol voor het verbeteren en verantwoorden van het afstuderen in het hbo. Kernboodschap is dat dit protocol een goed middel is om de systematische kwaliteitszorg van afstuderen te bevorderen en de externe validering te vergroten.

In deel 3 wordt specifiek ingegaan op de eis dat studenten moeten beschikken over onderzoekend vermogen, op de vraag welke rol dat speelt binnen het afstuderen en op de vraag welke eisen daaruit voortvloeien. Kernboodschap is dat onderzoekend vermogen in het hbo-onderwijs een ondersteunende rol heeft bij het maken van goede beroepsproducten en dat de lat daarom niet te hoog moet worden gelegd. Dit deel is geen onderdeel van haar opdracht maar is in de ogen van de expertgroep wel voorwaardelijk voor transparantie van afstuderen.

De expertgroep is van mening dat een goede verantwoording van opleidingen over het afstudeerproces bijdraagt aan externe validering en daarmee aan de kwaliteit en legitimiteit van het diploma. Landelijke afspraken hierover kunnen helpen onduidelijkheden weg te nemen en transparantie te bevorderen.

Namens de expertgroep,

Daan Andriessen
Voorzitter expertgroep

Petra Manders
Secretaris expertgroep

Januari 2013

DEEL 1:
WENSELIJKHEID EN MOGELIJKHEID VAN
EEN PROTOCOL

*Afstuderen in het hbo is veel meer dan het
schrijven van een werkstuk*

1. Aanleiding, context en opdracht

Sinds 2010 worden er in Nederland met enige regelmaat vraagtekens gezet bij de kwaliteit van de hbo-diploma's. Reden daarvoor waren onder andere incidenten met alternatieve afstudeertrajecten bij enkele hogescholen. Hoewel uit onderzoeken van de Inspectie HO en de NVAO duidelijk werd dat het geen hbo-breed probleem was, leidden deze onderzoeken er wel toe dat het beeld ontstond dat de waarde van de hbo-diploma's niet boven alle twijfel verheven was.

Dat was voor de Vereniging Hogescholen (toen nog HBO-raad) reden om een commissie in te stellen (de Commissie Bruijn) met als opdracht te inventariseren welke mogelijkheden voor externe validering van examens er zijn (waaronder centrale examinering), wat de voor- en nadelen van de verschillende alternatieven zijn en of er nog andere opties voor externe validering zijn.

In mei 2012 leverde de commissie Bruijn haar rapportage 'Vreemde ogen dwingen' op, met daarin een zevental aanbevelingen. De Vereniging Hogescholen heeft uit die zeven aanbevelingen een viertal, onderling samenhangende en in haar ogen meest urgente projecten, gedefinieerd (Brief HBO-Raad aan Staatssecretaris Zijlstra, referentienummer 12-2124, O&S):

1. Externe validering door middel van gezamenlijke toetsing in pilots.
2. Verantwoording van gezamenlijk toetsen in jaarverslagen hogescholen.
3. Externe validering door middel van een gemeenschappelijk protocol voor de beoordeling van werkstukken (expertgroep Protocol).
4. Externe validering door middel van vergroting toetsdeskundigheid van docenten (expertgroep BKE/SKE).

Voor de implementatie van de overige aanbevelingen van de commissie Bruijn draagt iedere hogeschool zelf verantwoordelijkheid. Acties die hieruit voortvloeien dienen aan te sluiten bij de eigen situatie, profiel en visie.

Voor punt 3 is de expertgroep Protocol ingesteld. De opdracht aan de expertgroep Protocol is op 29 maart 2013 als volgt geformuleerd: "*Onderzoek de wenselijkheid en mogelijkheid van een gezamenlijk, bottom-up opgesteld protocol of protocollen voor het beoordelen van (kern)werkstukken en adviseer over de wijze waarop deze tot stand dienen te komen en aan welke kwaliteitseisen deze dienen te voldoen*" (E-mail R. Smits aan de leden van de expertgroep d.d. 29 maart 2013).

Deel 1 van dit rapport geeft antwoord op deze vraag en is als volgt opgebouwd. Hoofdstuk 2 schetst de samenstelling en werkwijze van de expertgroep. In hoofdstuk 3 wordt een conceptueel model van het afstuderen geschetst waarmee antwoord kan worden gegeven op bovenstaande vraag. Dit model wordt in hoofdstuk 4 gebruikt voor een analyse van de vraag en het formuleren van een antwoord. Hoofdstuk 5 bevat de conclusies en de aanbevelingen aan de Vereniging Hogescholen.

2. Samenstelling en werkwijze expertgroep

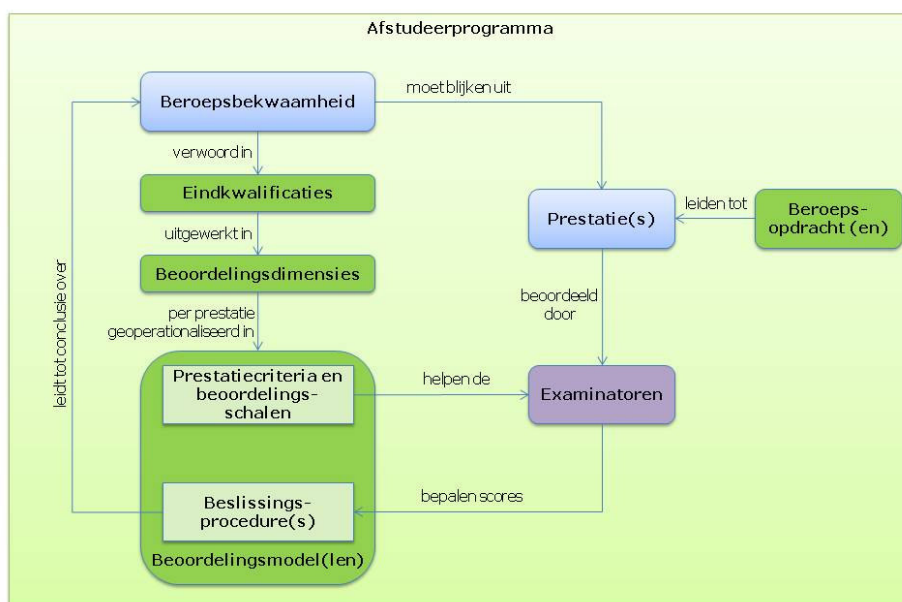
In overleg met de Vereniging Hogescholen is een expertgroep Protocol samengesteld (zie tabel 1). De expertgroep is acht maal bij elkaar geweest en heeft daarnaast gesprekken gevoerd (zie bijlage 1) met de NVAO, met twee evaluatiebureaus en met een lid van de Commissie Bruijn. In bijeenkomsten van de expertgroep is allereerst de opdracht verkend en gewerkt aan een gemeenschappelijk begrippenkader (zie bijlage 2). Vervolgens zijn ideeën geïnventariseerd voor het protocolleren van afstuderen. Ook zijn er casussen besproken uit verschillende instellingen waaruit blijkt tegen welke vraagstukken opleidingen aanlopen bij het afstuderen. Tussen de bijeenkomsten door hebben leden gewerkt aan teksten voor het eindrapport, zijn voorbeelden verzameld en is achtergrondinformatie gedeeld via Dropbox. Op basis hiervan is een protocol opgesteld dat is voorgelegd aan de NVAO, de drie evaluatiebureaus die actief zijn in het hbo en een zelfstandig opererende gecertificeerde secretaris.

Tabel 1: samenstelling van de expertgroep Protocol

Naam	Instelling	Hoedanigheid
Daan Andriessen (voorzitter)	Hogeschool Utrecht	Lector Methodologie van Praktijkgericht Onderzoek
Petra Manders (secretaris)	Saxion	Senior Beleidsmedewerker Kwaliteitszorg
Dan Greve	Hogeschool Utrecht	Onderwijsadviseur instituut ICT / Onderzoeker Methodologie van praktijkgericht onderzoek
Peter Hermans	Cito Artez	Knowledge Manager Research Program Manager Lecturer in Research in Arts Education
Linda Jakobs	Hogeschool Arnhem Nijmegen	Onderwijskundig adviseur Onderwijsontwikkeling
Lucie te Lintelo	Hogeschool van Amsterdam	Onderwijskundig adviseur en ontwikkelaar / trainer op gebied van (leerwegaafhankelijke) toetsen
Karin Neijenhuis	Hogeschool Rotterdam	Hoofddocent Evidence Based Care, Instituut voor Gezondheidszorg, opl. Logopedie/ Kenniscentrum Zorginnovatie
Gerard Straetmans	Saxion Cito	Lector Assessment Toetsdeskundige

3. Conceptueel model van afstuderen

Een onderzoek naar de wenselijkheid en mogelijkheid van een protocol vereist een helder begrippenkader. Daarom heeft de expertgroep een conceptueel model opgesteld dat aangeeft wat afstuderen in het hbo is en welke onderdelen en hulpmiddelen daarbij een rol spelen. Dit model is een abstractie en houdt geen rekening met de enorme variëteit aan afstudeerprogramma's die in het hbo voor komen. Figuur 1 geeft de essentie van het afstudeerproces grafisch weer vanuit het perspectief van de opleiding. In de tekst worden de kernbegrippen uit het model toegelicht.



Figuur 1. Modelmatige weergave van afstudeerprocessen in het hbo

Afstudeerprogramma

Het takenpakket van een beroepsbeoefenaar is veelal zo divers dat één prestatie daarvoor nooit representatief kan zijn (zie kader). Daarom hebben hbo-opleidingen vaak meerdere onderdelen in het curriculum aangewezen waarin de student laat zien te beschikken over de eindkwalificaties van de opleiding. Deze toetsonderdelen gezamenlijk noemt de expertgroep het afstudeerprogramma. Het programma bestaat uit één of meerdere beroepsopdrachten die uitgevoerd kunnen worden in stages, in opdrachten voor externe opdrachtgevers of in eigen projecten. Met het uitvoeren van de beroepsopdrachten wordt bewijsmateriaal verzameld op basis waarvan wordt besloten of de student de eindkwalificaties beheerst. Als ook alle andere toetsen in het toetsprogramma voldoende zijn kan besloten worden het diploma te verstrekken.

Als bijvoorbeeld een ambulanceverpleegkundige-in-opleiding laat zien dat hij adequaat kan handelen bij een geval van verdrinking, dan kan die prestatie hooguit worden opgevat als één bijdrage aan het bewijs voor beroepsbekwaamheid op het gebied van het verlenen van spoedeisende hulp door ambulanceverpleegkundigen. Het takenpakket van deze beroepsbeoefenaar en de omstandigheden waaronder hulp verleend moet worden zijn namelijk zo divers dat één prestatie daarvoor nooit representatief kan zijn. Het probleem van de beperkte generaliseerbaarheid van prestaties geldt voor de meeste bekwaamheden. De consequentie hiervan voor de beoordeling is dat er meerdere prestaties nodig zijn, die samen voldoende representatief zijn voor de taaksituaties waarin de betreffende beroepsbekwaamheid zich kan manifesteren.

Overigens is het vaak niet haalbaar om alle eindkwalificaties te toetsen met meerdere prestaties. Daarom wordt vaak een best practice per eindkwalificatie (bv in een portfolio-assessment) getoetst en/of de meest complexe beroepsopdracht.

Beroepsbekwaamheid

Het doel van het afstuderen is om vast te stellen of de student voldoende bekwaam is om adequaat te handelen in taaksituaties die representatief zijn voor de kerntaken van het beroep waarvoor wordt opgeleid. Dit gebeurt op basis van de beoordeling van door een student geleverde prestaties.

Eindkwalificaties

Een overzicht van de te toetsen eindkwalificaties vormt het uitgangspunt van ieder afstudeerprogramma. In de eindkwalificaties zijn de eisen vanuit het werkveld, de Dublin Descriptoren en de hbo-standaard uit Kwaliteit als Opdracht verwerkt (HBO-raad, 2009). In het afstudeerprogramma maakt de opleiding duidelijk hoe de eindkwalificaties zijn afgeleid van de competenties waarover een beginnende, hbo-gekwalficeerde beroepsbeoefenaar moet beschikken. Sommige opleidingen hanteren de competenties zoals vastgelegd in het landelijk opleidingsprofiel van de opleiding. Andere opleidingen vertalen de competenties in eigen eindkwalificaties om zich te kunnen profileren.

Beoordelingsdimensies

Beoordelingsdimensies geven globaal aan waarop het handelen van de kandidaat en/of resulterende producten daarvan moeten worden beoordeeld. 'Globaal' vanwege het feit dat deze beoordelingsdimensies bruikbaar moeten zijn voor elke prestatie waarmee de bekwaamheid ten aanzien van de kerntaak kan worden aangetoond. Beoordelingsdimensies worden niet rechtstreeks gebruikt om de prestaties van kandidaten mee te beoordelen; ze dienen in de eerste plaats als richtsnoer voor de formulering van prestatiecriteria (zie voor een toelichting bijlage 5).

In het beroepsprofiel van fysiotherapie is een van de benoemde competenties Onderzoeken. Deze competentie is als volgt omschreven: "De fysiotherapeut registreert systematisch patiënt- en behandelgegevens ten behoeve van kwaliteitszorg en onderzoek. Hij neemt deel aan wetenschappelijk onderzoek voor de verdere ontwikkeling van de beroepspraktijk en de wetenschappelijke fundering ervan." (KNGF, 2006, p. 23) Bij de competentie zijn meerdere beoordelingsdimensies aangegeven onder het kopje 'niveau-indicatoren'. Een daarvan is "Is goed ingevoerd in frequent gebruikte methoden en technieken van toegepast wetenschappelijk onderzoek". In het afstudeerprogramma van een opleiding kunnen meerdere prestaties zitten waarin deze beoordelingsdimensie kan worden getoetst. Indien het wordt getoetst in een schriftelijk verslag dient de dimensie vertaald te worden in een prestatie criterium dat verwoordt hoe in een schriftelijk verslag zichtbaar is dat de student hier aan voldoet.

Prestatie

Onder een prestatie wordt het resultaat verstaan van de uitvoering van een beroepsopdracht. De prestatie is een beroepsproduct. In navolging van Losse (Losse, 2012) onderscheidt de expertgroep vijf typen beroepsproducten: 1) advies, 2) ontwerp, 3) fysiek of digitaal product, 4) serie handelingen, 5) sec het antwoord op een onderzoeksvraag. Deze opsomming geeft aan hoe divers in het hbo de opleidingen zijn en daarmee ook de afstudeerprogramma's. Tabel 2 geeft van elk van de onderscheiden beroepsproducten een aantal voorbeelden.

Tabel 2: voorbeelden van beroepsproducten in het hbo

1. Advies	2. Ontwerp	3. Eindproduct	4. Handeling	5. Onderzoek
<i>Mogelijke oplossing of een plan</i>	<i>Verbeelding van een eindproduct</i>	<i>Fysiek of digitale uitwerking van een ontwerp</i>	<i>Professioneel gedrag naar cliënt of publiek</i>	<i>Antwoorden en conclusies</i>
• Organisatieadvies • Pedagogisch advies • Financieel advies	• Bouwontwerp • Bestemmingsplan • Ondernemingsplan	• Schilderij • ICT applicatie • Apparaat • Film • Journalistiek tekst	• Les geven • Voorstelling geven • Hulp verlenen • Verplegen • Therapie geven • Leiding geven • Ondernemen	• Archeologisch rapport • Laboratorium rapport • Forensisch rapport

Vaak wordt er naast het beroepsproduct ook een verantwoording van dat product gevraagd. De student verantwoordt de totstandkoming van het product en onderbouwt de keuzes die hij heeft gemaakt. Bij sommige beroepsproducten, zoals een advies en een

onderzoek, kan de verantwoording een integraal onderdeel vormen van het product. Bij een ontwerp, een fysiek of digitaal product of een handeling, legt de student de verantwoording meestal vast in een separaat verslag. De expertgroep hanteert voor het beroepsproduct en de bijbehorende verantwoording de term *kernwerkstuk*. Indien het maken van een kernwerkstuk helemaal aan het eind van de opleiding wordt gevraagd is er sprake van een *eindwerkstuk*.

Uit de opsomming van beroepsproducten mag duidelijk zijn dat bepaalde producten niet blijvend maar vluchtig zijn en alleen te beoordelen zijn door het handelen te observeren. Procesbeoordeling is sowieso aan te raden aangezien de kwaliteit van het product alleen niet altijd voldoende informatie geeft over de bekwaamheid. Voor het goed toetsen van beroepsbekwaamheid (de 'b' van hbo) is het noodzakelijk dat in een afstudeerprogramma voldoende bewijs van het juiste type (passend bij de te toetsen eindkwalificaties) wordt verzameld. Dat pleit voor een diversiteit aan producten.

Er wordt daarom in het hbo een scala aan bewijsvormen gehanteerd binnen het afstuderen waaronder handelingen binnen stages, ontwerpen, creatieve producten, gebruiksvoorwerpen, dienstverlening, films, composities, voorstellingen etc.. De expertgroep signaleert met zorg echter dat er onder druk van accreditaties de laatste jaren steeds meer nadruk wordt gelegd op de schriftelijke werkstukken. Men kan zich afvragen hoe representatief dit is voor het aantonen van de beroepsbekwaamheid.

Beroepsopdracht

Een beroepsopdracht zet de student ertoe aan om zijn bekwaamheid met betrekking tot een kerntaak van het beroep te demonstreren door het uitvoeren van meer of minder authentieke taken onder meer of minder authentieke omstandigheden, leidend tot een beroepsproduct. Uit deze formulering wordt duidelijk dat beroepsopdrachten kunnen variëren voor wat betreft hun natuurgetrouwheid.

Uiteraard geeft een prestatie die het resultaat is van een authentieke taak uitgevoerd onder authentieke omstandigheden het beste bewijs voor de betreffende beroepsbekwaamheid. Vaak zijn er echter dwingende redenen om af te wijken van deze 'perfect fit', zoals: hoge kosten van apparatuur en/of materialen, gevaar voor de kandidaat en/of diens omgeving, de beperkte beschikbaarheid van opdrachten in de reële werksituatie, de inefficiëntie van de authentieke taak. Een methodemix (combinatie van authentieke en minder authentieke assessmenttaken) zal vaak nodig zijn om efficiënte wijze de informatie te verzamelen die nodig is om met vertrouwen een conclusie te trekken over de bekwaamheid van kandidaten.

Het niveau van beroepsopdrachten dient representatief te zijn voor het niveau van de taken die beginnende professionals moeten kunnen uitvoeren in de beroepspraktijk. Bult-huis (2013) ontwikkelde een systematiek om het niveau van beroepsopdrachten te bepalen op grond van drie factoren:

- de cognitieve complexiteit van de taak;
- de context waarbinnen gehandeld moet worden;
- de zelfstandigheid van rol die de student moet vervullen.

De cognitieve complexiteit wordt doorgaans groter als de taak niet zonder meer uitvoerbaar is door de toepassing van geleerde standaardprocedures, maar vraagt om nieuwe, creatieve oplossingen. De context wordt complexer als er bijvoorbeeld meerdere partijen betrokken zijn bij de beroepsopdracht, die bovendien verschillende belangen kunnen hebben. De zelfstandigheid van een student is van een hoger niveau naarmate de vrijheid van handelen bij het ontwikkelen van de opdracht en de verantwoordelijkheid voor de kwaliteit van het eindresultaat toeneemt. Het verdient aanbeveling dat hogescholen die dezelfde beroepsopleiding aanbieden samen met vertegenwoordigers van de beroepsgroep het gewenste niveau van de beroepsopdrachten vastleggen. Dit met het oog op de vergelijkbaarheid van afstudeerresultaten en diploma's.

