

Expertiseontwikkeling en Professionalisering van de Assessor

Ontwikkeling van een rubriek voor assessorenkwaliteit
Criteriumgericht Beoordelen

Onderzoek door Inge Oudkerk Pool

Domein Onderwijs en Opvoeding
Hogeschool van Amsterdam

1 juni 2013



Ziggy, Tom Wilson, 1983

Expertiseontwikkeling en Professionalisering van de Assessor

Ontwikkeling van een rubriek voor assessorenkwaliteit
Criteriumgericht Beoordelen

Onderzoek door Inge Oudkerk Pool

Domein Onderwijs & Opvoeding
Hogeschool van Amsterdam

1 juni 2013

Colofon

Expertiseontwikkeling en Professionalisering van de Assessor
Ontwikkeling van een rubriek voor assessorenkwaliteit
Criteriumgericht Beoordelen

Onderzoek door Inge Oudkerk Pool binnen het lectoraat Leren & Innoveren
onder leiding van Marco Snoek, lector Kenniscentrum Onderwijs en Opvoeding
1 juni 2013

Uitgave

Kenniscentrum Onderwijs en Opvoeding, Hogeschool van Amsterdam

Drukwerk

OCE-Copyshop

Bestellen

U kunt deze publicatie bestellen bij het Kenniscentrum Onderwijs en Opvoeding via het bestelformulier op de website:
www.hva.nl/kenniscentrum-doo

Tekst-editing en Lay-out

Kenniscentrum Onderwijs en Opvoeding
Hogeschool van Amsterdam

Inhoudsopgave

Voorwoord	2
Introductie	3
1. Aanleiding, doel en vraagstelling van dit onderzoek	4
1.1 Aanleiding	4
1.2 Historie assessorentrainingen hoger onderwijs in het bijzonder de educatieve faculteiten	4
1.3 Doelstelling van het onderzoek	6
1.4 Onderzoeksvragen	6
1.5 Begripsverheldering.....	6
2. Theoretisch kader	8
2.1 Toetsvormen.....	8
2.2 Competentietoetsing binnen de lerarenopleidingen	9
2.3 Expertiseontwikkeling criteriumgericht beoordelen	13
3. Opzet en methodische verantwoording onderzoek.....	15
3.1 Type onderzoek.....	15
3.2 Opzet van het onderzoek	15
3.3 Respondenten	17
3.4 Dataverwerking.....	18
4. Resultaat kwaliteitsrubriek criteriumgericht beoordelen.....	19
4.1 Verwerking ruwe onderzoeksgegevens	19
4.2 Kwaliteitsrubriek criteriumgericht beoordelen.....	21
4.3 Kwaliteitsrubriek als ontwikkelingsgericht instrument.....	27
5. Conclusies en aanbevelingen.....	28
5.1 Antwoord op de onderzoeksvragen.....	28
5.2 Drie niveau's in assessoren.....	28
5.3 Vervolg vragen naar aanleiding van het onderzoek	29
5.4 Aanbevelingen	30
6. Samenvatting.....	31
7. Literatuur.....	33

Voorwoord

Toetsen is onlosmakelijk verbonden aan lesgeven en leren. Leren is immers een cyclisch proces waarin leeractiviteiten voortdurend geëvalueerd moeten worden om op basis van de evaluatie-uitkomsten het leerproces bij te sturen of om weer stappen verder te zetten. 'Zijn de doelen (in voldoende mate) bereikt? En welke stappen kunnen nu verder gezet worden?'

Die evaluatie heeft niet alleen een betekenis voor de leerder zelf, maar ook voor de omgeving: het is erkenning van behaald niveau, vaak met een civiel effect (in termen van formele afsluiting of juist toelating, een diploma, etc.).

In beide gevallen moet toetsing op een zorgvuldige wijze plaats vinden, maar als er voor de kandidaat veel van afhangt, dan luistert het extra nauw: de beoordeling moet transparant zijn, objectief (of minimaal intersubjectief) zijn, en de beoordeelde zou zich in het oordeel moeten kunnen herkennen. Dat stelt dus ook hoge eisen aan degene die de toetsing uitvoert en het oordeel uitspreekt: assessor.