Examinator

Examinator is de benaming voor de persoon die belast wordt met de beoordeling van prestaties. Om deugdelijke beoordelingen te geven moeten examinatoren deskundig op:

- A. de te beoordelen bekwaamheid. Studenten zullen niet veel vertrouwen hebben in een beoordeling van een examiner waarvan ze de vakdeskundigheid in twijfel trekken.
- B. het beoordelingsproces zelf. Dit houdt niet alleen in dat de examiner op voorhand kennis heeft genomen van het in te zetten beoordelingsmodel, maar ook dat hij/zij op de hoogte is van aspecten die van invloed zijn op de betrouwbaarheid en validiteit van de beoordeling.

De kwaliteit van beoordeling is voor een groot deel afhankelijk van de kwaliteit van de examiner. Beoordelen is mensenwerk. In de onderwijskundige literatuur bestaat dan ook grote overeenstemming over de noodzaak om examinatoren te trainen en, eventueel, te certificeren. Ook de commissie Bruijn adviseert om docenten toetsdeskundigheid te laten verwerven tot op het niveau van hun betrokkenheid bij het toets- en beoordelingsproces (Commissie externe validering examenkwaliteit hoger beroepsonderwijs, 2012). De expertgroep BKE/SKE heeft dit advies onlangs geconcretiseerd in een programma van eisen (Expertgroep BKE/SKE, 2013).

Beoordelingsmodel¹

Beroepsbekwaamheid moet blijken uit:

- de kwaliteit van een of meer vervaardigde beroepsproducten;
- verantwoordingsverslag(en);
- de kwaliteit van het proces dat tot het beroepsproduct heeft geleid.

Zowel product als proces dienen beoordeeld te worden door een ter zake deskundige, examiner genoemd. Door een beoordelingsmodel in te zetten kan men de beoordelingskwaliteit te bevorderen. Een beoordelingsmodel specificeert de kwaliteit waarop product en proces beoordeeld worden middels prestatiecriteria en beoordelingsschalen (American Educational Research Association, 1999).

Beoordelingsmodellen dienen valide te zijn, voor de student inzichtelijk en ze dienen bij te dragen aan een betrouwbare beoordeling van de geleverde prestaties. Tot slot moet het model ook praktisch hanteerbaar zijn.

Valide

Valide betekent dat het gebruik van het beoordelingsmodel bijdraagt aan een deugdelijke zak-/slaagbeslissing.

Inzichtelijk

Voor de student inzichtelijk betekent dat de prestatiecriteria voorafgaand aan de start van de beroepsopdracht aan de studenten worden gecommuniceerd en dat ze voor de student herkenbaar zijn uit het curriculum.

Betrouwbare beoordeling

Beoordelingsmodellen dragen bij aan een betrouwbare beoordeling wanneer de eindkwalificaties zijn vertaald in observeerbare criteria en ze examinatoren voldoende houvast bieden om tot eenzelfde interpretatie en oordeel te komen.

De score die een examiner toekent, krijgt pas betekenis als die wordt afgezet tegen een kritische score, ook wel cesuur genoemd. De cesuur is niets anders dan de kwantificering van de eindkwalificatie (dat wat de kandidaat moet beheersen). Naast de bepaling van de cesuur zullen voor een transparante beslissingsprocedure ook nog andere uitslagregels vastgelegd moeten worden. Bijvoorbeeld over de mate van compensatie die wordt toegestaan.

¹ De expertgroep volgt hierin de terminologie van de Expertgroep BKE/SKE (2013)

4. Analyse

Opdracht aan de expertgroep is het in kaart brengen van de mogelijkheid en wenselijkheid van bottom-up protocolleren en te adviseren over de wijze waarop protocollen tot stand kunnen komen en aan welke kwaliteitseisen deze moeten voldoen. Het conceptuele model van het afstudeerprogramma uit het vorige hoofdstuk maakt het mogelijk te kijken welke elementen in het afstuderen zich lenen voor protocolleren om de externe validering te vergroten.

In haar analyse heeft de expertgroep de term 'protocolleren' ruim opgevat. Het gaat om maatregelen die de kwaliteit en transparantie van het afstudeerprogramma verhogen door samen te werken. Hieronder valt:

1. Het gezamenlijk ontwikkelen en vaststellen van elementen van het afstudeerprogramma.
2. Het gezamenlijk uitvoeren van elementen van het afstudeerprogramma.
3. Het laten reviewen van elementen van het afstudeerprogramma. Reviewen is het beoordelen van elementen van het afstudeerprogramma aan de hand van kwaliteitscriteria met als doel deze te verbeteren.

Daarnaast heeft de expertgroep overwogen op welk niveau de samenwerking mogelijk is. Daarbij onderscheidt zij vier niveaus:

Niveau 1: Alle opleidingen van alle instellingen;

Niveau 2: Alle opleidingen met hetzelfde opleidingprofiel / CROHO-nummer van alle instellingen;

Niveau 3: Beperkt aantal opleidingen met hetzelfde opleidingprofiel / CROHO-nummer die onderwijsinhoudelijk overeenkomen;

Niveau 4: Binnen één instelling.

Mogelijkheden en wenselijkheid van protocolleren

De mogelijkheden voor protocolleren die de expertgroep als wenselijk bestempelt zijn samengevat in tabel 3. Deze worden hieronder één voor één besproken. Belangrijke overweging bij de analyse is de grote diversiteit aan hbo opleidingen met daarbinnen een grote diversiteit aan afstudeerprogramma's. Er wordt opgeleid tot vele verschillende beroepen met minstens evenzoveel verschillende eindkwalificaties. Ook binnen één opleidingsprofiel signaleert de expertgroep verschillen tussen hogescholen, zowel qua specifieke invulling van de eindkwalificaties als qua afstudeerprogramma's. Dit is onder meer het gevolg van het feit dat opleidingen een bepaalde opvatting hebben over onderwijs, een eigen gezicht willen hebben in de markt en willen inspelen op regionale verschillen.

Niveau 1: A- Gestandaardiseerde wijze van verslaglegging over het afstudeerprogramma als geheel

Op landelijk niveau is er volgens de expertgroep een gestandaardiseerde wijze van verantwoording van afstudeerprogramma's mogelijk. De wijze van verantwoording door een opleiding staat namelijk los van de te toetsen eindkwalificaties en de inrichting van het afstudeerprogramma en kan daarom gelden voor alle opleidingen van alle hogescholen.

Een goede verantwoording van het afstudeerprogramma bevordert de transparantie en draagt bij aan een proces van voortdurende verbetering. Een protocol 'Verbeteren en Verantwoorden van afstuderen in het hbo' kan behulpzaam zijn om te komen tot een dergelijke gestandaardiseerde wijze van verantwoording. In deel 2 van dit rapport doet de expertgroep een voorstel voor een dergelijk landelijk protocol.

De expertgroep erkent dat er een risico zit aan het gebruik van een dergelijk landelijk protocol, namelijk dat het protocol als checklist wordt gebruikt aan de hand waarvan opleidingen worden afgerekend. Op dat moment verliest het protocol een deel van zijn

waarde als verbeterinstrument en bestaat de kans dat opleidingen de werkelijkheid op papier te rooskleurig voorschotelen. De expertgroep ziet een landelijk protocol de komende jaren vooral als een groeimodel dat opleidingen kunnen gebruiken om hun afstuderen en de verantwoording daarover te verbeteren. De Vereniging Hogescholen zou na invoering van een dergelijk protocol na drie jaar kunnen reviewen hoe het staat met de kwaliteit van de verantwoording van de afstudeerprogramma's en op basis daarvan kunnen besluiten tot aanvullende maatregelen.

Tabel 3: Mogelijkheden voor protocolleren bij afstudeerprogramma's

Niveau	Gezamenlijk ontwikkelen & afspreken	Gezamenlijk uitvoeren	Reviewen
Niveau 1: Landelijk alle opleidingen van alle instellingen	<i>A. Op gelijksoortige wijze verantwoorden van het afstudeerprogramma aan de hand van het Protocol Verbeteren en verantwoorden afstuderen in het hbo</i> <i>B. Uitwerken betekenis en minimumeisen Onderzoekend Vermogen</i>		<i>A. Review in 2017 van de kwaliteit van de verantwoording van de afstudeerprogramma's</i>
Niveau 2: Alle opleidingen met hetzelfde opleidingprofiel / CROHO nummer van alle instellingen	<i>C. Gezamenlijk vaststellen body of knowledge & skills Onderzoekend Vermogen in het LOO</i>		<i>D. Gezamenlijk reviewen van de afstudeerprogramma's in het LOO</i>
Niveau 3: Beperkt aantal opleidingen met hetzelfde opleidingprofiel / CROHO nummer een groep opleidingen die onderwijsinhoudelijk overeenkomen	<i>E. Gezamenlijk vaststellen van eindkwalificaties</i> <i>E. Daarna eventueel gezamenlijk vaststellen van de beroepsopdrachten in het afstudeerprogramma</i> <i>E. Daarna eventueel gezamenlijk ontwikkelen en vaststellen van beoordelingsmodellen</i>	<i>F. Inzetten van waarnemers van andere hogescholen tijdens de beoordeling</i> <i>F. Inzetten van examinatoren van andere hogescholen</i>	<i>G. Gezamenlijk reviewen van afstudeerprogramma's</i>
Niveau 4: Binnen één instelling			<i>H. Reviewen van (delen van) het afstudeerprogramma door het werkveld</i> <i>H. Reviewen van (delen van) het afstudeerprogramma door stafmedewerkers en collega-opleidingen</i>

Niveau 1: B- Betekenis en minimumeisen Onderzoekend Vermogen

De expertgroep heeft gemerkt dat er in het land veel onduidelijk bestaat bij opleidingen over een specifiek onderdeel uit de hbo-standaard zoals weergegeven in de nota Kwaliteit als Opdracht (HBO-raad, 2009), namelijk het onderzoekend vermogen. De expertgroep is van mening dat het goed zou zijn wanneer er een landelijke visie wordt ontwikkeld op de rol van onderzoekend vermogen in het hoger beroepsonderwijs met daarin een aantal minimumeisen op het gebied van onderzoekend vermogen. In deel 3 van dit rapport doet de expertgroep hiervoor een voorstel.

Niveau 2: C- Body of knowledge & skills Onderzoekend Vermogen

In Landelijke OpleidingsOverleggen (LOO) worden de opleidingsprofielen ontwikkeld en vastgesteld conform de daarvoor opgestelde procedure (HBO-raad, 2010). De profielen beschrijven de basiscompetenties en veelal ook de kennis en vaardigheden waarover studenten dienen te beschikken in een body of knowledge & skills. Dit landelijke opleidingsprofiel is daarmee een reeds bestaande vorm van protocolleren.

Echter, in veel opleidingsprofielen is niet ingevuld wat de specifieke eisen op het gebied van onderzoekend vermogen zijn. Vaak bestaan er tussen hogescholen verschillende visies. De expertgroep acht het mogelijk en wenselijk dat in de LOO's vastgesteld wordt wat de voorgeschreven body of knowledge & skills is op het gebied van onderzoek (BoK-

SO). Een dergelijke BoKSO beschrijft bijvoorbeeld welke onderzoeksmethoden studenten dienen te beheersen en wat de eisen zijn op het gebied van het gebruik van literatuur. Deel 3 van het rapport beschrijft welke aspecten in de BoKSO kunnen worden meegenomen.

Niveau 2: D- Gezamenlijk reviewen van de afstudeerprogramma's

Gegeven de grote diversiteit tussen opleidingen, zelfs binnen hetzelfde opleidingsprofiel, acht de expertgroep het niet mogelijk en wenselijk om op niveau 2 het afstuderen te standaardiseren. Wel is het mogelijk en wenselijk om in de LOO's aandacht te besteden aan de verschillen en overeenkomsten tussen afstudeerprogramma's. Veel LOO's doen dat al. Daar kan dan ook het gezamenlijk reviewen van de individuele afstudeerprogramma's aan worden gekoppeld. Het Protocol Verbeteren en verantwoorden afstuderen kan daarbij als normkader worden gehanteerd.

Niveau 3: E- Gezamenlijke invulling delen van het afstuderen

De Commissie Bruijn pleit in haar rapport *Vreemde ogen dwingen* voor "Een gestandaardiseerde wijze van verslaglegging omtrent de beoordeling van (eind)werkstukken, stageverslagen, scripties etc." (p. 75²) en verwijst daarbij onder andere naar bijlage 3 van haar rapport waarin het beoordelingsmodel staat weergegeven dat de NVAO-Commissie Onderzoek Hogeschool Inholland (Inholland, 2011) heeft ontwikkeld.

Het standaardiseren van beoordelingsmodellen is volgens de expertgroep alleen mogelijk voor opleidingen die onderwijsinhoudelijk overeenkomen. Dat zijn volgens de expertgroep in het algemeen opleidingen met hetzelfde opleidingsprofiel die een sterk vergelijkbare profilering hebben. Beoordelingsmodellen zijn vrijwel altijd specifiek voor een bepaald opleidingsprofiel en dienen aan te sluiten bij de gehanteerde beoordelingsdimensies en beroepsopdrachten.

Daarmee vervalt naar de mening van de expertgroep de mogelijkheid om een valide beoordelingsmodel voor *alle* eindwerkstukken voor het *gehele* hbo te ontwikkelen. Toch zijn er thans meerdere lijsten met prestatiecriteria in omloop die beogen aan te geven wat een goed kernwerkstuk is. Voorbeeld hiervan is de bovengenoemde lijst van de NVAO-Commissie Onderzoek Hogeschool Inholland die in het land bekend is geworden onder de naam 'de Dunnewijk-criteria'. Deze criteria worden ook wel, maar ten onrechte, 'de NVAO-criteria' genoemd. De expertgroep vindt het belangrijk te benadrukken dat dit beoordelingsmodel is ontwikkeld voor een specifieke situatie waarin een commissie vanuit haar opdracht geen gebruik *mocht* maken van het beoordelingsmodel van de te onderzoeken economische opleidingen. Het beoordelingsmodel is dus heel specifiek en heeft geen officiële landelijke status.

Andere voorbeelden van landelijke beoordelingsmodellen zijn de lijsten die sommige evaluatiebureaus in een aantal gevallen hanteerden bij visitaties. Ook dit zijn algemene modellen die niet zijn afgeleid uit de specifieke te toetsen eindkwalificaties. Dit maakt dat ze alleen bij toeval valide zijn voor de te toetsen kwalificaties. Het feit dat de beoordelingsmodellen van evaluatiebureaus onderling ook nog eens sterk verschillen heeft bijgedragen aan de verwarring. Inmiddels zijn evaluatiebureaus deze modellen aan het heroverwegen.

De expertgroep is van mening dat de beoordelingsmodellen van de opleiding *zelf* leidend dienen te zijn bij de beoordeling alsook bij het vellen van een oordeel over het gerealiseerde eindniveau van een opleiding in het kader van een accreditatie. Het gebruik van andere beoordelingsmodellen dan die van de opleiding is alleen te billijken wanneer het eigen beoordelingsmodel ondeugdelijk is. Een ondeugdelijk beoordelingsmodel is overigens een signaal dat een opleiding het toetsbeleid nog niet helemaal op orde heeft.

De verschillen die tussen hogescholen bestaan binnen één en hetzelfde opleidingsprofiel maken het ontwikkelen van gezamenlijke beoordelingsmodellen op niveau 2 lastig. Een gezamenlijk beoordelingsmodel vereist afspraken over de beroepsopdracht waarvoor het

² Bladzijde nummers verwijzen naar de pdf versie

model wordt gebruikt en over de eindkwalificaties die met die beroepsopdracht worden getoetst. Opleidingen van verschillende hogescholen kunnen verkennen welke clusters van hogescholen op elkaar lijken qua eindkwalificaties en afstudeerprogramma en vervolgens op niveau 3 werken aan gezamenlijke beoordelingsmodellen.

De expertgroep ziet een aantal voordelen van het hanteren van een gezamenlijk beoordelingsmodel op niveau 3:

- Het maken van een goed beoordelingsmodel voor een valide en betrouwbare beoordeling is arbeidsintensief en vereist specifieke (toets)deskundigheid. Bundeling van krachten van meerdere opleidingen kan positieve invloed hebben op de kwaliteit hiervan.
- Door het gezamenlijk ontwikkelen van beoordelingsmodellen is er sprake van externe validering. Peers van hetzelfde type opleiding uit verschillende hogescholen werken mee aan de totstandkoming ervan en bevorderen de kwaliteit door de inbreng van hun verschillende visies.
- Het gebruiken van een gezamenlijk beoordelingsmodel maakt het makkelijker om examinatoren van verschillende hogescholen uit te wisselen. Dit bevordert de externe validering van de beoordeling.
- Wanneer er op deze wijze per type opleiding landelijk een beperkt aantal beoordelingsmodellen ontstaat, wordt het makkelijker en efficiënter om cluster-visitaties te houden.
- Wanneer er een gezamenlijk beoordelingsmodel is, wordt het makkelijker om docenten gezamenlijk te trainen in het begeleiden van beroepsopdrachten en om examinatoren te trainen in het beoordelen van afstudeerprestaties.

De expertgroep signaleert ook een aantal nadelen of risico's:

- Wijzigingen in het beoordelingsmodel hebben veelal gevolgen voor de inhoud en werkwijze in het curriculum. Er is een risico dat de gezamenlijke beoordelingsmodellen te ver af komen te staan van het curriculum waardoor er incongruentie ontstaat. Een gezamenlijk beoordelingsmodel is alleen mogelijk wanneer er gezamenlijkheid is qua eindkwalificaties, beoordelingsdimensies en beroepsopdrachten in het afstudeerprogramma.
- Over verbeteringen in het beoordelingsmodel moet altijd met meerdere hogescholen worden overlegd. Dat is minder efficiënt, kan leiden tot langere procedures en daarmee tot uitstel of zelfs afstel van noodzakelijke verbeteringen.
- Een goed beoordelingsmodel levert niet per definitie betrouwbare beoordelingen op. Daarvoor is ook nodig dat examinatoren deskundig zijn en tot een gemeenschappelijke interpretatie van criteria en beslisregels komen. In het kader hiervan is het belangrijk dat examinatoren periodiek met elkaar afstemmen hoe zij de criteria en beslisregels in de praktijk hanteren in kalibreersessies (in bijlage 6 is een handleiding voor het houden van een kalibreersessie opgenomen). Als een gezamenlijk beoordelingsmodel betekent dat opleidingen elkaars examinatoren inzetten, dan is de consequentie dat kalibreersessies ook gezamenlijk georganiseerd moeten worden. Dit kost tijd, en dus geld. De expertgroep signaleert dat veel opleidingen bovendien sporadisch of nog geen kalibreersessies houden met het eigen team van examinatoren. Het opleidingsoverstijgend organiseren hiervan is dan misschien nog een stap te ver.
- Het eigenaarschap van het beoordelingsmodel moet duidelijk blijven. Elke opleiding blijft verantwoordelijk voor het beoordelingsmodel dat wordt gehanteerd, ook als deze tot stand is gekomen met andere opleidingen. Dat betekent dat de examencommissie en examinatoren de validiteit, betrouwbaarheid en hanteerbaarheid bij audits moeten kunnen verantwoorden. Daarom is draagvlak voor het beoordelingsmodel essentieel. Voorkomen moet worden dat het ontwikkelen van het beoordelingsmodel de verantwoordelijkheid is van een enkele vertegenwoordiger van de deelnemende opleidingen, of uit-

mondt in een compromis dat de kwaliteit van de beoordeling uiteindelijk niet verbetert. Dit stelt hoge eisen aan het ontwikkelproces van het beoordelingsmodel.

Niveau 3: F- Inzetten van waarnemers en examinatoren

Opleidingen kunnen op niveau 3 docenten van elkaar inzetten om als waarnemer te fungeren tijdens het afstudeerproces. Zij kunnen aanwezig zijn wanneer studenten afstudeerprestaties leveren of wanneer prestaties worden beoordeeld. Opleidingen die eenzelfde beoordelingsmodel hanteren kunnen ook examinatoren uitwisselen. De expertgroep adviseert overigens wel dat deze examinatoren ook gezamenlijke kalibreersessies houden om tot een zelfde interpretatie van het beoordelingsmodel te komen.

Niveau 3: G- Gezamenlijk reviewen van afstudeerprogramma's

Op niveau 3 is het ook mogelijk om afstudeerprogramma's van elkaar te reviewen. Dit kan waarschijnlijk grondiger dan op niveau 2. De expertgroep acht het mogelijk en wenselijk als clusters van een aantal opleidingen periodiek de onderdelen van afstudeerprogramma's tegen het licht houden, bijvoorbeeld aan de hand van de vragen uit het protocol 'Verbeteren en Verantwoorden van Afstuderen in het hbo'.

Niveau 4: H- Reviewen van afstudeerprogramma's door collega's of het werkveld

Op het niveau van individuele opleidingen gebeurt op dit moment al veel om de kwaliteit van afstuderen te verbeteren. Er wordt veelal gewerkt met twee examinatoren, er worden beoordelaars uit het werkveld in het proces betrokken, en er zijn interne visitaties. In aanvulling daarop pleit de expertgroep voor het laten reviewen van afstudeerprogramma's door het werkveld, bijvoorbeeld binnen de beroepenveldcommissie of door collega's. Dit laatste kan bijvoorbeeld vorm worden gegeven door in de interne visitaties nadrukkelijk het afstudeerprogramma te reviewen aan de hand van de vragen uit het protocol.

Het tot stand komen van protocollen

De expertgroep is ook gevraagd te adviseren over het tot stand komen van protocollen. Opleidingen die op niveau 2 of 3 willen toewerken naar protocolleren van afstudeerprogramma's zouden de volgende stappen kunnen zetten:

1. In kaart brengen van de verschillende afstudeerprogramma's van de hogescholen en het analyseren van overeenkomsten en verschillen qua eindkwalificaties, beoordelingsdimensies, beroepsopdrachten en beoordelingsmodellen.
2. Indien nog niet gedaan: gezamenlijk ontwikkelen van een visie op de rol van onderzoekend vermogen in het beroep en de opleiding en het opstellen van een BoKSO.
3. Gezamenlijk reviewen van beroepsproducten uit het afstuderen om verschillen en overeenkomsten in beroepsopdrachten en in het gebruik van beoordelingsmodellen te onderzoeken.
4. Harmoniseren van de onderdelen van het afstudeerprogramma.
5. Gezamenlijk ontwikkelen van een beoordelingsmodel voor onderdelen van het afstudeerprogramma die opleidingen gemeenschappelijk hebben.

5. Conclusies en aanbevelingen

De Vereniging Hogescholen heeft de expertgroep Protocol gevraagd een antwoord te geven op de vraag wat de wenselijkheid en mogelijkheid is van een gezamenlijk, bottom-up opgesteld protocol of protocollen voor het beoordelen van (kern)werkstukken en te adviseren over de wijze waarop deze tot stand dienen te komen en aan welke kwaliteitseisen deze dienen te voldoen.

Overwegingen

De expertgroep komt tot de volgende overwegingen:

1. Afstudeerprogramma's in het hbo zijn zeer divers. Tussen opleidingsprofielen bestaan grote verschillen in eindkwalificaties en inrichting van afstudeerprogramma's. Ook binnen hetzelfde opleidingsprofiel zijn er verschillen tussen hogescholen.
2. Opleidingen in het hbo leiden op tot beroepsbekwaamheid. Dat men gekwalificeerd is voor een bepaald beroep moet blijken uit het feit dat men de kerntaken van dat beroep op adequate wijze kan uitvoeren. Toetsing daarvan vraagt om de inzet van een scala aan toetsvormen binnen het afstudeerprogramma waarvan het werken met (kern)werkstukken er slechts één is.
3. Het is wenselijk dat verschillende hogescholen voor hetzelfde type opleiding streven naar meer overeenkomstige afstudeerprogramma's, zowel qua eindkwalificaties, verstrekte beroepsopdrachten als beoordelingsmodellen. Dit mag echter niet ten koste gaan van de gewenste profilering die opleidingen nastreven.
4. De kwaliteit van beoordeling is voor het grootste deel afhankelijk van de kwaliteit van de examinerator. Beoordelen is mensenwerk.
5. De beoordelingsmodellen die binnen het afstuderen gebruikt worden dienen valide en inzichtelijk te zijn, en dienen bij te dragen aan een betrouwbare beoordeling. Dit vereist onder andere dat de prestatiecriteria zijn afgeleid uit de te toetsen eindkwalificaties en ook dat ze zijn afgestemd op de verstrekte beroepsopdrachten. Opleidingen die willen werken aan gezamenlijke beoordelingsmodellen zullen dus eerst overeenstemming moeten bereiken over dezelfde eindkwalificaties en beroepsopdrachten.
6. Er zijn landelijke beoordelingsmodellen in omloop. Deze hebben geen officiële status en kunnen hooguit bij toeval valide zijn voor de te toetsen kwalificaties. Het gebruik van deze modellen bij zak-/slaagbeslissingen en bij visitaties dient zoveel mogelijk te worden vermeden.

Conclusie

Deze overwegingen leiden tot het volgende antwoord op de vraag die aan de expertgroep is gesteld. Het is *niet* wenselijk en mogelijk om te gaan werken met één landelijk, voor alle opleidingen geldend protocol, indien protocol wordt uitgelegd als een beoordelingsmodel voor afstudeerprestaties (waaronder kernwerkstukken) als omschreven in hoofdstuk 3.

Landelijke beoordelingsmodellen per opleidingsprofiel / croho-nummer zijn alléén mogelijk voor opleidingen die onderwijsinhoudelijk overeenkomen, dat wil zeggen opleidingen die werken met dezelfde eindkwalificaties, beoordelingsdimensies en afstudeerprogramma's. Dat zijn er op dit moment niet veel.

Andere vormen van protocollering zijn *wél* wenselijk en mogelijk, te weten:

- Op landelijk niveau (niveau 1) is het wenselijk en mogelijk om tot afspraken te komen over een eenvormige wijze van verantwoording van afstudeerprogramma's en om tot een visie te komen op de betekenis van onderzoekend vermogen. Delen 2 en 3 van dit rapport bevatten hiervoor voorstellen.

- Op niveau van opleidingsprofielen (niveau 2) is het wenselijk en mogelijk dat wordt gewerkt aan invulling van de body of knowledge & skills voor onderzoek. Op dit niveau kunnen ook afstudeerprogramma's worden gereviewd.
- Op niveau van een beperkt aantal opleidingen met het zelfde opleidingsprofiel die onderwijsinhoudelijk overeenkomen (niveau 3) is het mogelijk (delen) van het afstudeerprogramma gezamenlijk te ontwikkelen. Of dit wenselijk is hangt af van de vraag of de gezamenlijkheid niet ten koste gaat van de individuele profilering.
- Op het niveau van de individuele opleiding (niveau 4) is het wenselijk en mogelijk dat het afstudeerprogramma wordt gereviewd door het werkveld of door collega's binnen de hogeschool.

Aanbevelingen

De expertgroep adviseert de Vereniging Hogescholen op basis van deze conclusies het volgende:

1. Stel het protocol 'Verbeteren en Verantwoorden van het afstuderen in het hbo' uit deel 2 vast. Hierin is onder andere de eis opgenomen dat opleidingen hun afstudeerprogramma periodiek laten reviewen op minimaal niveau 4. Hanteer dit model voortsnog als groeimodel dat opleidingen kunnen gebruiken om hun afstudeerprogramma's en de verantwoording daarover te verbeteren. Onderzoek over drie jaar de implementatie van het protocol en de kwaliteit van de verantwoording, alsmede de mogelijkheden om tot verdere protocollering van afstuderen te komen.
2. Geef duidelijkheid over de rol van onderzoekend vermogen in het hoger beroepsonderwijs. Geef tevens aan wat het minimumniveau is voor het beheersen van onderzoekend vermogen op bachelor- en masterniveau. Neem hiervoor het voorstel uit deel 3 over.
3. Bepleit bij de NVAO, de evaluatiebureaus en de zelfstandige secretarissen een uniforme wijze van werken bij het beoordelen van standaard 3/16, 2^e deel waarin:
 - a. er gebruik wordt gemaakt van de beoordelingsmodellen van de opleiding, tenzij deze van onvoldoende kwaliteit zijn;
 - b. er bij de visitatie gekeken wordt naar ál het bewijs dat in het afstudeerprogramma wordt verzameld en niet alleen naar de eindwerkstukken.
4. Stimuleer bij opleidingen gezamenlijkheid in afstudeerprogramma's, zonder dit verplichtend op te leggen. Start hiervoor bij de LOO's³ en vraag hen invulling te geven aan het gezamenlijk vaststellen van de body of knowledge & skills onderzoekend vermogen en het reviewen van de afstudeerprogramma's.

Consequenties van de aanbevelingen

Bovengenoemde aanbevelingen hebben consequenties voor LOO's, opleidingen en visitatiepanels. De LOO's krijgen een nadrukkelijke rol bij het bepalen van het eindniveau op het gebied van onderzoekend vermogen en bij het zoeken naar mogelijkheden om tot gezamenlijkheid te komen.

Opleidingen zijn thans al druk bezig met het verbeteren van hun toetsbeleid. Zij zullen nog nadrukkelijker dan thans het geval is daarin aandacht moeten besteden aan de invulling van het afstudeerprogramma. De expertgroep schat in dat dit voor veel opleidingen onder andere betekent dat zij hun beoordelingsmodellen moeten aanscherpen.

Visitatiepanels kunnen niet langer volstaan met het beoordelen van 15 eindwerkstukken aan de hand van een eigen beoordelingsmodel zoals in het verleden wel voor kwam. Voor een goede beoordeling van standaard 3/16, 2^e deel zullen zij meer bewijsmateriaal moeten bestuderen, hetgeen uiteraard ook meer tijd zal kosten. Daar staat tegenover dat op

³ Commerciële aanbieders van hbo-onderwijs zijn op dit moment nog niet vertegenwoordigd in de LOO's. Bij een grotere rol van LOO's in het protocolleren van het afstuderen dient hier een oplossing voor gevonden te worden.

het moment dat een protocol voor het verbeteren en verantwoorden afstuderen in het hbo heeft geleid tot een sterke verbetering van de kwaliteit en de verantwoording van het afstudeerprogramma's, het systeem mogelijk zo veranderd zou kunnen worden dat visitatiepanels alleen nog maar de uitvoering van het afstudeerprogramma via een systeemaudit hoeven te beoordelen en niet meer het gerealiseerde eindniveau zelf.

DEEL 2: PROTOCOL VERBETEREN EN VERANTWOOR- DEN VAN AFSTUDEREN IN HET HBO

Protocol als middel voor systematische
kwaliteitszorg en externe validering
van afstuderen

Inleiding

De expertgroep heeft in deel 1 geconstateerd dat het mogelijk en wenselijk is om op landelijk niveau tot afspraken te komen over de wijze waarop opleidingen zich over hun afstudeerprogramma verantwoorden. Een goede verantwoording van het afstudeerprogramma bevordert de transparantie en draagt bij aan een proces van voortdurende verbetering. Een protocol 'Verbeteren en Verantwoorden van Afstuderen in het hbo' kan behulpzaam zijn om te komen tot een dergelijke gestandaardiseerde wijze van verantwoording.

De bedoeling van dit protocol is om opleidingen in het hbo te helpen samen werken aan het valide, betrouwbaar en voor studenten inzichtelijk toetsen van de beoogde eindkwalificaties, zoals beoogd in standaard 3/16 van de NVAO en het rapport 'Vreemde ogen dwingen' van de commissie Bruijn.

Gegeven de grote variëteit aan opleidingen en afstudeerprogramma's zal dit protocol vooral beschrijven op welke punten opleidingen zich dienen te verantwoorden over de kwaliteit van het afstudeerprogramma en waarop ze zich mogelijk ook kunnen verbeteren.

Om dit te bereiken heeft het protocol drie doelen:

1. Duidelijkheid scheppen over terminologie en inrichting van het afstuderen.
2. Een kader scheppen voor gezamenlijk leren en verbeteren.
3. Een kader scheppen voor borging en verantwoording.

In deze doelen zit iets paradoxaals: Borgen en verantwoorden suggereert dat precies duidelijk is waar de lat ligt, terwijl gezamenlijk leren en verbeteren ook over het vinden en verhogen van de lat gaat. Bovendien suggereert het leggen van een lat dat alles wat er onder valt niet goed is, terwijl daar juist de ruimte voor ontwikkeling zit. Gelukkig maakt het verschil tussen voldoende, goed en excellent in de beoordelingskaders van de NVAO het mogelijk hier rekening mee te houden (zie bijlage 4).

Het protocol bevat daarom een opsomming van vragen over het afstudeerprogramma waar een opleiding een goed antwoord op moet hebben om een goede beoordelings- en beslissingskwaliteit in het afstuderen te garanderen. In een verantwoording van het afstudeerprogramma beschrijft en verantwoordt de opleiding op inzichtelijke wijze voor alle betrokkenen:

- de eindkwalificaties die worden getoetst,
- op welke manier toetsing plaatsvindt, in het bijzonder de beroepsopdrachten die worden uitgevoerd,
- door wie en op welke manier prestaties worden beoordeeld, en
- hoe de informatie die het afstudeerprogramma oplevert wordt geaggregeerd tot een zak-/ slaagbeslissing.

Tegelijkertijd is het protocol geen checklist aangezien er vaak meerdere goede methoden zijn om de stappen van het protocol in te vullen. Het is aan de opleiding en hun (staf)adviseurs om hier keuzes in te maken die afhankelijk zijn van het profiel van de opleiding maar ook van de organisatorische context. Daarbij verbeteren opleidingen en (staf)adviseurs de gebruikte methoden of ontdekken nieuwe mogelijkheden. Dit heeft weer nieuwe keuzevragen en verbeteracties tot gevolg. Dit hele proces past in de door de NVAO gewenste systematische kwaliteitszorg en een consequent verbeterbeleid. Hierdoor draagt het werken met het protocol bij aan een groeiende kwaliteitscultuur rond afstudeerprogramma's.

De vragen in het protocol zijn afgeleid uit standaard 1 en 3/16 van de NVAO (NVAO, 2011) en de relatie tussen die twee:

Standaard 1:

"De beoogde eindkwalificaties van de opleiding zijn wat betreft inhoud, niveau en oriëntatie geconcretiseerd en voldoen aan internationale eisen."

"De beoogde eindkwalificaties passen wat betreft niveau en oriëntatie (bachelor of master; hbo of wo) binnen het Nederlandse kwalificatieraamwerk. Ze sluiten bovendien aan bij de actuele eisen die in internationaal perspectief vanuit het beroepenveld en het vakgebied worden gesteld aan de inhoud van de opleiding."

Standaard 3/16 (1^e deel):

"De opleiding beschikt over een adequaat systeem van toetsing."

"De toetsen en de beoordeling zijn valide, betrouwbaar en voor studenten inzichtelijk."

Daarnaast is bij het opstellen van het protocol gebruik gemaakt van theorie over toetsen en van beschrijvingen van good practices van diverse opleidingen. Het protocol bevat daardoor nauwelijks nieuwe eisen voor opleidingen ten opzichte van de huidige situatie. De enige uitzonderingen hierop zijn dat het protocol helderheid eist over het specificeren van eindkwalificaties voor onderzoekend vermogen en de eis om 'vreemde ogen' het afstudeerprogramma te laten reviewen.

Op welke manier opleidingen ervoor zorgen dat de vragen uit het protocol worden beantwoord, is aan de opleidingen zelf. De expertgroep heeft voor twee punten een handreiking geschreven om opleidingen daarbij te helpen. Bijlage 5 gaat over het ontwikkelen van goede beoordelingsmodellen en bijlage 6 gaat over het organiseren van kalibreersessies.

Wat betreft het beoordelen zelf maakt het protocol geen onderscheid tussen kernwerkstukken en andere prestaties die binnen het afstudeerprogramma worden beoordeeld. De expertgroep is van mening dat alle beoordelingen die deel uitmaken van het afstudeerprogramma aan dezelfde kwaliteitsstandaarden moeten voldoen.

Hoewel de expertgroep zich richt op de bacheloropleidingen in Nederland, is gedurende het ontwikkeltraject gebleken dat zowel de Associate degree-programma's als de hbo-masteropleidingen met dezelfde problematiek worstelen. Het protocol is volgens de expertgroep ook toe te passen op de afstudeerprogramma's van deze opleidingen.

Protocol

Eindkwalificaties

1. Weerspiegelen de eindkwalificaties van de opleiding zowel de eisen vanuit het werkveld als de eisen aan het hbo-niveau?
2. Heeft de opleiding het vereiste niveau voor onderzoekend vermogen beschreven in de eindkwalificaties?
3. Toetst de opleiding alle eindkwalificaties in het afstudeerprogramma en, indien het meerdere afstudeeronderdelen betreft, is het duidelijk waar wat getoetst wordt?

Beroepsopdrachten

4. Zijn de beroepsopdrachten die studenten uitvoeren in het afstudeerprogramma geschikt voor het aantonen van de te verwerven eindkwalificaties?
5. Bewaakt de opleiding dat de complexiteit van de beroepsopdrachten voor alle studenten dezelfde is en dat er een vergelijkbare mate van zelfstandigheid van hen verwacht wordt bij de uitvoering hiervan?

Beoordeling

6. Waarborgt de opleiding dat elke examinerator bekwaam is om een onderbouwd oordeel van de prestaties van studenten tot stand te brengen?
7. Bieden de gehanteerde beoordelingsmodellen voldoende garantie op een valide, betrouwbare en transparante beoordeling en zijn ze tegelijkertijd werkbaar voor examinatoren?
8. Borgt de opleiding een gemeenschappelijke interpretatie van de beoordelingsmodellen door de examinatoren?
9. Is de beoordelingsprocedure voor alle betrokkenen transparant en werkbaar en bevordert deze een zo betrouwbaar mogelijke beoordeling?

Randvoorwaarden

10. Zorgt de opleiding ervoor dat het afstudeerprogramma door alle betrokkenen binnen de beschikbare tijd en mogelijkheden uitgevoerd kan worden?