Aandacht voor de kwaliteit van de assessor en voor de ontwikkeling daarvan is daarom belangrijk. In dit verslag vertelt Inge Oudkerk Pool over haar onderzoek naar kwaliteit en kwaliteitsniveaus van assessoren. Dat onderzoek is uitgevoerd binnen de kenniskring van het lectoraat Leren & Innoveren van de HvA. Achtergrond is haar werk als motor achter de ontwikkeling van integratieve assessments binnen de lerarenopleidingen, binnen het domein Onderwijs en Opvoeding van de HvA, als ook landelijk. Vanuit die rol heeft ze ook aan de basis gestaan van de assessorentrainingen die in de afgelopen jaren ontwikkeld en uitgevoerd zijn.

Het ontwikkelings- en uitvoeringswerk is nu onderbouwd door middel van onderzoek. Centrale vraag is of er verschil is tussen assessoren met en zonder training en met en zonder certificering. Die onderzoeksvraag is niet alleen belangrijk voor de legitimering van de trainingen en certificeringstrajecten voor assessoren, maar ook om zicht te krijgen op professionele ontwikkeling van assessoren en om een instrument te ontwikkelen dat die professionele ontwikkeling kan ondersteunen.

Het resultaat van dit onderzoek is een kwaliteitsrubriek (rubric) die gebruikt kan worden om assessoren te beoordelen, maar ook om assessoren te helpen in de reflectie op hun professionele ontwikkeling.

Daarmee is het een mooi voorbeeld van een praktijkgericht onderzoek dat uitmondt in een concreet instrument dat een belangrijke bijdrage levert voor de ondersteuning van de praktijk, in dit geval de assessorentrainingen.

Het onderzoek laat ook zien dat hbo-opleidingen en hun assessoren eigenlijk niet tevreden kunnen zijn met een tussenniveau, maar dat alle assessoren (uiteindelijk) het eindniveau van expert assessor zouden moeten behalen om studenten op een transparante, eerlijke en herkenbare wijze te kunnen beoordelen.

En daar ligt een belangrijke uitdaging voor het hbo!

Marco Snoek
Lector Leren & Innoveren
Kenniscentrum Onderwijs en Opvoeding HvA

Introductie

Vanaf 1998 heb ik me binnen het domein Onderwijs en Opvoeding van de HvA en haar rechtsvoorgangers bezig gehouden met portfolio assessments en de daarbij behorende assessorentrainingen. Als je al zolang bezig bent om het leerpotentieel en/ of het actuele handelingsniveau van kandidaten en studenten vast te stellen en te verantwoorden, dan is het onderzoek doen naar expertiseontwikkeling en professionalisering van de assessor eigenlijk heel vanzelfsprekend.

Toch had ik het zetje van mijn toenmalige domeinvoorzitter Marjan Freriks nodig om de stap naar praktijkgericht onderzoek te maken. Ik ben van nature niet zo'n onderzoeker en zie mezelf veel meer als ontwikkelaar, trainster en assessor.

Terugkijkend op afgelopen tweeënhalf jaar voel ik de voldoening van het vastleggen van de opgedane leerervaringen van vijftien jaar assessorentrainingen in onderzoeksresultaten. Het feit dat er nu een eerste stap is gezet in het benoemen van de verschillende interventies, die assessoren plegen tijdens de beschreven professionaliseringsfasen ongetraind, getraind en expert assessor zie ik als een startpunt om de beoordelingskwaliteit van opleiders transparant te maken.

Hierdoor wordt zichtbaar op welke manier en hoe een beoordelaar tot zijn beoordeling komt en of dit met voldoende kwaliteit gebeurt.

Hopelijk leidt de kwaliteitsrubriek criteriumgericht beoordelen tot discussie, zelfevaluatie en professionalisering van beoordelaars die criteriumgericht aan het begeleiden dan wel aan het beoordelen zijn.

Ik verwacht dat de kwaliteitsrubriek komende maanden in het assessorenveld verder aangescherpt zal gaan worden. Het is bedoeld als hulpmiddel om de kwaliteit van de beoordelaar mede te bepalen en het ontwikkelingspad op het gebied van criteriumgericht beoordelen vast te stellen.

Als het goed is wordt de rubriek in de praktijk toegepast, pas dan is dit praktijkgericht onderzoek geslaagd!