Verantwoording en ontwikkeling kwaliteit

11. Geeft de opleiding invulling aan het 'vreemde ogen'-principe om de kwaliteit van het afstudeerprogramma aantoonbaar te bevorderen?
12. Overlegt de opleiding bij een visitatie bewijzen die samen een transparant en representatief beeld geven van het afstudeerprogramma en het gerealiseerde eindniveau van studenten?

Uitwerking en toelichting

Eindkwalificaties

Een overzicht van de te toetsen eindkwalificaties vormt het startpunt van elk afstudeerprogramma. Uitgangspunt is dat de opleiding alle eindkwalificaties van de opleiding in het afstudeerprogramma toetst.

1. Weerspiegelen de eindkwalificaties van de opleiding zowel de eisen vanuit het werkveld als de eisen aan het hbo-niveau?

In de eindkwalificaties zijn de eisen vanuit het werkveld, de Dublin Descriptoren en de hbo-standaard uit Kwaliteit als Opdracht verwerkt. Een landelijk opleidingsprofiel vormt meestal de basis voor deze eindkwalificaties, maar opleidingen geven hier vaak een eigen inkleuring aan. Behalve dat de eindkwalificaties relevant en op het hbo-niveau dienen te zijn, is het ook belangrijk dat ze bruikbaar zijn voor de vormgeving van leeractiviteiten en toetsen.

2. Heeft de opleiding het vereiste niveau voor onderzoekend vermogen beschreven in de eindkwalificaties?

De hbo-standaard uit Kwaliteit als Opdracht kent vier onderdelen: een gedegen theoretische basis, onderzoekend vermogen, professioneel vakmanschap en beroepsethiek & maatschappelijke oriëntatie. Van deze vier is onderzoekend vermogen het onderdeel waarover de grootste onduidelijkheid bestaat terwijl het grote invloed heeft op het te realiseren eindniveau. Daarom formuleert de opleiding een visie op de rol van onderzoekend vermogen in de opleiding en in het afstuderen en verwerkt het vereiste niveau in de eindkwalificaties (zie deel 3 van dit rapport).

3. Toetst de opleiding alle eindkwalificaties in het afstudeerprogramma en, indien het meerdere afstudeeronderdelen betreft, is het duidelijk waar wat getoetst wordt?

Bij afstuderen is het uiteindelijke doel van de beoordeling vast te stellen of de student voldoende bekwaam is om professioneel en op hbo-niveau te handelen in taaksituaties die representatief zijn voor de kerntaken van het beroep waarvoor deze wordt opgeleid. Die taaksituaties en de te beoordelen eindkwalificaties zijn veelal zo divers dat één prestatie daarvoor niet representatief kan zijn. Dit pleit voor een afstudeerprogramma dat uit meerdere onderdelen bestaat en waarin de student in verschillende taaksituaties moet handelen. Stages zijn hiervoor vaak bij uitstek geschikte gelegenheden die soms nog maar beperkt worden benut. Voorwaarde voor het gebruik van stages als onderdeel van het afstudeerprogramma is dat de competenties op het vereiste eindniveau worden getoetst.

De opleiding moet kunnen aangeven welke eindkwalificaties in welke afstudeeronderdelen worden getoetst en moet deze keuze kunnen verantwoorden. Het toetsen van alle eindkwalificaties in het afstudeerprogramma is alleen mogelijk wanneer de opleiding een realistisch aantal eindkwalificaties heeft gedefinieerd.

Beroepsopdrachten

4. Zijn de beroepsopdrachten die studenten uitvoeren in het afstudeerprogramma geschikt voor het aantonen van de te verwerven eindkwalificaties?

Het gaat hierbij om de volgende vragen:

- Hebben de beroepsopdrachten waarmee de beoogde eindkwalificaties moeten worden aangetoond voldoende gelijkenis met de taaksituaties waarvoor een beginnend beroepsbeoefenaar zich gesteld ziet in de werkelijke beroepspraktijk?

tijk? Dit betreft de authenticiteit van beroepsopdrachten en de situatie waarin ze worden uitgevoerd.

- Worden de eindkwalificaties in voldoende mate gedekt door de beroepsopdrachten?
- Is sprake van voldoende opdrachten om een uitspraak te kunnen doen over de bekwaamheid van een student?
- Zijn de uitgevoerde beroepsopdrachten samen voldoende representatief voor alle mogelijke opdrachten die studenten voorgelegd zouden kunnen krijgen? Anders gezegd: Is het aannemelijk dat een student dezelfde beoordeling zou hebben gekregen ten aanzien van de betreffende eindkwalificatie als hij iets andere beroepsopdrachten had moeten uitvoeren?
- Is de zelfstandigheid waarmee de student uitvoering geeft aan beroepsopdrachten vergelijkbaar met die van een beginnend professional?

5. Bewaakt de opleiding dat de complexiteit van de beroepsopdrachten voor alle studenten dezelfde is en dat er een vergelijkbare mate van zelfstandigheid van hen verwacht wordt bij de uitvoering?

Omdat studenten in de afstudeerfase niet allemaal dezelfde beroepsopdrachten uitvoeren (of niet in dezelfde context) en zij vaak zelf een opdrachtgever moeten verwerven, is het belangrijk dat de opleiding de kwaliteit van de beroepsopdrachten goed bewaakt. Opdrachten dienen tijdig beoordeeld te worden op complexiteit en zelfstandigheid.

Beoordeling

Beoordelen is in de eerste plaats het werk van vakbekwame professionals. Examinatoren hebben een belangrijke maatschappelijke opdracht. Zij beoordelen de bekwaamheid van afstudeerders en bepalen daardoor mede de kwaliteit van de beginnende beroepsbeoefenaar op de arbeidsmarkt. Beoordelingsmodellen zijn daarbij een hulpmiddel.

Terugkerend vraagstuk bij het ontwikkelen van beoordelingsmodellen is de mate waarin een opleiding ruimte biedt aan de professionaliteit van de examinatoren. Zowel het volledig vertrouwen op het 'timmermansoog' van de examinerator als het 'dichttimmeren' van procedures en criteria is ongewenst. Het is belangrijk dat examinatoren deskundigen zijn die in staat worden gesteld hun – door jarenlang opgebouwde expertise en ervaring ontwikkelde – 'timmermansoog' in te zetten bij de beoordelingen.

Tegelijkertijd is het nodig de transparantie en herleidbaarheid van beslissingen te vergroten. De expertise en ervaring van de examinatoren is nodig voor de ontwikkeling van duidelijke prestatiecriteria, beoordelingsprocedures en beoordelingsformulieren, eenduidige en professionele uitvoering van de beoordeling, alsmede kalibratie met collega-examinatoren om tot de juiste interpretatie van criteria en de prestaties van studenten te komen. Dit leidt tot de onderstaande vragen.

6. Waarborgt de opleiding dat elke examinerator aantoonbaar bekwaam is in het beoordelen van prestaties van studenten?

Examinatoren hebben de taak om de prestaties van studenten te analyseren en te interpreteren, leidend tot het formuleren van onderbouwde oordelen. Omdat beoordelingsmodellen voor complexe prestaties nooit op zo'n manier uitgewerkt kunnen worden dat verschil in interpretatie uitgesloten kan worden, is de deskundigheid van examinatoren cruciaal voor het vertrouwen in een betrouwbare beoordeling.

- Beschikken examinatoren over de inhoudelijke deskundigheid om de prestaties van studenten te beoordelen? Hebben zij, zoals de Expertgroep BKE/SKE (2013, p. 28) aangeeft 'een doorleefd beeld van wat in een bepaalde professie aan kennis en vaardigheden wordt verwacht en nodig is'?
- Zijn de examinatoren bekwaam in de toetsvormen die gebruikt worden bij het beoordelen van de eindkwalificaties en zijn ze zich bewust van de beoordelarseffecten die kunnen optreden (zie bijlage 3)?

- Voldoen de examinatoren aan de Basiskwalificaties Examinatoren zoals de Expertgroep BKE/SKE (2013) heeft vastgesteld?

7. Bieden de gehanteerde beoordelingsmodellen voldoende garantie op een valide, betrouwbare en transparante beoordeling en zijn ze tegelijkertijd werkbaar voor examinatoren?

Beroepsbekwaamheid moet blijken uit de kwaliteit van een (of meer) vervaardigd(e) beroepsproduct(en) en uit de kwaliteit van het proces dat tot het beroepsproduct heeft geleid. Beide dienen beoordeeld te worden door een ter zake deskundige examinator. Door een beoordelingsmodel in te zetten kan de beoordelingskwaliteit worden bevorderd.

Een beoordelingsmodel specificceert:

- de prestatiecriteria die worden gebruikt om de kwaliteit van product en proces te beoordelen;
- de beoordelingsschalen (normering) waarmee kwaliteitsverschillen tussen studenten uitgedrukt worden;
- de wijze waarop de resultaten die door studenten zijn behaald worden geaggregeerd tot een eindresultaat;
- de regels voor een zak- / slaagbeslissing.

In het kader van een valide, betrouwbare en transparante beoordeling zijn de volgende vragen aan de orde:

- Zijn de prestatiecriteria die in het beoordelingsmodel worden gebruikt op een goede en navolgbare manier afgeleid van de in het afstudeerprogramma te beoordelen eindkwalificaties?
- Zijn de eindkwalificaties die in de afstudeerwerken beoordeeld worden, zoveel mogelijk uitgewerkt in observeerbare prestatiecriteria?
- Is duidelijk hoe de scores op prestatiecriteria van één beroepsproduct gecombineerd worden tot conclusies over de kwaliteit van een beroepsproduct?
- Is duidelijk hoe de beoordelingen van een reeks prestaties gecombineerd worden tot een eindoordeel over de verwerving van de eindkwalificatie(s)?
- Is duidelijk bij welke (geaggregeerde) score de grens tussen zakken of slaagen ligt?
- Is het beoordelingsmodel praktisch hanteerbaar voor examinatoren om tot een oordeel en de onderbouwing daarvan te komen?
- Maakt het beoordelingsmodel aan alle betrokkenen (studenten, examinatoren en derden) inzichtelijk wat het oordeel is over de kwaliteit van een prestatie, waarom dit oordeel is toegekend en hoe dit tot stand is gekomen?

Voor het maken van beoordelingsmodellen heeft de expertgroep een handreiking ontwikkeld (zie bijlage 5).

8. Borgt de opleiding een gemeenschappelijke interpretatie van de beoordelingsmodellen door de examinatoren?

Zijn de examinatoren in staat om tot eenzelfde interpretatie van een oordeel over een prestatie te komen? Vinden er bijvoorbeeld regelmatig bijeenkomsten plaats waar de examinatoren aan de hand van authentiek materiaal de deugdelijkheid van hun beoordelingen checken en eventueel (de toelichting bij) prestatiecriteria vervangen of bijstellen ('kalibreersessies')? De expertgroep heeft hiervoor een Handreiking Kalibreersessies ontwikkeld (zie bijlage 6).

9. Is de beoordelingsprocedure voor alle betrokkenen transparant en werkbaar en bevordert deze tegelijkertijd een zo betrouwbaar mogelijke beoordeling?

- Zijn de rollen van de verschillende betrokkenen bij de beoordeling van afstudeerprestaties duidelijk? Is duidelijk wie op welk moment een rol heeft in het

beoordelingsproces, en wat de status van een eventuele (praktijk)begeleider en/of opdrachtgever hierbij is?

- Worden er minstens twee examinatoren ingezet voor de beoordeling van prestaties van een student?
- Is de beoordelingsprocedure zo ingericht dat ongewenste beoordelaarseffecten worden bestreden (zie bijlage 3)?
- Zijn de examinatoren zo onafhankelijk mogelijk? Dat betekent dat ten minste één examiner niet betrokken is geweest bij de begeleiding van afstudeeronderdelen.
- Zijn er afspraken gemaakt over wat te doen bij (grote) discrepantie in oordelen tussen twee examinatoren? Is duidelijk wanneer eventueel een onafhankelijke derde examiner moet worden ingeschakeld?

Randvoorwaarden

10.Zorgt de opleiding ervoor dat het afstudeerprogramma door alle betrokkenen binnen de beschikbare tijd en mogelijkheden uitgevoerd kan worden?

Het is de verantwoordelijkheid van het opleidingsmanagement om alle betrokkenen voldoende te faciliteren. Voor studenten geldt:

- Is sprake van voldoende tijd en de juiste ondersteuning om een de verlangde prestaties te kunnen leveren?
- Heeft het onderwijs voorafgaand aan het afstudeerprogramma hen voldoende voorbereid om zelfstandig de beroepsopdrachten uit te voeren die nodig zijn om de eindkwalificaties aan te tonen?

Voor examinatoren geldt:

- Is er voldoende tijd, zowel in aantal uren als in doorlooptijd, om prestaties van studenten te beoordelen en een correcte naleving van beoordelingsprocedures te bewerkstelligen? Als er onder een te grote tijdsdruk wordt gewerkt, is de kans groot dat examinatoren niet alle informatie even zorgvuldig wegen, dat ze niet eerst individueel tot een weloverwogen oordeel komen, of onvoldoende onderbouwing geven.
- Worden zij voldoende gefaciliteerd, in tijd en middelen, om de benodigde deskundigheid te verwerven en op peil te houden? Beoordelen is moeilijk en er is veel training en ervaring voor nodig om te groeien van beginnend beoordelaar naar expertbeoordelaar (Expertgroep BKE/SKE, 2013).

Verantwoording en ontwikkeling kwaliteit

11.Geeft de opleiding invulling aan het 'vreemde ogen'-principe om de kwaliteit van het afstudeerprogramma aantoonbaar te bevorderen?

Andere mensen laten meekijken bij het opzetten en uitvoeren van het afstudeerprogramma kan de kwaliteit ervan bevorderen. 'Vreemde ogen' geven vaak een positieve impuls aan kwaliteitsontwikkeling. Zoals in deel 1 is aangegeven zijn er meerdere mogelijkheden:

- Reviewen: het beoordelen van het afstudeerprogramma aan de hand van kwaliteitscriteria met als doel deze te verbeteren;
- gezamenlijk uitvoeren;
- gezamenlijk ontwikkelen & afspreken.

Dit kan op meerdere niveaus:

- Niveau 1: Landelijk alle opleidingen van alle instellingen;
- Niveau 2: Alle opleidingen met hetzelfde opleidingprofiel / CROHO-nummer van alle instellingen;

- Niveau 3: Beperkt aantal opleidingen met hetzelfde opleidingprofiel;
- Niveau 4: Binnen één instelling.

Hoe geeft de opleiding hieraan invulling? Is minimaal sprake van het periodiek laten reviewen van het afstudeerprogramma op niveau 4?

12. Overlegt de opleiding bij een visitatie bewijzen die samen een transparant en representatief beeld geven van het afstudeerprogramma en het gerealiseerde eindniveau van studenten?

Visitaties zijn een belangrijk instrument om de kwaliteit van afstudeerprogramma's tegen het licht te houden en te verbeteren. Opleidingen worden iedere zes jaar extern geïnspecteerd waarna een beslissing volgt over de accreditatie. Daarnaast organiseren veel hogescholen tussentijdse interne visitaties.

Onderdeel van de visitatie is een oordeel over het toetsbeleid van de opleiding. Dit betreft dus ook een oordeel over de opzet en motivatie van het afstudeerprogramma. De NVAO eist dat de opleiding beschikt over een adequaat systeem van toetsing. Wat onder 'adequaat' wordt verstaan bij de toetsing van eindkwalificaties is naar de mening van de expertgroep in dit protocol gespecificeerd. Dat betekent dat de opleiding bij een visitatie een verantwoording kan aanleveren van de opzet van het afstudeerprogramma langs de lijnen van dit protocol.

In het kader van de verantwoording is het niet voldoende dat de informatie over het afstuderen op papier deugt. Uiteindelijk moet zichtbaar zijn dat de uitvoering overeenkomt met wat afgesproken is. Daaruit blijkt dat procedures en instrumenten werkbaar zijn, examinatoren deskundig zijn en alle betrokkenen tijdig en volledig zijn geïnformeerd. Beoordelen blijft mensenwerk. Openheid in de uitvoering wordt bevorderd als het opleidingsteam achter de uitgangspunten en opzet staat. Het opleidingsteam betrekken bij de totstandkoming van het afstudeerprogramma is één van de manieren om dit te realiseren.

Naast een oordeel over het systeem van toetsing wordt bij een visitatie ook een oordeel geveld over het gerealiseerde eindniveau. De opleiding toont aan dat de beoogde eindkwalificaties worden gerealiseerd. Dat betekent dat beslissingen van de opleiding over de beroepsbekwaamheid van studenten herleidbaar moeten zijn voor derden. Deze standaard wordt in de praktijk geoperationaliseerd door een herbeoordeling door het visitatiepanel van het bewijsmateriaal van 15 studenten. De richtlijn NVAO beoordeling afstudeerwerken instrueert panels een aselechte en gestratificeerde steekproef te trekken "van 15 verschillende eindwerken (of portfolio's) uit de grootste variant van de opleiding" (NVAO, 2012, p.1).

Zoals in deel 1 is betoogd tonen opleidingen dat veelal aan door het verzamelen van diverse vormen van bewijsmateriaal in meerdere afstudeeronderdelen. Om een representatief beeld te geven van de gerealiseerde eindkwalificaties is het dus zaak om van de geselecteerde studenten al dit bewijsmateriaal aan te leveren, samen met de beoordelingsmodellen die de panels dienen te hanteren bij de herbeoordeling van het materiaal. Het is voor opleidingen dus belangrijk om goed na te denken wat ze in het dossier van de student opnemen om eindkwalificaties aan te tonen.

Het is de expertgroep opgevallen dat er veel onduidelijkheid bestaat over de wijze waarop visitatiepanels het bereikte eindniveau van een opleiding toetsen. Daarom staat deze werkwijze in bijlage 4 beschreven.

DEEL 3:

VISIE OP ONDERZOEKEND VERMOGEN

Onderzoekend vermogen is ondersteunend
bij het maken van een beroepsproduct

1. Inleiding

Afgestudeerde hbo-ers dienen te beschikken over de bekwaamheid 'onderzoekend vermogen' (HBO-raad, 2009) maar er is veel onduidelijkheid over wat dit precies betekent. Wel is duidelijk dat het hbo niet opleidt tot onderzoeker (behalve opleidingen zoals die tot laboratoriumonderzoeker) maar tot beroepsbeoefenaar. Wat dan wél de rol van onderzoekend vermogen is, is voor veel opleidingen onduidelijk. Een gevolg van die onduidelijkheid is dat opleidingen in het hbo moeite hebben onderzoek te integreren in de rest van het curriculum en niet goed weten hoe onderzoekend vermogen te toetsen. Een ander gevolg is dat de eisen aan het werk van studenten op het gebied van onderzoek onrealistisch hoog worden. Het woord onderzoek roept veel associaties op met kwaliteitseisen die maar beperkt relevant zijn voor beroepsonderwijs op bachelorniveau.

De expertgroep probeert in onderstaande tekst te verhelderen wat kan worden verstaan onder onderzoekend vermogen, welke rol dit vermogen speelt bij het uitvoeren van beroepsopdrachten en hoe dit vermogen kan worden geëxamineerd. De tekst sluit af met een opsomming van minimumeisen voor onderzoekend vermogen op bachelorniveau die volgens de expertgroep realistisch zijn.

2. Waar komt de eis van onderzoekend vermogen vandaan?

We gaan hier niet in op het grote debat over de rol van onderzoek in het hbo in het algemeen en het karakter van dat onderzoek (zie voor een overzicht Griffioen, Visser-Wijnveen, & Willems, 2013). We richten ons op de eis dat bachelorstudenten onderzoekend vermogen moeten hebben. Deze eis is als eerste geformuleerd in de nota 'Kwaliteit als opdracht' (2009, p.17) van de toenmalige HBO-raad: "In onze moderne samenleving is het cruciaal dat hbo-bachelors over een onderzoekend vermogen beschikken dat leidt tot reflectie, tot evidence based practice, en tot innovatie". Deze eis komt voort uit de visie dat onze economie in toenemende mate is gebaseerd op kennis. Het innovatievermogen van de economie is sterk afhankelijk van het vermogen om nieuwe kennis te ontwikkelen. De motor voor die nieuwe kennis is niet alleen de wetenschap maar ook de kennisintensieve organisaties in het bedrijfsleven en bij de overheid waarin de hbo-opgeleide zo'n belangrijke rol speelt.

De eis van onderzoekend vermogen is overigens niet nieuw. De Dublin Descriptoren (Joint Quality Initiative, 2004) - die ook voortkomen uit de visie van Europa als de kennis-economie - wezen ook al in de richting van een dergelijk vermogen. De descriptor 'Oordeelsvorming' wordt daarin als volgt omschreven: "Is in staat om relevante gegevens te verzamelen en interpreteren (meestal op het vakgebied) met het doel een oordeel te vormen dat mede gebaseerd is op het afwegen van relevante sociaalmaatschappelijke, wetenschappelijke of ethische aspecten." Gegevensverzameling en interpretatie is een belangrijk aspect van onderzoek.

Ook de Commissie Franssen (2001) die de generieke kwalificaties voor de hbo-bachelor heeft opgesteld, wees al in de richting van onderzoekend vermogen. Zo moesten studenten volgens de commissie complexe probleemsituaties kunnen analyseren, relevante (wetenschappelijk) inzichten, theorieën concepten en onderzoeksresultaten kunnen toepassen, werkzaamheden planmatig kunnen aanpakken, oplossingen kunnen ontwikkelen en beoordelen en kunnen reflecteren op het handelen.

Op basis van de Dublin Descriptoren en de generieke kwalificaties zou men nog kunnen concluderen dat een bachelorstudent alleen onderzoek hoeft te kunnen toepassen en geen onderzoek zelfstandig hoeft te kunnen uitvoeren. Met de introductie van het begrip onderzoekend vermogen geeft de HBO-raad aan dat studenten ook in staat moeten zijn om zelf een onderzoek uit te voeren. Dit gaat verder dan het beschikken over een kritische houding ten opzichte van de beroepspraktijk en het toepassen van onderzoeksresultaten van anderen.

De HBO-raad geeft in 'Kwaliteit als opdracht' aan dat onderzoekend vermogen meer is dan een onderzoekende houding wanneer ze stelt dat onderzoekend vermogen moet leiden tot reflectie, evidence-based practice en innovatie (HBO-raad, 2009). Deze termen

worden in het document niet nader omschreven. De expertgroep interpreteert deze termen als volgt:

1. *Reflectie*: dit betreft in dit verband het terugkijken op het eigen handelen in de beroepspraktijk, signaleren wat er niet goed ging, dat proberen te verklaren vanuit de kennisbasis en in diezelfde kennisbasis ook de uitgangspunten vinden voor een betere aanpak. Reflectie is alleen mogelijk als het handelen ook goed is *verantwoord*.
2. *Evidence-based practice*: dit betreft het gebruiken van de kennisbasis om de juiste handelingen te kiezen. Het gaat dus om het goed *onderbouwen* van het handelen. De kennisbasis mag daarbij divers zijn: literatuur, kennis uit de praktijk en kennis van de patiënt, cliënt of opdrachtgever.
3. *Innovatie*: dit betreft het *vernieuwen* van de beroepspraktijk om deze te verbeteren.

Met name het derde punt laat zien dat onderzoekend vermogen ook meer is dan onderzoeksresultaten kunnen toepassen. Vernieuwing om de beroepspraktijk te verbeteren vereist het genereren van kennis naast het benutten van bestaande kennis. Hoe ver die vernieuwing dient te gaan komt in hoofdstuk 5 aan de orde.

3. Hoe kunnen we onderzoekend vermogen definiëren?

Onderzoekend vermogen bestaat uit drie componenten:

Onderzoekende houding

In de eerste plaats betekent onderzoekend vermogen het beschikken over een onderzoekende houding. Voor de invulling van dit begrip kunnen we gebruik maken van het werk van Van der Rijst (2009). Hij onderscheidt zes aspecten aan een onderzoekende houding:

- kritisch zijn;
- willen begrijpen;
- willen bereiken;
- willen delen;
- willen vernieuwen;
- en willen weten.

Kennis uit onderzoek van anderen toepassen

De tweede component is het beschikken over het vermogen om bij het maken van beroepsproducten gebruik te maken van de kennis van anderen. Dat kunnen resultaten zijn van ander onderzoek of de praktijkkennis van professionals in de beroepspraktijk.

Zelf onderzoek doen

De derde component is het beschikken over het vermogen om zelf de onderzoekscyclus te doorlopen, die loopt van het formuleren van een vraag via het verzamelen van data tot het geven van een antwoord. De data die verzameld moeten worden mogen daarbij ook uit secundaire bronnen of uit de literatuur komen. Het gaat erom dat de student het principe van de rode draad die in de onderzoekscyclus zit begrijpt en kan hanteren.

4. Wat is dan onderzoek?

Een definitie voor onderzoek mag alleen definiërende eigenschappen van onderzoek bevatten en geen kwaliteitseisen voor goed onderzoek. De meeste definities bevatten dergelijke kwaliteitseisen wel. Bij het zoeken van een goede, oordeelsneutrale, definitie van onderzoek kan de definitie van Verschuren (2009) als startpunt dienen. Verschuren stelt dat onderzoek is: “het doelbewust en methodisch zoeken naar nieuwe kennis in de vorm van antwoorden op vooraf gestelde vragen volgens een tevoren opgesteld plan”.

Deze definitie bevat nog wel een aantal kwaliteitseisen waarover niet iedereen hetzelfde denkt. Niet in alle onderzoeksparadigma's is men het eens met de eis dat onderzoek altijd doelbewust is, dat vragen van tevoren moeten worden gesteld en dat het mogelijk is

van tevoren een plan op te stellen. Daarbij komt dat het bij praktijkgericht onderzoek in het hbo niet primair gaat om het bijdragen van nieuwe kennis aan wat er in de theorie al bekend is maar om het ontwikkelen van kennis die relevant is voor de beroepspraktijk. Dit leidt tot de volgende, door de expertgroep aangepaste, definitie van onderzoek die specifiek is voor het hbo onderwijs:

Onderzoek is het methodisch beantwoorden van vragen dat leidt tot relevante kennis

5. Wat is het vereiste niveau van onderzoek in bachelor en master?

Over het verschil in vereist niveau tussen de professional bachelor- en de professional masteropleiding is elders al veel gezegd (Bulthuis, 2013). Wat betreft onderzoekend vermogen wijzen Griffioen en Wortman (2013) er terecht op dat het verschil tussen bachelor en master zit in de complexiteit van de taak & context en de mate van zelfstandigheid die wordt verwacht. Bij onderzoek is de complexiteit van de taak primair een functie van twee variabelen: 1) de mate waarin gestreefd wordt naar relevantie van de resultaten voor een bredere groep (de relevantie-claim) en 2) de mate waarin gestreefd wordt naar grondigheid van de methode (de rigor-claim) (Butter, 2013).

Rigor claim:	Bescheiden	Diepgaander	Uitgebreid
Relevantie claim:			
Relevant voor de situatie / opdrachtgever	BA		
Relevant voor het vakgebied		MA	
Relevant voor maatschappij en wetenschap			PhD

Figuur 2: Niveaus van onderzoekend vermogen

Relevantie-claim

Het doel van onderzoek kan zijn om resultaten op te leveren die alleen relevant zijn voor een specifieke situatie die is onderzocht. Dat kan bijvoorbeeld de situatie van een patiënt, cliënt of opdrachtgever zijn. De relevantie-claim is dan bescheiden. Hieraan kunnen we ook de eis van innovatie koppelen. Bij een bescheiden relevantie-claim hoeft de innovatie niet nieuw te zijn voor de hele wereld maar wel nieuw voor de specifieke situatie.

Het doel kan echter ook zijn om resultaten op te leveren die in potentie relevant zijn voor andere situaties, patiënten, cliënten of opdrachtgevers en daarmee voor de collega beroepsbeoefenaren in hele vakgebied. De relevantie-claim is dan groter.

Ook kan er sprake zijn van een nog bredere doelgroep. Dan is de claim om iets bij te dragen aan de maatschappij als geheel of een wetenschappelijk forum, zoals in het geval van een promotieonderzoek. Op deze manier zijn drie niveaus van relevantie te onderscheiden.

Rigor-claim

Er zijn ook meerdere niveaus te onderscheiden in de mate waarin het onderzoek grondig wordt uitgevoerd. Schön (1991) hanteert hiervoor de term 'rigor'.

Basisvoorwaarde is dat er een duidelijke ketting van bewijsmateriaal ('chain of evidence') ten grondslag ligt aan de uitspraken die op basis van het onderzoek worden gedaan.

Het niveau van rigor hangt vervolgens af van onder andere:

	Bescheiden	Diepgaander	Uitgebreid
Gebruik literatuur	Vakliteratuur vs Wetenschappelijke literatuur Nederlands vs Engels		
Bewijskracht (Van Yperen et al., 2013)	Theoretisch, Indicatief, Causaal		
Generalisatie	Geen, Beperkte groep, Grote groep		
Body of Knowledge Onderzoek	Vereiste kennis van onderzoeksdesigns en -methoden		

Figuur 3: Mogelijke dimensies van de rigor-claim

- welke literatuur wordt geraadpleegd (vakliteratuur, wetenschappelijke literatuur, buitenlandse literatuur);
- welke bewijskracht wordt nagestreefd (Van Yperen, Veerman, & Bijl, 2013) (alleen theoretische bewijzen, empirische indicaties verzamelen of causaal toetsen);
- hoe wordt omgegaan met het uitsluiten van rivaliserende verklaringen (Yin, 2000);
- welke ambitie er is tot generaliseren en welke eisen er gesteld worden aan de body of knowledge & skills die de student moet beheersen op het gebied van onderzoek.

Combineren we beide dimensies en plaatsen we daar bachelor, master en PhD niveau in dan ontstaat figuur 2. Dimensies voor de rigor-claim staan weergegeven in figuur 3.

Bachelor

De expertgroep stelt voor af te spreken dat voor de bachelor geldt dat de relevantieclaim minimaal 'relevant voor de situatie / opdrachtgever' is en dat de rigor-claim bescheiden is. Dit betekent in ieder geval een duidelijke 'chain of evidence'. Wat verder onder bescheiden wordt verstaan hangt af van het opleidingsprofiel en de body of knowledge & skills voor onderzoek (BoKSO) die (landelijk) binnen de betreffende opleiding wordt afgesproken.

Master

In de master geldt als minimumniveau wél dat de student laat zien dat hij kan reflecteren op de relevantie van zijn onderzoek voor andere situaties dan die zijn onderzocht. Het zal niet altijd mogelijk zijn om een concrete bijdrage te leveren aan de beroepspraktijk die in meerdere gevallen toepasbaar is, maar een reflectie daarop is wel altijd mogelijk. Tevens dient de rigor diepgaander te zijn dan bij onderzoek op bachelorniveau.

Phd

Op PhD-niveau geldt dat de promovendus een bijdrage moet leveren die relevant is voor de maatschappij als geheel en de wetenschap. Dit gaat vaak gepaard met een uitgebreide rigor-claim.

6. Wat is de relatie tussen onderzoek en maken van beroepsproducten?

Het hbo leidt meestal niet op tot onderzoeker. Het hbo verzorgt beroepsopleidingen. Studenten verwerven bekwaamheden die hen in staat stellen de beroepsproducten te realiseren die in de latere beroepsuitoefening van hen worden verlangd. Wat is dan de relatie tussen onderzoek en het maken van beroepsproducten?

Deze relatie verschilt naar type opleiding en beroep. Griffioen & Wortman (2013) maken onderscheid tussen *onderzoeksafhankelijke beroepen*, *onderzoekende beroepen*, *reflectieve beroepen* en *ontwerpende beroepen*. De expertgroep hanteert een vergelijkbaar onderscheid dat is gebaseerd op het soort beroepsproduct dat in een beroep dominant is. Zoals reeds genoemd in deel 1 zijn er volgens (Losse, 2012) vijf soorten beroepsproducten (zie tabel 4).

Tabel 4: soorten beroepsproducten

Beroepsproduct	Omschrijving	Type opleiding (Griffioen en Wortman, 2013)	Voorbeeld van opleiding waarin dit product dominant is	Processen
1. Advies	Mogelijke oplossing of plan	onderzoeksafhankelijke beroepen	Commerciële economie	Onderzoekscyclus
2. Ontwerp	Verbeelding van een eindproduct	onderzoeksafhankelijke beroepen ontwerpende beroepen	Product design	Ontwerpcyclus en onderzoekcycli
3. Eindproduct	Fysieke of digitale uitwerking van ontwerp	onderzoeksafhankelijke beroepen ontwerpende beroepen	Engineering	Ontwerpcyclus, onderzoekcycli en realisatieproces
4. Handeling	Professioneel gedrag naar cliënt/publiek	onderzoeksafhankelijke beroepen reflectieve beroepen	Verpleegkunde Kunstonderwijs Maatschappelijk werk	Ontwerpcyclus, onderzoekcycli en realisatieproces
5. Onderzoek	Antwoorden en conclusies	onderzoeksafhankelijke beroepen onderzoekende beroepen	Chemisch analist	Onderzoekscyclus

De rol van onderzoek verschilt per type beroepsproduct. Zo komen een advies (1) en onderzoek (5) volledig tot stand door slechts het doorlopen van een onderzoekscyclus, terwijl een ontwerp (2) tot stand komt door het doorlopen van een ontwerpcyclus waarin een aantal beslissingen wordt gebaseerd op onderzoek. Een eindproduct (3) komt tot stand door het doorlopen van een ontwerpcyclus, gecombineerd met onderzoek en een realisatieproces. Bij een handeling (4) ligt dit weer anders. In een handelingsproces vindt op onderdelen onderzoek plaats. Zo is in de zorg de keuze voor een handeling vaak gebaseerd op onderzoek bij de cliënt. Handelingen in de kunst komen tot stand in een constant proces van maken en reflecteren waarbij het reflecteren gezien kan worden als onderzoek in bovenstaande betekenis van het woord, namelijk als het methodisch beantwoorden van een vraag.

Wat deze vijf typen beroepsproducten gemeenschappelijk hebben, is dat ze tot stand komen door methodisch te werken. Onderzoek in het hbo ondersteunt hierbij het methodische werken. Bachelorstudenten leren in hun opleiding de methodische werkwijzen waarmee ze beroepsproducten kunnen realiseren. Deze werkwijzen zijn beroepsspecifiek. Een elektrotechnicus leert andere werkwijzen dan een verpleegkundige. De expertgroep noemt deze werkwijzen 'beroepsmethoden'.

Wat alle beroepsmethoden gemeenschappelijk hebben is dat ze bestaan uit stappen. Bij veel stappen moet er, voordat er kan worden gehandeld, eerst één of meerdere vragen worden beantwoord, bijvoorbeeld over de beginsituatie, over het te bereiken doel etc. Wanneer dit vragen betreft waar de kennisbasis of de reguliere werkprocedures geen antwoord op geven én wanneer het tevens voor de kwaliteit van het beroepsproduct belangrijk is dat het antwoord van hoge kwaliteit is, dan is er sprake van een onderzoeksvraag en moet er onderzoek worden gedaan. Oskam, Cowan, Hoiting, & Souren (2012) maken in dit verband onderscheid tussen onderzoeken en uitzoeken. Onderzoeken is nodig wanneer de vraag belangrijk is en het antwoord niet direct voor het grijpen ligt. In andere gevallen volstaat uitzoeken. Uitzoeken is geen onderzoek omdat de vraag niet methodisch (dus volgens bepaalde methoden met bijbehorende kwaliteitscriteria) wordt beantwoord. Tabel 5 geeft een aantal voorbeelden van onderzoeksvragen die aan de orde kunnen komen in verschillende beroepsmethoden. Voor ieder soort beroepsproduct uit tabel 4 is een voorbeeld gezocht. Deze voorbeelden komen uit verschillende opleidingen.

Tabel 5: beroepsproducten, methoden en vragen

Beroeps-product	Omschrijving van het beroeps-product	Opleiding waar het voorbeeld uit komt	Voorbeeld van een opdracht	Voorbeeld van beroeps-methode	Voorbeeld van een onderzoeksvraag
1. Advies	Mogelijke oplossing of plan; resultaat op papier	Vrijtijds-management	Manieren vinden om meer klanten aan te trekken voor een reisbureau	Eerste twee stappen van de regulatieve cyclus: diagnose en ontwerp	Wat zijn de oorzaken van het dalende aantal klanten?
2. Ontwerp	Verbeelding van een eindproduct; resultaat op papier	Engineering	Bouwen van een sorter-machine	Good Automated Manufacturing Practice (GAMP)	Hoe ontwerp je het systeem zodanig dat het de User acceptance test en Factory acceptance test doorstaat?
3. Eindproduct	Fysieke of digitale uitwerking van ontwerp; resultaat in praktijk gerealiseerd	Informatica	Maken van een mobile bestelsysteem voor een café	OpenUp-Inception fase	Wat zijn de belangrijkste eisen (use-cases) waaraan het ontwerp moet voldoen?
4. Handeling	Professioneel gedrag naar cliënt/publiek; resultaat in praktijk gerealiseerd	Logopedie	Behandeling kiezen bij een patiënt met Multiple Sclerose	Behandeling van slikstoornissen	Is behandeling met behulp van neuromusculaire elektrostimulatie (NMES) bij een patiënt met Multiple Sclerose (MS) effectief in het verminderen van aspiratie?
5. Onderzoek	Antwoorden en conclusies; resultaat op papier	Chemisch analist	De bepaling van vrij cortisol in serum met behulp van evenwichtsdialyse	Empirische cyclus	Wat zijn de invloed van de dialyseverhouding serum/buffer, de temperatuur tijdens de dialyse, en de dialysetijd op de validiteit?

Bij de meeste beroepsmethoden kunnen in elk van de stappen één of meerdere vragen opkomen die niet uitgezocht maar onderzocht moeten worden. De student zal dus veelal meerdere keren de onderzoekscyclus moeten doorlopen. Tabel 6 geeft hiervan een voorbeeld voor de ontwerpcyclus als beroepsmethode.