Een woord van dank is hier op zijn plaats. Ik dank in de eerste plaats de kenniskring van het lectoraat Leren & Innoveren van de HvA. De kenniskring en met name Marco Snoek zijn vanaf de eerste concepten betrokken geweest en hebben kritisch en constructief mijn onderzoeksproces geobserveerd en besproken. Het gevolg van de inspirerende bijeenkomsten was dat ik met nieuwe ideeën aan de slag kon.

Naast de kenniskring heb ik veel steun en inspiratie ondervonden van mijn drie collega's, gecertificeerd en expert assessoren Anja Stigter, Paulien Langedijk en Annelien Haitink. Zij hebben at random DVD-banden geanalyseerd en beoordeeld, waardoor mijn onderzoek de zo belangrijke inter-beoordelaars betrouwbaarheid verkreeg. Samen met Annelien heb ik de kwaliteitsrubriek meerdere malen aangepast en verbeterd tot de versie die het nu heeft. Ik dank ook Antoinette van Berkel van het EVC Centrum van de HvA, die mij op het spoor heeft gezet om de kwaliteitsrubriek ook als ontwikkelingsgericht instrument te gebruiken. Zij liet mij hiervoor inspiratie op doen bij de reflectievragen uit het profiel docent Basiskwalificatie Onderwijs (BKO) van de HvA.

Dank ook aan de assessoren (binnen en buiten de HvA) die hun assessmentopnames ter beschikking stelden voor het onderzoek en aan de assessoren die ik 'live' heb kunnen observeren. Vanwege de gegarandeerde anonimiteit worden ze niet genoemd. Ze weten zelf wel dat ik hen bedoel.

Als laatste dank ik Annette Dietze van de afdeling Strategie en Ontwikkeling van het domein Onderwijs en Opvoeding voor het dagdeel dat ik afgelopen twee jaar heb mogen en kunnen gebruiken voor mijn onderzoek en participatie in de kenniskring Leren & Innoveren.

Inge Oudkerk Pool

1. Aanleiding, doel en vraagstelling van dit onderzoek

1.1 *Aanleiding*

Toetsen is een vak apart en niet elke goede opleider is direct ook een goede beoordelaar. Toetsen en leren zijn onlosmakelijk met elkaar verbonden en voor het verkrijgen van beter onderwijs zullen onderwijs en toetsing steeds beter geïntegreerd en op elkaar afgestemd dienen te worden.

Professionalisering van de beoordelaars is daarvoor noodzakelijk. Naast de kwaliteitsslag bij bestaande docenten zou toetsen een onderdeel moeten gaan uitmaken van het BKO-traject voor nieuwe docenten (Basiskwalificatie Onderwijs). Op deze manier verzekert het hoger onderwijs zich ervan dat iedere opleider geschoold is in het beoordelen van studenten.

Naast de eigen interne kritische blik van hogescholen, opleidingen en kwaliteitcommissies (denk aan de interne audits), worden er vanuit de politiek en maatschappij ook steeds meer vraagtekens gezet bij de kwaliteit van toetsing. Een recent rapport is 'Vreemde ogen dwingen' (Eindrapport Commissie externe validering examenkwaliteit Hoger beroepsonderwijs april 2012), waar onder andere adviezen staan geformuleerd over de externe validering van de kwaliteit van toetsen, zoals bijvoorbeeld door middel van het certificeren van beoordelaars.

Iedereen is het er over eens: de subjectiviteit dient uit de beoordeling te verdwijnen.

In het curriculum van het hoger onderwijs nemen integratieve toetsen een bijzondere plaats in. Integratieve toetsen zijn toetsen die aan studenten vragen om het geleerde in complexe beroepscontexten op te roepen en toe te passen. Het is voor studenten niet meer toereikend om een reeks toetsbriefjes per vak te verzamelen en deze in te leveren ter verkrijgen van het diploma. Er is meer nodig om het hbo-niveau aan te tonen. Het integraal kunnen terugkijken op het leertraject, zodat de student beargumenteerd weet aan te geven waar hij staat en welke stappen nodig zijn voor een volgende fase van de opleiding is een belangrijk onderdeel van het kwalificeren als hbo-professional. Juist de integratieve toetsvormen bieden zicht op het hbo- en het ontwikkelde competentieniveau van studenten. Voorbeelden van integratieve toetsvormen zijn portfolio's, bekwaamheidsproeven, stagebeoordelingen.