Tabel 6: onderzoeksvragen in de ontwerpcyclus

Fase in de ontwerpcyclus	Onderzoeksvragen	Onderzoeksfunctie (Oost & Markenhof, 2002)
1. Analyseren en beschrijven	Wat is het probleem? Wat zijn de oorzaken van het probleem?	Beschrijven, evalueren Verklaren
2. Programma van eisen opstellen	Wat zijn de eisen van de gebruikers?	Beschrijven
3. (deel)uitwerkingen bedenken	Welke oplossingen zijn bekend? Wat zijn hun voor- en nadelen?	Beschrijven Evalueren
4. Ontwerpvoorstel formuleren	Welk oplossing voldoet aan de eisen?	Evalueren
5. Ontwerp realiseren	Nvt	
6. Ontwerp testen en evalueren	Hoe goed is de oplossing? Wat zijn randvoorwaarden voor het gebruik van de oplossing?	Evalueren Verklaren

7. Wat betekent dit voor onderwijs van onderzoekend vermogen?

Deze nauwe koppeling tussen het realiseren van beroepsproducten en het doen van onderzoek in de bachelor heeft gevolgen voor de manier waarop onderwijs in onderzoekend vermogen in het curriculum vorm moet krijgen. Het startpunt voor onderwijs in onderzoekend vermogen is het uitvoeren van kerntaken c.q. het maken van beroepsproducten en het toepassen van beroepsmethoden daarbij. Bij het toepassen van beroepsmethoden

komen vragen naar boven. Vervolgens kan in het onderwijs worden stilgestaan bij het methodisch beantwoorden van die vragen door het doorlopen van de onderzoekscyclus. Onderzoeksonderwijs dient in deze visie zoveel mogelijk geïntegreerd te zijn in het reguliere onderwijs.

Wanneer er toch voor wordt gekozen in aparte modules les te geven over de onderzoekscyclus, dan moet bij voorkeur iedere keer de relatie worden gelegd met het realiseren van beroepsproducten (bijvoorbeeld door middel van het geven van voorbeelden). Op die manier wordt voorkomen dat het aanleren van onderzoekend vermogen losstaat van de opleiding tot het beroep (Van der Linden, 2012).

8. Wat betekent dit voor het toetsen van onderzoekend vermogen in het afstuderen?

Onderzoekend vermogen *kan* getoetst worden in aparte onderzoeksopdrachten, los van het maken van een beroepsproduct. Echter, veelal zal onderzoekend vermogen samen met andere eindkwalificaties worden getoetst in authentieke beroepsopdrachten waarin studenten een beroepsproduct maken en waarin ze als onderdeel daarvan onderzoek doen.

De kwaliteitseisen die aan het onderzoek binnen de authentieke beroepsopdracht worden gesteld zullen verschillen per beroepsproduct. Hierbij kan een onderscheid worden gemaakt tussen enerzijds de beroepsproducten advies (1), ontwerp (2) en onderzoek (5) en anderzijds eindproduct (3) en handeling (4).

- Bij de eerste drie beroepsproducten blijft het resultaat papier. Het advies of ontwerp wordt niet gerealiseerd. De kwaliteit van het resultaat is daarmee alleen te verantwoorden en beoordelen door te kijken naar de manier waarop het tot stand is gekomen en de onderbouwing: is het beroepsproduct gebaseerd op betrouwbare en valide informatie?
- Het beroepsproduct eindproduct (3) en handeling (4) bestaan in tegenstelling tot de vorige drie niet alleen op papier. Zij zijn in de praktijk gerealiseerd en hun kwaliteit is óók af te leiden uit hun functioneren: doet het apparaat wat het moet doen? Leidt de handeling tot het beoogde resultaat? Bij onderzoek, advies en ontwerp zullen opleidingen dan ook andere kwaliteitseisen stellen aan de rigor van het onderzoek dan bij een eindproduct en een handeling.

9. Welke minimeisen aan onderzoekend vermogen zijn uit het voorgaande af te leiden?

Uit het voorgaande is een aantal eisen af te leiden voor het afstudeerniveau. Deze eisen liggen hoger dan vóór het verschijnen van 'Kwaliteit als opdracht'.

Bachelor

De eis van onderzoekend vermogen betekent volgens de expertgroep op bachelorniveau:

1. Het *verantwoorden* van de gekozen aanpak en de resultaten.
2. Het *reflecteren* op de gekozen aanpak en de resultaten.
3. Het *toepassen* van modellen, theorieën en onderzoeksresultaten van anderen.
4. Een *kritische houding* ten aanzien van modellen, theorieën en onderzoeksresultaten van anderen.
5. Het realiseren van *vernieuwing* in één specifieke situatie.
6. Zelf kunnen *uitvoeren* van een onderzoekscyclus met een rode draad van vraag naar conclusie.

Daarnaast vult iedere opleiding dit aan met het beschrijven van de body of knowledge & skills op het gebied van onderzoek.

Master

Voor master niveau gelden bovenstaande eisen ook, behalve dat punt 5 wordt uitgebreid tot het realiseren van *vernieuwing* in één specifieke situatie én het reflecteren op de relevantie van het onderzoek voor andere situaties dan die zijn onderzocht. Tevens zal de body of knowledge & skills op het gebied van onderzoek bij de master uitgebreider zijn.

Associate degree

Een student met een Associate degree (Ad) hoeft volgens de expertgroep niet te beschikken over onderzoekend vermogen zoals hierboven beschreven. De student hoeft niet de gehele cyclus zelfstandig te kunnen doorlopen. Wel wordt een onderzoekende houding verwacht en moet de student in staat zijn binnen een onderzoek dat iemand anders uitvoert gegevens te verzamelen. Ook moet de student standaardmodellen en theorieën kunnen toepassen en de keuze kunnen verantwoorden. Daarmee zijn alleen eisen 1 t/m 4 op de Ad van toepassing.

BIJLAGEN

Bijlage 1: Lijst van gesprekspartners

Henri Ponds	NVAO
Fred de Bruijn	Hobéon
Ruud van der Herberg	Hobéon
Paul Thijsen	NQA
Adriaan Horrevorts	NQA
Cees van de Vleuten	Universiteit Maastricht
René Kloosterman	Aequi
Tineke Kleene	Zelfstandig NVAO gecertificeerd secretaris

Bijlage 2: Verklarende woordenlijst

Term	Betekenis
Afstudeeronderdeel	Onderdeel van het afstudeerprogramma waarin één specifieke beroepsopdracht aan de orde is.
Afstudeerprestatie	Het resultaat van de uitvoering van een beroepsopdracht. De prestatie is een beroepsproduct, al dan niet voorzien van een verantwoordingsverslag.
Afstudeerprogramma	Onderdelen in het toetsprogramma van een opleiding waarin de student laat zien te beschikken over de eindkwalificaties van de opleiding.
Afstuderen	Het doorlopen van het afstudeerprogramma.
Beoordelingsdimensie	Beschrijft een van de aspecten waarop het handelen van de kandidaat en/of resulterende producten daarvan moeten worden beoordeeld en waaraan dit aspect moet voldoen.
Beoordelingsformulier	Formulier waarop het beoordelingsmodel staat weergegeven en dat door examinatoren wordt gebruikt bij het beoordelen van afstudeerprestaties.
Beoordelingsmodel	Een model dat de kwaliteit specificeert waarop bij een afstudeerprestatie product en proces beoordeeld worden, middels prestatiecriteria en beoordelingsschalen en beslisregels.
Beoordelingsprocedure	Werkwijze die de opleiding hanteert bij het beoordelen van afstudeerprestaties.
Beoordelingsschaal	Nominale of ordinale ordening van prestatieniveaus die gebruikt wordt om uit te drukken in welke mate de afstudeerprestatie van een student voldoet aan de prestatiecriteria.
Beroepsopdracht	Een opdracht die de student ertoe aan zet om zijn bekwaamheid met betrekking tot een kerntaak van het beroep te demonstreren door het uitvoeren van meer of minder authentieke taken onder meer of minder authentieke omstandigheden, leidend tot een beroepsproduct en soms ook een verantwoordingsverslag.
Beroepsproduct	Een prestatie naar aanleiding van een beroepsopdracht die grote gelijkenis vertoont met prestaties die bij de latere uitoefening van het beroep ook geleverd zal moeten kunnen worden.
Bewijsmateriaal	De gedocumenteerde afstudeerprestaties van studenten en hun beoordeling.
Body of Knowledge & Skills	Een beschrijving van de beroepskennis en –vaardigheden waarover de student als beginnende beroepsbeoefenaar moet beschikken.
Body of Knowledge & Skills Onderzoek (BoKSO)	Een beschrijving van de beroepskennis en –vaardigheden op het gebied van onderzoekend vermogen waarover de student als beginnende beroepsbeoefenaar moet beschikken.
Competentie	Het vermogen van iemand om op adequate wijze uitvoering te geven aan de kerntaken van een specifiek beroep of specifieke functie.
Eindkwalificaties	De beschrijving van wat een student op het niveau van de startbekwame beroepsbeoefenaar moet beheersen.
Eindniveau	Indicatie van wat de studenten van de opleiding aan het einde van de studie beheersen.
Eindwerkstuk	Beroepsproduct met verantwoordingsverslag dat aan het einde van het afstudeerprogramma van een opleiding wordt vervaardigd en zicht moet geven op de mate van beheersing van een deel van de eindkwalificaties.
Evidence-based practice	Het gebruiken van de kennisbasis om de juiste handelingen te kiezen.
Examinator	Iedere docent / beoordelaar die eindverantwoordelijk is voor één of meer fasen van de toetscyclus.
Herleidbaar	De mate waarin de beoordeling door derden te controleren en te reproduceren is.
Kalibreersessie	Bijeenkomst van examinatoren waarin gewerkt wordt aan harmoniseren van de interpretatie van het beoordelingsmodel.
Kernwerkstuk	Beroepsproduct met verantwoordingsverslag dat ergens in het afstudeerprogramma van een opleiding wordt vervaardigd en zicht moet geven op de mate van beheersing van een deel van de eindkwalificaties.
Observeerbare criteria	Prestatiecriteria die kenmerken beschrijven van de te leveren afstudeerprestatie. Bij een advies en een verantwoordingsverslag gaan de observeerbare criteria over eigenschappen van het document. Bij een ontwerp gaan ze over eigenschappen van het ontwerp. Bij een product gaan ze over eigenschappen van het product. Bij een handeling zoals een behandeling of een gegeven presentatie gaan ze over eigenschappen van de handeling. Bij een verantwoordingsverslag gaan ze over eigenschappen van het verslag.
Ondersteunende hulpmiddelen	De beoordelingsschalen waarmee kwaliteitsverschillen tussen studenten uitgedrukt worden.
Onderzoek	Het methodisch beantwoorden van vragen dat leidt tot relevante kennis.
Onderzoekend vermogen	De combinatie van een onderzoekende houding, het kunnen toepassen van onderzoeksresultaten van anderen en het kunnen doorlopen van de onderzoekscyclus.
Portfolio	Het geheel aan afstudeerprestaties.
Prestatie	De uitvoering van een beroepsopdracht. De prestatie is een beroepsproduct, al dan niet voorzien van een verantwoordingsverslag.
Prestatiecriteria	Algemene aanduiding voor gewenste kwaliteiten waaraan het professionele gedrag en vervaardigde beroepsproducten en verantwoordingsverslagen van studenten dienen te voldoen. Ook wel beoordelingscriteria genoemd.
Prestatieniveau	Kwaliteitsniveau waarop de prestatie wordt geleverd.
Protocol	Een inventarisatie van punten waar een te beoordelen object of proces aan moet voldoen om een minimale kwaliteit te garanderen.
Relevantie-claim	De mate waarin gestreefd wordt naar relevantie van de resultaten van onderzoek voor een bredere groep.
Rigor-claim	De mate waarin bij onderzoek gestreefd wordt naar grondigheid van de methode.
Verantwoordingsverslag	Verslag van de student waarin wordt verantwoord hoe het beroepsproduct tot stand is gekomen en waarin de tijdens het proces genomen beslissingen worden onderbouwd.
Werkstuk	Zie Kernwerkstuk.

Bijlage 3: Overzicht beoordelaarseffecten

Uit Van Berkel & Bax (2006).

In algemene zin zijn verschillen tussen beoordelaars die hetzelfde werk beoordelen niet wenselijk. De verschillen zorgen voor ongewenste effecten op de score van de student. De persoon van de beoordelaar mag immers, zo is de algemene opvatting, bij het geven van scores geen rol spelen. Toch is in de meeste gevallen niet uit te sluiten dat de persoon van de beoordelaar het uiteindelijk oordeel ongewenst deels heeft beïnvloed. Dat zal ertoe leiden dat de beoordeling minder betrouwbaar wordt. Er is veel onderzoek verricht naar deze ongewenste effecten. In zijn standaardwerk, *Methodologie* (1961), benoemt en definieert De Groot deze effecten. Het onderstaande is daaraan deels ontleend.

Contaminatie-effect

Het contaminatie-effect is het effect dat optreedt bij beoordelen wanneer de vrijheid in de beoordeling, onwillekeurig of willekeurig, wordt gebruikt voor andere doeleinden dan die van een onbevangen, onbevooroordeelde beoordeling. Dit is bijvoorbeeld het geval als beoordelaars lagere scores geven om aan te tonen dat hun vak moeilijk is.

Halo-effect

Het halo-effect treedt op wanneer beoordelaars hun oordeel mede laten beïnvloeden door andere prestaties van de student dan geuit op de te beoordelen prestatie. De beoordelaars hebben in dat geval de neiging om een goede student bij een wat mindere prestatie wat hoger te beoordelen dan de prestatie van de student rechtvaardigt. De prestatie wordt overgewaardeerd.

Horn-effect

Het horn-effect is het tegenovergestelde van het halo-effect. Het horn-effect treedt op wanneer beoordelaars hun oordeel mede laten beïnvloeden door andere prestaties van de student dan geuit op de te beoordelen prestatie. De beoordelaars hebben in dat geval de neiging om een slechte student bij een wat hogere prestatie wat lager te beoordelen dan de prestatie van de student rechtvaardigt. De prestatie wordt ondergewaardeerd.

Normverschuiving

De normverschuiving is het effect dat optreedt wanneer een beoordelaar zich aanpast aan de geleverde prestaties van studenten. Beoordelaars kunnen bijvoorbeeld minder streng worden als na een aantal beoordelingen blijkt dat de meeste studenten eenzelfde vraag fout beantwoorden.

Restriction-of-range

Restriction-of-range wordt veroorzaakt door de vrijheid in de beoordelingstaak die willekeurig of onwillekeurig leidt tot bepaalde verdelingen van de beoordelingen die algemeen menselijke of persoonlijke neigingen uitdrukken. Zo zal de ene beoordelaar wel alle schaalwaarden (scores van 1 tot 10) gebruiken, terwijl een andere beoordelaar steeds waarden zal toekennen in de buurt van het midden.

Sequentie-effect

Het sequentie-effect treedt op wanneer een beoordelaar de beoordeling ten onrechte mede laat leiden door een of meer voorafgaande beoordelingen. Bijvoorbeeld, beoordelaars kijken toetsvragen in een bepaalde volgorde na. Een groot aantal slechte antwoorden dat wordt gevolgd door een correct antwoord, kan er dan toe leiden dat het correcte antwoord onevenredig hoog wordt beoordeeld.

Signifisch effect

Het significisch effect treedt op wanneer beoordelaars letten op verschillende aspecten, of als ze de aspecten verschillend wegen in hun oordeel. Dit effect treedt bijvoorbeeld op wanneer beoordelaars schrijfproducten beoordelen. De ene beoordelaar let vooral op de grammaticale (on)juistheden en de andere op structuur of inhoud.

Bijlage 4: Toetsing en beoordeling van eindniveau bij een visitatie

De expertgroep heeft in de gevoerde gesprekken gemerkt dat over de wijze waarop visitatiepanels van evaluatiebureaus het bereikte eindniveau van een opleiding toetsen en beoordelen nogal wat onduidelijkheid bestaat. Om de door de bureaus gevolgde werkwijze te doorgronden is het goed hun positie in het accreditatieproces helder te hebben.

Positie van de evaluatiebureaus

Nederland kent in totaal acht evaluatiebureaus. Evaluatiebureaus zijn zelfstandige bureaus die op verzoek van opleidingen in het hoger onderwijs de visitatie uitvoeren door middel van externe en onafhankelijke panels van deskundigen. Opleidingen zijn overigens niet verplicht een evaluatiebureau in te schakelen; ook gecertificeerde secretarissen⁴ en andere bureaus kunnen door een opleiding worden ingehuurd om de visitatie te organiseren. Daarbij zijn dezelfde regels van toepassing als bij het inhuren van een evaluatiebureau. Voorwaarde is dat er altijd een gecertificeerd secretaris bij de visitatie betrokken is en dat de samenstelling van het panel voldoet aan de voorwaarden van de NVAO. De samenstelling van ieder panel wordt voorafgaand aan de visitatie door de NVAO goedgekeurd. Het panel brengt door middel van een visitatierapport advies uit aan de NVAO met betrekking tot de accreditatie van de opleiding. De opleiding kan met een positief visitatierapport accreditatie aanvragen bij de NVAO⁵.

Evaluatiebureaus hebben daarmee geen formele relatie met de NVAO, maar worden, omwille van vergelijkbaarheid van de beoordelingen, wel door de NVAO geïnstrueerd. Een sleutelpositie in deze constructie wordt ingenomen door de gecertificeerde secretaris. Secretarissen van panels zijn door de NVAO getraind. De secretaris is geen lid van het panel, maar houdt het proces in de gaten. De panelleden gaan over de inhoud en de oordelen.

Samenstelling van de panels⁶

Voor het accrediteren van bestaande opleidingen stelt de instelling, vaak in samenspraak met het evaluatiebureau, het panel samen, wijst een secretaris aan en draagt het panel vervolgens ter instemming voor aan de NVAO. De instelling levert hiervoor gegevens aan over de deskundigheden en onafhankelijkheid van de panelleden en de secretaris. Aan de samenstelling van het beoordelingspanel worden door de NVAO de volgende eisen gesteld:

1. het panel bestaat uit ten minste vier leden, onder wie ten minste twee gezaghebbende domeindeskundigen en een student;
2. het panel beschikt in zijn geheel over de volgende deskundigheden:
 - a. deskundigheid ten aanzien van de ontwikkelingen in het vakgebied;
 - b. internationale deskundigheid;
 - c. werkvelddeskundigheid in het voor de opleiding relevante beroepenveld (indien van toepassing);
 - d. ervaring met het geven en ontwikkelen van onderwijs op het desbetreffende opleidingsniveau en deskundigheid ten aanzien van de door de opleiding gehanteerde onderwijsvorm(en);
 - e. studentgebonden deskundigheid;
 - f. visitatie- of auditdeskundigheid.
3. het panel is onafhankelijk (de leden hebben ten minste vijf jaar geen banden gehad met de instelling waartoe de opleiding behoort);

⁴ http://www.nvao.net/gecertificeerde_secretarissen

⁵ <http://www.nvao.net/evaluatiebureaus>

⁶ <http://www.nvao.net/panels>

4. het panel wordt bijgestaan door een onafhankelijke, externe secretaris die is getraind en gecertificeerd door de NVAO.

Standaarden

De panels kijken naar 3 of 16 standaarden, afhankelijk van het soort accreditatie dat plaatsvindt. Specifieke aandacht gaat uit naar de consistentie tussen standaard 1: “De beoogde eindkwalificaties van de opleiding zijn wat betreft inhoud, niveau en oriëntatie geconcretiseerd en voldoen aan internationale eisen” en standaard 3 (of in het geval van een uitgebreide opleidingsbeoordeling standaard 16): “De opleiding beschikt over een adequaat systeem van toetsing en toont aan dat de beoogde eindkwalificaties worden gerealiseerd”. Om standaard 3 of 16 te beoordelen kijkt het panel enerzijds naar de manier waarop de opleiding toetst en waarop zij de kwaliteit van het examenprogramma borgt. Dit dient valide, betrouwbaar en voor de student inzichtelijk te zijn. Anderzijds vormt het panel zich een oordeel over het eindniveau van de studenten. Voor het beoordelen van het eindniveau kan het panel kijken naar de resultaten van alle afstudeeronderdelen, waaronder eventuele kernwerkstukken. In de Wet versterking kwaliteitsborging hoger onderwijs zijn deze onderdelen inmiddels gescheiden.

Te hanteren beoordelingsmodel

De instructie van de NVAO aan evaluatiebureaus bij de beoordeling van het eindniveau is dat het beoordelingsmodel van de opleiding zelf het uitgangspunt is voor de beoordeling van het eindniveau door de panels. De panels beoordelen in eerste instantie de prestatiecriteria die de opleiding zelf hanteert. Zij nemen er kennis van en stellen vast of deze passen binnen de door NVAO gehanteerde kwaliteitskaders, Dublin Descriptoren, hbo-standaard en beroepsprofiel van de opleiding en beoordelen of ze valide, betrouwbaar en voor de student inzichtelijk zijn.

Indien de prestatiecriteria niet verantwoord zijn, of niet aanwezig zijn, bespreekt het visitatiepanel dat met de opleiding. Het panel gebruikt dan (en alleen dan) in aanvulling op of in plaats van de prestatiecriteria van de opleiding eigen prestatiecriteria. Pas als het beoordelingsmodel is beoordeeld, kan het bewijsmateriaal beoordeeld worden. De NVAO spreekt in dit verband van “...de verzameling van (eind)werken die als geheel het bewijs vormen dat de student de eindkwalificaties heeft gerealiseerd” (NVAO, 2012).

De beoordeling van het bewijsmateriaal door de opleiding dient voor de student inzichtelijk te zijn. Daartoe wordt o.a. naar het volgende gekeken:

- Is het proces waarlangs de beoordeling tot stand is gekomen inzichtelijk gemaakt voor derden?
- Is duidelijk op welke criteria de beoordeling heeft plaatsgevonden?
- Zijn er meerdere beoordelaars betrokken geweest bij de beoordeling en wat is de kwaliteit van die beoordelaars?
- Is er een duidelijke afstemming tussen beoordelaars geweest en is die afstemming vastgelegd?
- Is duidelijk hoe het cijfer / de beoordeling tot stand is gekomen en is dat resultaat schriftelijk onderbouwd met heldere feedback?
- Getuigt het bewijsmateriaal van voldoende onderzoekend vermogen van de student?

Beoordelingsschalen bij de opleidingsbeoordeling⁷

De NVAO hanteert een viertal oordelen: onvoldoende, voldoende, goed en excellent. In deze paragraaf wordt bij iedere schaal allereerst een typering gegeven. Bij ieder oordeel wordt ter illustratie een aantal voorbeelden gegeven dat dienstig kan zijn bij de beoordeling van de opleiding. Bij voorbeelden die enkel gelden voor de uitgebreide opleidingsbeoordeling staat 'UOB' vermeld.

Oordeel 1: onvoldoende

De opleiding voldoet niet aan de gangbare basiskwaliteit en vertoont op meerdere vlakken ernstige tekortkomingen.

Er kan hierbij gedacht worden aan het volgende:

- De eindkwalificaties passen qua niveau en/of oriëntatie niet binnen de (inter)nationale kwalificatieraamwerken en zijn niet geconcretiseerd in domein- of opleidingsspecifieke leerresultaten.
- Programma, personeel en voorzieningen vormen samen geen faciliterende leeromgeving.
- Er is geen opleidingsbreed, transparant en coherent toetsbeleid.
- De beoogde eindkwalificaties worden niet gerealiseerd.
- De kwaliteitszorg in de opleiding is niet systematisch en vertaalt zich in een gebrek aan verbeterbeleid. (UOB)

Oordeel 2: Voldoende

De opleiding voldoet aan de gangbare basiskwaliteit en vertoont over de volle breedte een acceptabel niveau.

Er kan hierbij gedacht worden aan:

- De eindkwalificaties passen qua niveau en/of oriëntatie binnen de (inter)nationale kwalificatieraamwerken en zijn geconcretiseerd in domein- of opleidingsspecifieke leerresultaten.
- Programma, personeel en voorzieningen vormen samen een faciliterende leeromgeving die toelaat de eindkwalificaties te realiseren.
- Er is een opleidingsbreed, transparant en coherent toetsbeleid, dat evenwel nog niet consequent door iedereen gevolgd wordt.
- De beoogde eindkwalificaties worden gerealiseerd.
- De kwaliteitszorg in de opleiding is systematisch en vertaalt zich in een consequent verbeterbeleid. (UOB)

Oordeel 3: Goed

De opleiding steekt systematisch en over de volle breedte uit boven de gangbare basiskwaliteit.

Er kan hierbij gedacht worden aan:

- De beoogde eindkwalificaties passen qua niveau en/of oriëntatie binnen de (inter)nationale kwalificatieraamwerken en zijn geconcretiseerd in domein- of opleidingsspecifieke leerresultaten. Vanuit een geëxpliciteerde visie geeft de opleiding er een specifieke invulling aan.
- Programma, personeel en voorzieningen vormen samen een uitdagende leeromgeving.
- Er is een opleidingsbreed, transparant en coherent toetsbeleid dat door iedereen gevolgd wordt.

⁷ NVAO | Beoordelingskaders opleidingsbeoordeling (beperkt) | 22 november 2011, pp. 16-18

- De gerealiseerde eindkwalificaties vertalen zich in producten die systematisch bovengemiddeld zijn.
- De kwaliteitszorg in de opleiding is systematisch en vertaalt zich in een consequent verbeterbeleid dat zicht uit in een groeiende kwaliteitscultuur. (UOB)

Oordeel 4: Excellent

De opleiding steekt systematisch en over de volle breedte ver uit boven de gangbare basiskwaliteit en geldt als een (inter)nationaal voorbeeld.

Er kan hierbij gedacht worden aan:

- De eindkwalificaties passen qua niveau en/of oriëntatie binnen de (inter)nationale kwalificatieraamwerken en zijn geconcretiseerd in domein- of opleidingsspecifieke leerresultaten. Vanuit een geëxpliciteerde en unieke visie geeft de opleiding er een specifieke invulling aan en fungeert zij (inter)nationaal als een voorbeeld.
- Programma, personeel en voorzieningen vormen samen een innovatieve en originele leeromgeving.
- De gerealiseerde eindkwalificaties zijn van uitmuntende kwaliteit en vertalen zich bijvoorbeeld in awards en (inter)nationale publicaties.
- De kwaliteitszorg in de opleiding is systematisch en vertaalt zich in een consequent verbeterbeleid en sterk zelfevaluerend vermogen van de opleiding, wat zich uit in een sterke kwaliteitscultuur. (UOB)

De NVAO geeft hierbij de volgende beslisregels mee:

Beperkte opleidingsbeoordeling

- Het eindoordeel over een opleiding is in elk geval 'onvoldoende' indien standaard 3 als 'onvoldoende' beoordeeld wordt.
- Het eindoordeel over een opleiding kan alleen 'voldoende' zijn indien tenminste twee standaarden als 'voldoende' worden beoordeeld; waaronder in elk geval standaard 3.
- Het eindoordeel over een opleiding kan alleen 'goed' zijn indien tenminste twee standaarden als 'goed' worden beoordeeld; waaronder in elk geval standaard 3.
- Het eindoordeel over een opleiding kan alleen 'excellent' zijn indien tenminste twee standaarden als 'excellent' worden beoordeeld; waaronder in elk geval standaard 3.

Uitgebreide opleidingsbeoordeling

- Het eindoordeel over een opleiding is in elk geval 'onvoldoende' indien standaard 16 als 'onvoldoende' beoordeeld wordt.
- Het eindoordeel over een opleiding kan alleen 'voldoende' zijn indien tenminste de standaarden 1,3,6,9,13,14,15 en 16 als 'voldoende' worden beoordeeld.
- Het eindoordeel over een opleiding kan alleen 'goed' zijn indien tenminste de standaarden 1,3,6,9,13,14,15 en 16 als 'goed' worden beoordeeld.
- Het eindoordeel over een opleiding kan alleen 'excellent' zijn indien de standaarden 1,3,6,9,13,14,15 en 16 als 'excellent' worden beoordeeld.

Bijlage 5: Handreiking Ontwikkelen beoordelingsmodellen

In deze handreiking worden de belangrijkste stappen besproken die gezet moeten worden om te komen tot goede beoordelingsmodellen.

1. Inventariseren van taaksituaties

Een beoordelingsmodel bevat prestatiecriteria waarmee bekwaamheid met betrekking tot het uitvoeren van de kerntaak van een beroep of functie beoordeeld wordt. Het spreekt voor zich dat deze prestatiecriteria inhoudelijk goed moeten zijn afgestemd op de producten en processen die het resultaat zijn van de uitvoering van de beroepsopdracht(en). Om dat te bewerkstelligen is de eerste stap van de ontwikkelprocedure het beschrijven van de te beoordelen bekwaamheid aan de hand van een opsomming van taaksituaties waarmee de toekomstige beroepsbeoefenaar geconfronteerd kan worden. Een taaksituatie is de combinatie van een taak waarmee een beroepsbeoefenaar geconfronteerd kan worden en de situatie waarin die taak moeten worden uitgevoerd. De geïnventariseerde taaksituaties zijn een inspiratiebron voor zowel de ontwikkelaars van beroepsopdrachten als voor de ontwikkelaars van de bijbehorende beoordelingsmodellen.

Ter illustratie: De belangrijkste te verwerven bekwaamheid van een ambulanceverpleegkundige is dat hij/zij spoedeisende hulp kan verlenen. Daarvoor moet een groot aantal diagnostische en therapeutische handelingen worden beheerst, zoals: ademweg vrijmaken en vrijhouden, bevalling begeleiden, kunstmatige beademing, stelpen van bloedingen, wervel-immobilisatie, wondverzorging, medicatie bepalen en toedienen, enz. Deze taken moeten onder zeer uiteenlopende omstandigheden kunnen worden uitgevoerd, zoals: onder tijdsdruk, op een gevaarlijke locatie, in slechte weersomstandigheden, in aanwezigheid van agressieve omstanders, enz.

2. Identificeren van beoordelingsdimensies

Beoordelingsdimensies geven globaal aan waarop het handelen van de kandidaat en/of resulterende producten daarvan moeten worden beoordeeld. 'Globaal' vanwege het feit dat deze beoordelingsdimensies bruikbaar moeten zijn voor elke prestatie waarmee de bekwaamheid ten aanzien van de kerntaak kan worden aangetoond. Beoordelingsdimensies worden niet rechtstreeks gebruikt om de prestaties van kandidaten mee te beoordelen; ze dienen in de eerste plaats als richtsnoer voor de formulering van prestatiecriteria.

Ter illustratie: Eén van de kerntaken van een fysiotherapeut is therapeutisch handelen. Dat wil zeggen dat de fysiotherapeut in coöperatie met de cliënt op methodische wijze het behandelplan uitvoert. Er zijn talloze taaksituaties denkbaar waarin bekwaamheid ten aanzien van deze kerntaak zich kan manifesteren. Immers er zijn verschillende klachten waarmee een cliënt zich kan melden bij een fysiotherapeut, zoals pijnklachten, zwellingen, krachtsvermindering, ademhalingsproblematiek, enz. De cliënt zelf is eveneens een factor die van invloed is op het therapeutisch handelen. Een oudere cliënt dient anders behandeld te worden dan een volwassen cliënt of een cliënt met een verstandelijke beperking. Ook zijn er verschillende behandelmethoden die beheerst dienen te worden, zoals: oefentherapie (bijv. krachttraining, groepstraining), massagetherapie, ademhalingstherapie en dergelijke. Ten slotte kan de context waarin behandeld wordt verschil maken: gezondheidscentrum, praktijk voor sportfysiotherapie, ziekenhuis, verpleeghuis, etc.

De opstellers van het beoordelingsmodel dienen deze verschillende taaksituaties te beschouwen waarna ieder voor zich beschrijft waarin de bekwame en onbekwame beroepsbeoefenaar zich naar verwachting het meest nadrukkelijk van elkaar onderscheiden. De resultaten van de groepsleden worden vergeleken en de verschillen bediscussieerd. De aspecten waarover men overeenstemming weet te bereiken worden bestempeld als de beoordelingsdimensies van de betreffende bekwaamheid.

Voorbeelden zijn voor fysiotherapie zijn:

- Kiest behandelingen die goed zijn afgestemd op de behandeldoelen, zowel op het niveau van stoornissen, activiteiten als participatie;
- Voert behandelingen vloeiend uit en in een tempo dat past bij de beroepsuitoefening;
- Kan gemaakte keuzes voor behandelingen verantwoorden;
- Evalueert de effectiviteit en efficiëntie van het therapeutisch proces;
- Stelt het behandelplan zo nodig bij;
- Informeert de cliënt over (de effecten van) de behandeling.

Er is nog een andere manier om beoordelingsdimensies te identificeren. Die is echter alleen toe te passen als er al beoordeelde prestaties van studenten zijn. Beknopt weergegeven houdt deze methode het volgende in:

- Selecteer minimaal tien afstudeerprestaties van studenten. Zorg ervoor dat ze verschillende prestatieniveaus representeren.
- Laat elke prestatie door minimaal drie onafhankelijk van elkaar werkende examinatoren toewijzen aan één van de volgende categorieën: hoge, middelmatige en lage kwaliteit.
- Selecteer die prestaties (minimaal zes) waarover de examinatoren de grootste overeenstemming wisten te bereiken.
- De argumenten voor indeling worden uitgewisseld en bediscussieerd. De argumenten waarover de grootste overeenstemming is, worden bestempeld als de beoordelingsdimensies.

Het identificeren van de beoordelingsdimensies is, evenals het inventariseren van de taaksituaties, een belangrijke stap in de operationalisatie van de opleidingsdoelen. Reden waarom dit geen taak behoort te zijn van individuele docenten maar een gezamenlijke taak, gecoördineerd door een vertegenwoordiging van de opleidingsstaf, zoals een curriculum- of toetscommissie. Bij voorkeur wordt dit landelijk georganiseerd met een groep bestaande uit vertegenwoordigers van soortgelijke opleidingen en van de beroepsgroep.

3. Formuleren van prestatiecriteria

Beoordelingsdimensies geven globaal de verlangde kwaliteit van handelen en/of vervaardigde producten aan, maar niet waaruit die kwaliteit precies moet blijken. Daarvoor zijn prestatiecriteria nodig. Prestatiecriteria zijn nadere uitwerkingen van een beoordelingsdimensie en beschrijven waaraan het te beoordelen handelen of de kenmerken van een te beoordelen product moeten voldoen.

De accuraatheid van de beoordeling is erbij gebaat als de prestatiecriteria inhoudelijk zijn afgestemd op de door de kandidaten uit te voeren beroepsopdracht. Het ligt daarom voor de hand om prestatiecriteria te laten formuleren door de constructeurs van de beroepsopdracht(en). Deze moeten zich daarbij echter ook laten leiden door de reeds vastgestelde beoordelingsdimensies. Anders gezegd: de prestatiecriteria zijn taakspecifieke uitwerkingen van de beoordelingsdimensies. Dit betekent dat de prestatiecriteria als uitwerking van, bijvoorbeeld, de beoordelingsdimensie 'Evalueert de effectiviteit en efficiëntie van het therapeutisch proces' voor beroepsopdracht A (taaksituatie waarbij een bejaarde vrouw met ademhalingsklachten is opgenomen in een verzorgingshuis) anders zijn dan voor beroepsopdracht B (taaksituatie waarbij een amateurvoetballer met een hamstringblessure gemasseerd moet worden).

Om de objectiviteit van de beoordeling te bevorderen dienen prestatiecriteria zoveel mogelijk geformuleerd te worden in observeerbare criteria. 'Zoveel mogelijk' omdat het streven naar volledig objectieve formuleringen ontraden moet worden, aangezien men zich daarmee vaak beperkt tot het beoordelen van kwaliteiten die gemakkelijk waarneembaar zijn en dat zijn niet altijd de aspecten die het verschil maken tussen bekwaam en niet bekwaam handelen.

Ter illustratie: Neem de situatie dat de zangprestatie van een conservatoriumstudent beoordeeld moet worden. Met de moderne ICT-middelen waarover tegenwoordig beschikt kan worden, is het vrij eenvoudig om een gezongen lied op te nemen en dit met gespecialiseerde software te analyseren op kenmerken als toon, pitch, timbre, volume. Met enige inspanning lukt het ook nog wel om voor een specifiek lied per passage de gewenste meetwaarden ten aanzien van die kenmerken te beschrijven waarmee de weg naar een objectieve beoordeling van een creatief product zou zijn vrijgemaakt. Gelukkig willen de meesten niet zover gaan, het staat veel te ver af van waar het bij een lied eigenlijk om moet gaan, namelijk of je er door geraakt wordt. Maar vanwege het subjectieve karakter wordt dit in het geval een formele beoordeling geen goed idee gevonden. Er zal een tussenweg gevonden moeten worden die de te beoordelen bekwaamheid niet kapot analyseert en tegelijkertijd zoveel steun biedt aan de examinerator dat de beoordeling niet te veel mening en te weinig feitelijk is.

4. Construeren van een beoordelingsschaal bij een eenvoudige prestatie

De prestaties worden beoordeeld aan de hand van prestatiecriteria. De observaties die daarvan het resultaat zijn worden uitgedrukt in een score op een beoordelingsschaal. De eenvoudigste beoordelingsschaal heeft twee schaalpunten: 'voldaan/voldoende' of 'niet voldaan/onvoldoende'. Maar gelet op de ambitie van veel opleidingen om af te rekenen met de zesjes-cultuur is het raadzaam om meer schaalpunten, ofwel prestatieniveaus, te onderscheiden. De literatuur is niet eenduidig over het aantal prestatieniveaus dat onderscheiden moet worden, maar veel auteurs houden het op drie tot maximaal zes niveaus.

Naarmate er meer prestatieniveaus onderscheiden moeten worden, nemen de problemen om die prestatieniveaus zo helder en objectief mogelijk van elkaar te onderscheiden toe. Zo wordt bij nauwgezette bestudering van voorbeeldmatige beoordelingsmodellen in de literatuur al snel duidelijk dat het onderscheid in prestatieniveaus vaak gebaseerd is op één of hooguit enkele termen, die vaak verschillend geïnterpreteerd kunnen worden.

In figuur 4 wordt een beoordelingsmodel gepresenteerd voor het beoordelen van mondelinge presentaties. Een dergelijk beoordelingsmodel wordt een rubric genoemd. Het voorbeeld laat goed zien welke nadelen er kleven aan zo'n rubric. In de rubric worden de prestatieverschillen uitsluitend bepaald door de termen 'coherent', 'goed', 'gebrekig' en 'schiet ernstig tekort'. De vraag rijst wat hier onder verstaan moet worden.