De meeste lerarenopleidingen kennen twee en soms zelfs drie portfolio assessments. Daarnaast wordt de methodiek van de portfolio-assessments volop ingezet bij de EVC en de Zij-instroom in Beroep (ZiB) assessments. Elke lerarenopleider krijgt dan ook te maken met studenten die een portfolio assessment hebben gedaan. In bijna alle gevallen krijgt de docent zelf de taak als beoordelaar (assessor) om studenten te beoordelen aan de hand van portfolio assessments. Omdat toetsen onder andere de functie hebben om misclassificatie te voorkomen, is het van belang dat toetsen niet alleen zorgvuldig geconstrueerd zijn maar ook zorgvuldig uitgevoerd en geïnterpreteerd worden, zodat de juiste beslissingen worden genomen.

Bij deze integratieve toetsvormen hangt de kwaliteit van de toets voor een groot deel af van de uitvoerder. De opleider zal zich moeten professionaliseren in het afnemen van integratieve toetsen, zoals bijvoorbeeld het analyseren van een portfolio. Toetsdeskundigen in het hoger onderwijs constateren dat het met de kennis en kunde van opleiders over toetsen niet goed is gesteld (Riksen, 2010). Opleiders in het hoger onderwijs zijn vaak vakexperts en zijn niet of nauwelijks opgeleid in toetsing. Resultaat is dat opleiders onvoldoende kennis hebben van de basisprincipes, procedures en technieken die bij integratieve toetsen horen. Gevolg hiervan is dat de beoordeling sterk subjectief getint is en te afhankelijk van de opleider. Studenten ervaren dan ook grote willekeur bij beoordelingen.

Daarom is het belangrijk om aandacht te besteden aan de kwaliteit van de assessor. Die kwaliteit staat in dit onderzoek centraal.

1.2 *Historie assessorentrainingen hoger onderwijs in het bijzonder de educatieve faculteiten*

Al vanaf het jaar 2000 is er aandacht voor professionele kwaliteit en de professionele ontwikkeling van assessoren.

De kwaliteit van het assessment en het gebruik van het instrumentarium lijkt immers voor een groot deel af te hangen van de kwaliteit van de assessor. Assessoren hebben de taak om vast te stellen of het niveau dat de student aantoonde in zijn portfolio en tijdens het criteriumgericht interview (CGI) voldoende is om een volgende fase in te gaan. Verder dienen de beoordelaars te verantwoorden waarop de beoordeling is gestoeld.

Elke student heeft recht op een assessor die de EVC's dan wel het competentieniveau her- en erkent en de kwaliteiten ten toon spreidt, die een student van een assessor mag verwachten.

De student verwacht dat de beoordelaar op basis van een analyse van het portfolio weet wat hij nog wilt weten, focus en regie houdt en specifieke vraagmethodieken gebruikt om tot de vereiste informatie te komen.

De assessor voegt de verkregen informatie uit het portfolio en het CGI samen en komt tot een verantwoorde beoordeling en heeft hiermee de input voor het nemen van de juiste beslissingen.

Ook de examencommissie moet op de betrouwbaarheid en de validiteit van een assessmentinstrument en de kwaliteit van de assessor kunnen vertrouwen. Zeker bij EVC- en ZiB assessments waar leerwegaafhankelijk getoetst wordt en waarbij het erkennen en waarderen van opgedane leerervaringen en competenties het doel is, dient de assessorenkwaliteit een voorwaarde te zijn, waar examencommissies vanuit moeten kunnen gaan.

Het bewaken van de kwaliteit van de beoordelaar en de beoordelingen (mondeling en schriftelijk) is dan ook cruciaal voor alle betrokkenen (kandidaten/studenten, werkgevers, opleiding).

Vanaf de start van het competentiegerichte toetsen en beoordelen worden er (basis)trainingen georganiseerd voor de assessor/beoordelaar. Hoofddoel van de training is om de assessoren zodanig te bekwamen, dat de portfolio beoordeling op een verantwoorde en bekwame manier plaatsvindt. Om dit te bereiken richt de training zich op het eenduidig leren toepassen van de procedure en de methodiek van het assessen. Via specifieke vraagmethodieken (zie pagina 9 en 10) komt de assessor – en nogmaals: dat is in bijna alle situaties “de opleider zelf”- tot de kern van het niveau waar een student zich bevindt.

De assessorentraining is een eerste begin van de ontwikkeling van de beoordelingskwaliteit van opleiders. Het certificeringstraject dat assessoren daarna kunnen doorlopen, biedt opleidingsverantwoordelijken de waarborg van de vereiste kwaliteit van assessoren. Hiermee heeft men een professionaliseringstraject in handen, waarmee de beoordelingskwaliteit van opleiders geborgd wordt.

De basis assessorentraining heeft zich afgelopen tien jaar steeds verder ontwikkeld. De assessorentraining omvat minimaal vijf dagdelen. De opleiders maken kennis met de beoordelingstheorie en methodiek en oefenen de beoordelingsvaardigheden eerst op elkaar, daarna op een acteur en vervolgens op de studenten in authentieke toetsituatie(s) van de opleiding. Aan het eind van de training verkrijgen de deelnemers het deelnemerscertificaat. Dit is een eerste stap tot professionele beoordelingskwaliteit. Het doen van de basis-assessorentraining en het afnemen van assessments geeft nog geen garantie tot bekwame uitvoering van criteriumgericht beoordelen. De assessment methodiek is nog onvoldoende verinnerlijkt. Het blijkt dat het gericht verder oefenen en het krijgen van feedback door collega – assessoren en assessorenexperts een essentiële vervolgstap is op weg naar het beoogde kwaliteitscertificaat.

In de afgelopen jaren is daarom door het HvA EVC Centrum het assessorenprofiel vastgesteld. Dit profiel beschrijft de kerntaken die de assessor uitvoert, de basiskwalificaties en de kwaliteiten waarover de assessor dient te beschikken (hierover later meer, pagina 11). Dit is de basis geweest voor het ontwikkelen van een beoordelingsprocedure op basis waarvan een certificaat uitgereikt kan worden. Het niveau van de beoordeling en het certificaat overstijgt het niveau van de basis-assessorentraining.

Uit de eerste ervaringen van assessoren (uit domein Onderwijs en Opvoeding (DOO) van de HvA) met de certificering (2010-2013) blijkt dat de “kwaliteitsstandaard- assessoren” een haalbare maar ook hoge standaard is. Het opgaan voor certificering bij het EVC Centrum zonder voorbereidend coachingstraject, gaf een situatie waarin teveel assessoren zakten voor de certificering (drie van de vijf assessoren slaagden niet voor certificering in 2010). In dat zelfde jaar is in reactie hierop, ter voorbereiding op de certificering is het coachingstraject ontwikkeld en uitgevoerd.

Er is sinds 2011 sprake van een certificeringstraject. De cijfers zijn nu heel anders, assessoren die het certificeringstraject hebben doorlopen slagen bijna altijd (tot december 2012 zijn veertien van de vijftien assessoren van DOO direct geslaagd en één assessor is via de herkansing gecertificeerd). De verwachting is dat door certificering de kwaliteit van het criteriumgericht beoordelen in het gehele HBO-curriculum bevorderd wordt, aangezien dezelfde kwaliteiten gebruikt worden bij het beoordelen van andere complexe integratieve toetsen. Bijvoorbeeld bij het beoordelen (en begeleiden) van de stage- en projectopdrachten.

Het onderzoek dat in dit rapport beschreven wordt, wil antwoord geven op de vraag of er inderdaad verschil in interventies waar te nemen is tijdens het criteriumgericht interview (CGI) tussen de verschillende groepen van assessoren (ongetraind, getraind en expert), en op de vraag of er verschillende professionele niveaus van criteriumgericht beoordelen zijn te beschrijven in de vorm van een kwaliteitsrubriek (rubric) criteriumgericht beoordelen?