Bekwaamheid: Mondeling presenteren			
<i>Beoordelingsschaal</i>	<i>Beoordelingsdimensies:</i>		
	<i>Structuur</i>	<i>Presentatie</i>	<i>Taalgebruik</i>
4	Gedurende de hele presentatie was sprake van een coherente organisatie; overgangen waren naadloos; introductie en conclusie waren effectief.	Voor iedereen uitstekend verstaanbaar; vloeiende presentatie met gevarieerde intonatie; ondersteunende lichaamstaal; oogcontact met alle toehoorders.	Gekozen woorden en samenstelling van de zinnen versterken de boodschap; nauwelijks grammaticale fouten.
3	Doorgaans een goede organisatie maar met hier en daar een niet-logische ordening van ideeën; duidelijk herkenbare introductie en conclusie.	Voldoende verstaanbaar; doorgaans vlotte presentatie; doorgaans effectieve lichaamstaal; voldoende oogcontact.	Effectief gebruik van taal ondersteunt de boodschap; enkele grammaticale fouten hebben geen invloed op de begrijpelijkheid.
2	Organisatie is gebrekkig; gepresenteerde ideeën zijn onvoldoende uitgewerkt; zwakke overgangen; introductie en conclusie bereiken doel niet.	Verstaanbaar schiet regelmatig tekort; presentatie stukt regelmatig; lichaamstaal en oogcontact werken niet ondersteunend voor de boodschap.	Het taalgebruik was niet altijd passend voor de over te brengen boodschap; grammaticale fouten zijn zo ernstig dat ze de begrijpelijkheid benadelen.
1	Organisatie schiet ernstig tekort; gepresenteerde ideeën zijn moeilijk te volgen; overgangen zijn plotseling en onlogisch; introductie en conclusie ontbreken.	Zeer slecht verstaanbaar; hakkende presentatie; storende lichaamstaal en ontbreken van oogcontact.	Taalgebruik detoneert met de over te brengen boodschap; zware grammaticale fouten maken het erg moeilijk om de boodschap te volgen.

*Figuur 4: Traditionele rubric voor het beoordelen van mondelinge presentaties.
(Voorbeeld aangepast overgenomen uit McMillan (2006, p.223))*

Een ander probleem met deze rubrics is dat de scoring moeilijker is dan op het eerste gezicht lijkt. In figuur 4 is te zien dat voor toekenning van score '4' de structuur van een mondelinge presentatie ten aanzien van alle drie prestatiecriteria gunstig beoordeeld

wordt. Echter, doordat de prestatiecriteria binnen een bepaalde beoordelingsdimensie niet afzonderlijk te scoren zijn, kan een examiner in verwarring raken over de toe te kennen score als bijvoorbeeld één van de prestatiecriteria niet gunstig beoordeeld kan worden. Nog een ander probleem van dergelijke rubrics is dat ze, met name als het gaat om complexe bekwaamheden, zeer omvangrijk kunnen zijn en dus veel leestijd vragen van een examiner. Zo neemt een veel gebruikte rubric voor het beoordelen van de schrijfvaardigheid van middelbare scholieren in Oregon (USA) zeven pagina's A4 in beslag.

Genoemde problemen geven voldoende aanleiding om te kiezen voor een ander, eenvoudiger type rubric. Een waarbij alle prestatiecriteria afzonderlijk gescoord dienen te worden en waarvan alleen het 'voldoende' prestatieniveau in zo objectief mogelijke termen beschreven is. Figuur 5 licht dit principe toe aan de hand van een (fictief) beoordelingsmodel voor het uitvoeren van een fysiotherapeutisch behandelplan. De waardering 'voldoende' wordt toegekend als het informeren van de cliënt over de (effecten van de) behandeling volgens de examiner voldoet aan de eisen zoals geformuleerd in een prestatie-indicator. Indien niet aan de betreffende eisen wordt voldaan, kent de examiner de waardering 'onvoldoende' toe. Als de examiner meent dat er ruimschoots aan de betreffende eisen voldaan wordt, kan hij de waardering 'goed' toekennen. En wanneer het informeren van de cliënt ernstig tekort schiet ten aanzien van de betreffende eisen, kan de examiner besluiten tot de waardering 'zwak'. Uiteraard zijn dergelijke beoordelingen niet vrij van subjectiviteit, maar de verschillen tussen de prestatieniveaus in de eerder besproken rubrics zijn ook maar zelden in volledig objectieve termen te beschrijven.

Beroepsbekwaamheid: Fysiotherapeutisch behandelplan uitvoeren				
<i>Beoordelingsdimensie: Informeert de cliënt over (de effecten van) de behandeling.</i>	<i>Beoordelingsschaal</i>			
Beroepsopdracht: A (Taaksituatie: bejaarde in verzorgingshuis met ademhalingsklachten)	4 <i>Goed</i>	3 <i>Voldoende</i>	2 <i>Onvoldoende</i>	1 <i>Zwak</i>
Prestatiecriteria:				
Stelt verschillende vragen om te achterhalen of de cliënt de huiswerk oefeningen uitvoert.				
Legt vooraf elke behandeling of oefening uit op een voor de cliënt begrijpelijke wijze.				
Enz.				

Figuur 5: Deel van een beoordelingsmodel voor 'Fysiotherapeutisch behandelplan uitvoeren'.

5. Construeren van een beoordelingsschaal bij een complexe prestatie

Bij een complexe prestatie, zoals bijvoorbeeld het uitvoeren van een onderzoek, kan het beoordelingsmodel snel erg uitgebreid worden. Op dat moment is het verstandig om de prestatiecriteria per beoordelingsdimensie te clusteren. Figuur 6 geeft een voorbeeld van een beoordelingsformulier voor het beoordelen van een onderzoeksverslag waarmee het onderzoekend vermogen van een student kan worden getoetst. De beoordelingsdimensies zijn gegroepeerd naar de fasen in de onderzoekscyclus. De prestatiecriteria benoemen wanneer het resultaat van die fasen goed is. De criteria zijn gebaseerd op het werk van Oost en Markenhof (2002) en Andriessen en Van Weert (2007, 2008). Meerdere prestatiecriteria zijn gecombineerd in één standaard per onderzoeksfase.

De beoordelingsschaal bestaat uit vier punten:

- OV onvoldoende: het resultaat van de fase voldoet niet aan de standaard
- V voldoende: het resultaat van de fase voldoet niet op alle punten aan de standaard maar wel in voldoende mate
- G goed: het resultaat van de fase voldoet op alle punten aan de standaard
- ZG zeer goed: het resultaat van de fase overtreft de standaard

Standaard per fase	Oordeel OV / V / G / ZG	Toelichting
Aanleiding en te bestuderen kwestie zijn inzichtelijk beschreven		
Doelstelling is onderbouwd, open, afgebakend en op hbo-niveau		
Het literatuuronderzoek en het theoretisch kader is inhoudelijk verankerd, relevant, precies en kritisch		
Centrale vraag en deelvragen zijn inhoudelijk verankerd, relevant, afgebakend, precies en functioneel		
De gebruikte methodologie is geëxpliciteerd, passend, onderbouwd en haalbaar		
Het onderzoek is vakkundig, betrouwbaar en controleerbaar uitgevoerd		
De conclusies zijn inhoudelijk verankerd, adequaat en valide		
De aanbevelingen zijn inhoudelijk verankerd, valide, adequaat, haalbaar, duurzaam en creatief		

Figuur 6: Beoordelingsformulier voor een onderzoeksverslag

Bij dit voorbeeld kan de volgende beslisregel worden gehanteerd om tot een beslissing en een cijfer te komen:

- Eén van de standaarden O -> cijfer onder de 5,5
- Alle standaarden V -> cijfer tussen 5,5 en 7,0
- Standaarden V en G -> cijfer tussen 7,0 en 8,0
- Alle standaarden G -> cijfer een 8,0
- Standaarden G en ZG -> cijfer tussen 8,0 en 9,0
- Alle standaarden ZG -> cijfer > 9,0

In kalibreersessies (zie bijlage 6) werkt het team van examinatoren van de opleiding aan een document dat een toelichting vormt op de standaarden die in het beoordelingsformulier worden gebruikt. In dat document worden alle prestatiecriteria toegelicht. Ook wordt per standaard geïllustreerd hoe een onderzoeksverslag er uit ziet dat voldoende is (de indicaties) en wat hoe een onderzoeksverslag er uit ziet dat onvoldoende is (contra-indicaties). De termen uit het toelichtende document kunnen een rol spelen bij:

1. Het beoordelen van het onderzoeksverslag;
2. Het geven van de toelichting op het oordeel;
3. Het geven van feedback aan de student bij de begeleiding van het onderzoek;
4. In het onderwijs over onderzoek.

Bijlage 6: Handreiking Kalibreersessies

1. Doel

De betrouwbaarheid van de beoordeling van eindwerkstukken kan worden verhoogd door de examinatoren gezamenlijk de prestatiecriteria en de score die behaald moet worden voor een voldoende resultaat te laten verhelderen. Een praktische vorm waarin dat kan is om eindwerkstukken van studenten gezamenlijk te reviewen in een kalibreersessie. Deze handreiking bevat hiervoor een opzet.

Aanvullend doel kan zijn om elkaar te trainen in de begeleiding van afstudeerders. In dat geval wordt niet alleen de vraag gesteld: 'Voldoet dit (onderdeel van het) eindwerkstuk aan de prestatiecriteria?', maar ook: 'Stel dat jij deze student begeleidt, welke begeleidingsstrategie zou jij kiezen om deze student te helpen om het resultaat nog beter te maken?'

2. Deelnemers

Aan de review nemen de examinatoren van de opleiding deel. Indien het doel ook het verbeteren van de begeleiding is, dan nemen ook begeleiders deel die geen examiner zijn. De reviewsessie wordt geleid door een moderator. De taak van de moderator is het gesprek te leiden en iedere keer te vragen naar verheldering van de argumentatie waarom een prestatiecriterium een bepaald oordeel heeft gekregen. Daarnaast is er een notulist aanwezig die alle punten notuleert die (kunnen) leiden tot verduidelijking van de prestatiecriteria. Tot slot is er een organisator die de sessies organiseert.

3. Verslaglegging

Verslaglegging kan in de vorm van een toelichting op het beoordelingsmodel met per prestatiecriterium een toelichting. Er kan gewerkt worden met een tabel. In de eerste kolom staat de prestatie-indicator. In de tweede kolom staan de 'indicaties' voor een positieve score op het criterium. Dit zijn eigenschappen van het werkstuk waaraan is te zien dat het aan het prestatiecriterium voldoet. In de derde kolom staan de 'contra-indicaties'. Dit zijn eigenschappen van het werkstuk waaraan je kan zien dat het werkstuk niet aan het prestatiecriterium voldoet.

4. Voorwaarden

Voor een succesvolle kalibreersessie is het volgende noodzakelijk:

1. Er zijn eindwerkstukken beschikbaar die reeds zijn beoordeeld. Het is aan te raden werkstukken van verschillend niveau te reviewen (bijvoorbeeld een onvoldoende, een voldoende en een goed). Bij voorkeur zijn de gegeven cijfers niet bekend bij de deelnemers.
2. De groep is veilig voor de deelnemers om te kunnen zeggen wat ze willen. Er mogen fouten worden gemaakt.
3. Er wordt afgesproken dat het doel van de sessie is om te komen tot verheldering van de criteria en de cesuur en niet om een oordeel te vellen over de student die het eindwerkstuk heeft gemaakt of over de docent die het heeft begeleid of beoordeeld.
4. De groep is in staat tot het voeren van een dialoog en voorkomt dat discussie ontstaat. Bij een discussie gaat het erom wie gelijk heeft en worden argumenten als wapens gebruikt. Een discussie is een handige gespreksvorm wanneer er een besluit moet worden genomen. Bij een kalibreersessie gaat het er echter om zoveel mogelijk van elkaar te leren. Voor dat doel is een dialoog veel geschikter. Die kenmerkt zich door het stellen van verdiepende vragen aan elkaar om op die manier er achter te komen waarom iemand een bepaald oordeel heeft.

5. De moderator heeft de vaardigheden om in de groep discussies om te vormen in dialogen door het stellen van vragen.
6. Er is genoeg tijd beschikbaar (2-2,5 uur).
7. Inhoudelijke deskundige examinatoren die op de hoogte zijn welke eindkwalificaties dienen te worden getoetst en inhoudelijke kennis hebben van de beroepsopdracht en hoe die kan worden uitgevoerd.

5. Voorbereiding

- De organisator zoekt geschikte eindwerkstukken om te reviewen. De organisator plant de bijeenkomst en zendt ter voorbereiding de werkstukken en het beoordelingsmodel met formulier toe.
- Alle deelnemers en de moderator reviewen vooraf de eindwerkstukken aan de hand van het beoordelingsmodel, geven een oordeel (onvoldoende / voldoende / goed of een cijfer, afhankelijk van welke schaal wordt gehanteerd) en nemen het ingevulde beoordelingsformulier mee naar de sessie.
- De moderator maakt een presentatie met daarin per prestatie criterium per werkstuk de belangrijkste informatie uit het werkstuk.

6. Uitvoering

1. Bij de start van de sessie vraagt de moderator naar het oordeel dat iedere deelnemer heeft gegeven. Zo kan een indruk verkregen worden van de interbeoordelaars-overeenstemming over de eindbeoordeling. Dit geeft aan wat de noodzaak is van verdere kalibratie.
2. Vervolgens vraagt de moderator om het oordeel op het eerste prestatie criterium voor de respectievelijke werkstukken. Wanneer er verschillen zijn in oordelen wordt om een argumentatie gevraagd. De moderator stimuleert de deelnemers om zoveel mogelijk de termen van de prestatie criteria bij hun argumentatie te gebruiken. Hierdoor ontstaat zicht op de betekenis die de deelnemers geven aan de prestatie criteria.
3. De deelnemers en de moderator gaan in de argumentatie op zoek naar aanvullende criteria die worden genoemd. Dit zoeken gebeurt door het stellen van vragen: waarom vind je dat? Waaruit leid je af dat het niet goed is? Wat versta jij onder... Wat vind jij goed genoeg e.d.? Deze criteria worden benoemd en beoordeeld wordt of die breed worden gedragen. Zo ja, dan worden ze toegevoegd aan het toelichtingsformulier.
4. Stap 2 en 3 worden herhaald totdat alle prestatie criteria zijn besproken.
5. Door de punten te noteren waarover afstemming is verkregen betreffende onvoldoende / voldoende en deze toe te voegen aan een toelichting bij het beoordelingsmodel ontstaat duidelijkheid over de standaard.

7. Afronding

Na afloop stuurt de notulist de toelichting die is ontwikkeld c.q. bijgesteld aan alle deelnemers toe.

Gebruikte bronnen

- American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC.
- Andriessen, D. G., & Van Weert, T. (2007). Kwaliteit van ontwerpprojecten. Wanneer is ontwerpgerichte kennisontwikkeling goed? Utrecht: Digitale Universiteit.
- Andriessen, D. G., & Van Weert, T. (2008). Kwaliteitszorg voor praktijkgericht onderzoek aan hogescholen. *Onderzoek van Onderwijs*, 37(1).
- Bulthuis, P. (2013). Het ZelCommodel, grip op competentieniveaus. De samenhang tussen zelfstandigheid en complexiteit. *Examens*, 10(2), 5–10.
- Butter, R. (2013). Role of research in higher prof. educ. on BA, MA & PhD levels is demarcated by the scope of the required relevance&rigour claims. #mmucarpe. Retrieved from <https://twitter.com/renebutter/status/398695835288223744>
- Commissie externe validering examenkwaliteit hoger beroepsonderwijs. (2012). *Vreemde ogen dwingen* (pp. 1–80). Den Haag. Retrieved from www.vereniginghogescholen.nl
- Expertgroep BKE/ SKE. (2013). *BESLISSEN IN HET HOGER BEROEPSONDERWIJS een basis- en seniorkwalificatie examinering*. Retrieved from www.vereniginghogescholen.nl
- Franssen, J. (2001). *Prikkelen presteren profileren*. Retrieved from <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2006/01/09/prikkelen-presteren-profileren.html>
- Griffioen, D., Visser- Wijnveen, G. J., & Willems, J. (2013). *Integratie van onderzoek in het onderwijs*. Groningen: Noordhoff Uitgevers.
- Griffioen, D., & Wortman, O. (2013). Onderzoek in het onderwijs van de Hogeschool van Amsterdam: op weg naar diversiteit in een uniform model. *Tijdschrift Voor Hoger Onderwijs*, 31(1/ 2), 16–31.
- HBO- raad. (2009). *Kwaliteit als opdracht*. Den Haag. Retrieved from www.vereniginghogescholen.nl
- HBO- raad. (2010). *Procedure op- en vaststelling landelijke opleidingsprofielen bachelor- opleidingen hogescholen*. Retrieved from http://www.vereniginghogescholen.nl/vereniging-hogescholen/publicaties/cat_view/43-publicaties/116-verenigingsafspraken-en-procedures
- Inholland, N.- C. O. H. (2011). *Rapport van Bevindingen Onderzoek Hogeschool Inholland*. Retrieved from www.nvao.net

- Joint Quality Initiative. (2004). Shared "Dublin" descriptors for Short Cycle, First Cycle, Second Cycle and Third Cycle Awards. Retrieved September 28, 2013, from <http://www.eua.be/>
- KNGF. (2006). *Het beroepsprofiel van de fysiotherapeut*. Amersfoort.
- Losse, M. (2012). Verbinding tussen onderzoek en onderwijs. Amersfoort: Presentatie Facta conferentie 11 december 2012.
- McMillan, J. H. (2006). *Classroom Assessment; Principles and practice for effective standards-based instruction* (4th ed.). Needham Heights (MA): Allyn & Bacon.
- NVAO. (2011). *Beoordelingskaders accreditatiestelsel hoger onderwijs*. Retrieved from www.nvao.nl
- NVAO. (2012). *Richtlijn NVAO beoordeling afstudeerwerken door panels tijdens het visitatieproces voor accreditatie en (indien van toepassing) een toets nieuwe opleiding*. Retrieved from www.nvao.net
- Oost, H., & Markenhof, A. (2002). *Een onderzoek voorbereiden*. Baarn: HBUitgevers.
- Oskam, I., Cowan, K., Hoiting, L., & Souren, P. (2012). *Ontwerpen van technische innovaties* (p. 352). Groningen: Noordhoff Uitgevers.
- Schon, D. (1991). *The reflective practitioner*. New York, NY: Basic Books.
- Van Berkel, H., & Bax, A. (2006). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu Van Loghum.
- Van der Linden, W. (2012). *A design-based approach to introducing student teachers in conducting and using research*. Eindhoven University of Technology.
- Van der Rijst, R. M. (2009). *De zes aspecten van een onderzoekende houding*. Retrieved from <http://www.ecent.nl/artikel/2000/view.do>
- Van Yperen, T., Veerman, J. W., & Bijl, B. (2013). Praktijkgestuurd effectonderzoek in de jeugdzorg. In H. De Jong, P. Tops, & M. Van der Land (Eds.), *Prikken in praktijken* (pp. 47–61). Den Haag: Boom Lemma Uitgevers.
- Verschuren, P. (2009). *De probleemstelling voor een onderzoek*. Utrecht: Spectrum.
- Yin, R. K. (2000). Rival explanations as an alternative to reforms as "experiments." In L. Bickman (Ed.), *Validity and social experimentation* (pp. 239–266). Thousand Oaks: Sage.