1.3 Doelstelling van het onderzoek

Er zijn twee doelstellingen in dit onderzoek:

- Aan de hand van observeerbare kenmerken van interventies de professionele niveau - verschillen tussen assessoren (ongetraind, getraind en expert) beschrijven. De beschrijving van de professionele ontwikkeling levert zo mogelijk een kwaliteitsrubriek op van interventies, die je mag verwachten van beoordelaars die in één van drie professionele ontwikkelingsniveaus zitten. In de kwaliteitsrubriek worden kwaliteiten beschreven aan de hand van de interventies die de assessor pleegt tijdens het CGI: het maakt transparant op welke manier de assessor zijn kwaliteiten inzet om te weten te komen wat hij wilt weten. Wanneer de assessor de interventies uitvoert die verwacht worden bij het gecertificeerd niveau kan men spreken van "expert assessor". In de onderzoeksgroep zitten assessoren die zich in de basistraining bevinden (ongetraind) of in het certificeringstraject zitten (getraind) dan wel al gecertificeerd assessor zijn. De laatste groep heeft het kwaliteitscertificaat assessor ontvangen.
- Als tweede hoopt dit onderzoek legitimering en rechtvaardiging te bieden voor de investering in trainingen en certificering van de assessoren. Dit alles heeft ten doel de kwaliteit van competentiegericht beoordelen en begeleiden te bewaken en te borgen.

1.4 Onderzoeksvragen

Op basis van de bovenstaande doelstelling komen we tot de volgende onderzoeksvragen: Is er verschil te observeren bij de 3 onderzoeksgroepen (ongetraind, getraind en expert) in het gebruik van de interventies die de assessoren bij het criteriumgericht beoordelen hanteren tijdens de afname van het CGI? Het onderzoek richt zich vooral op de eerste vijf onderdelen van het assessorenprofiel¹ (zie noot 1).

Als blijkt dat er verschil is in de hantering van de onderdelen van het assessorenprofiel, zijn er dan patronen waar te nemen in de interventies die de assessoren gebruiken?

Is er een kwaliteitsrubriek criteriumgericht beoordelen te beschrijven waarin de drie verschillende professionele niveaus (ongetraind, getraind en gecertificeerd) qua interventies met betrekking tot criteriumgericht beoordelen zichtbaar zijn.

1.5 Begripsverheldering

Ongetrainde assessor:

Beoordelaar heeft nog geen basistraining assessor gevolgd of bevindt zich in de basis assessorentraining.

Getrainde assessor:

Beoordelaar heeft de basistraining assessor afgerond en wordt ingezet bij de uitvoering van assessments. Dit kunnen zowel opleidingsassessments als ook EVC -en/of ZiB assessments zijn. De beoordelaars kunnen in het voorbereidende certificeringstraject zitten en hebben op eigen initiatief gekozen om zich verder te ontwikkelen tot gecertificeerd assessor.

¹ Opgesteld door het HvA EVC Centrum en opgenomen in het erkende assessorenprofiel assessor.

1. *Gespreksstructuur*

2. *Gesprekstechniek (START, turbo, LSD, metacommuniceren, etc.)*

3. *Relatie bewijsstukken, beroepshandelingen en beoordelingscriteria*

4. *Beoordelingstechniek*

5. *Feedback geven*

6. *Omgaan met bezwaren (indien van toepassing)*

Expert assessor:

Assessoren zijn gecertificeerd als assessor via de beoordelaars en de certificeringsprocedure HvA EVC Centrum² dan wel via beoordelaars van het ADEF³ netwerk Assessments en de ADEF certificeringsprocedure.

Certificeringstraject tot assessor:

Drie bijeenkomsten waarvan de eerste twee collegiale consultatie centraal staat aan de hand van DVD-opnamen en het assessorenprofiel. De laatste keer krijgt de assessor individuele feedback tijdens een CGI afname van één van de trainers.

Certificeringsprocedure HvA EVC Centrum:

Op basis van het assessorenprofiel (pagina 11 en 12) wordt de assessor tijdens een opleidingsassessment beoordeeld door de beoordelaar. Dit is een authentieke beoordelingssituatie.

Certificeringsprocedure ADEF netwerk Assessments:

Op basis van het assessorenprofiel (pagina 11 en 12) stelt de assessor zijn assessorenportfolio samen en wordt hierop via het CGI door twee beoordelaars beoordeeld.

² Het HvA EVC Centrum heeft het Hobéon Keurmerk voor Kwaliteit voor het Kwaliteitskader Assessmentcentra ontvangen in juli 2012. De beoordeling van assessoren tot gecertificeerd assessor is hier een onderdeel van.

³ ADEF staat voor Algemeen Directeurenoverleg Educatieve Faculteiten

2. Theoretisch kader

2.1 Toetsvormen

Onderwijs heeft haar uitgangspunten in leerdoelen. Leerdoelen zijn de start van het onderwijsproces, toetsen zijn het eindpunt (Tyler, R.W. 1949).

Toetsen staat aldus niet los van het onderwijs en de kwaliteit van de toetsing dient verantwoord te worden. Dit betekent dat elke opleider zich zal moeten professionaliseren in toetsing. Leren en toetsen horen altijd bij elkaar.

Het beoordelen van ontwikkelingen van het leerproces (formatieve toetsing) en van de behaalde resultaten van dat leerproces (summatieve toetsing) neemt een wezenlijk deel in van de curricula binnen het hoger onderwijs. Er zijn vier verschillende toetsvormen te onderscheiden, die gezamenlijk het toetshuis vormen van een opleiding (zie het cahier Toetsen en Beoordelen, Kayzel, Nieweg & Kok, 2006).

Toetsing van kennis

De theoretische concepten van het onderwijsprogramma worden getoetst. De kennis die essentieel geacht wordt voor een bepaald beroep wordt ook wel de kennisbasis van een beroep genoemd. Deze kennisbasis is een elementair onderdeel van elke beroepsopleiding. Gedurende de gehele opleiding wordt de kennisbasis getoetst om het niveau van de kennisbasis bij de student vast te stellen.

Toetsen die hier onder vallen zijn de zogenaamde kennis-, essay-, casus-, voortgang- en overall toetsen.

De onderwijspsycholoog Benjamin S. Bloom ontwikkelde in de periode tussen 1940 en 1959 zijn taxonomie (1956) die nog steeds veel gebruikt wordt om het niveau van toetsing van kennisvelden te onderzoeken. Er was destijds behoefte aan een referentiekader dat standaard-classificaties bevat(te).

De redenering was (is): als er duidelijkheid is, maakt dat de onderlinge uitwisseling van informatie tussen onderwijsgevend en -gevend gemakkelijker. Dit leidde tot de taxonomie die een classificatie geeft van leergedrag(ingen) dat (die) de lerende persoon kan tonen als bedoelde uitkomst(en) van onderwijsprocessen (Oudkerk Pool, T. 1994). De taxonomie van Bloom geeft een classificatie voor het cognitieve domein, waarbij zes hoofdcategorieën en een flink aantal onderverdelingen beschreven zijn. De categorieën lopen op in moeilijkheid en de mate waarin ze prikkelen tot nadenken.

Vaak roept een hoger niveau van vragen vanzelf vragen op een lager niveau op. Hierdoor worden verbanden gelegd en wordt het begrip versterkt.

De zes hoofdcategorieën zijn:

- Kennis
- Begrip
- Toepassing
- Analyse
- Synthese
- Evaluatie

Toetsing van vaardigheden en attitude

Toetsen die hier bij passen zijn practicumtoetsen, hands-on demonstraties, presentaties van generieke vaardigheden, reflectieverslagen. De kernvaardigheden die de student moet beheersen voor de uitoefening van het beroep, worden geoefend binnen het opleidingscurriculum en zijn bij voorkeur beroepsgerelateerde vaardigheden. Hoe verder men in de opleiding komt, des te moeilijker en complexer de opdrachten worden tot uiteindelijk het startbekwaamheidsniveau van het beroep is bereikt.

Toetsing van beroepsmatig handelen

Toetsen die hier onder vallen zijn stage-, praktijk-, projectopdrachten, beroepssimulatie assessments en het afstudeerwerkstuk. Hier wordt de ontwikkeling en het resultaat van het werkplekleren getoetst met behulp van de feedback van de stagebegeleider op de werkplek en beoordeling van het stageverslag.

Een voorbeeld van een praktische toets - zoals het gedragsassessment- is een probleem dat is geformuleerd in een casus: "Hoe zou jij een uit de handlopende discussie stoppen tussen twee leerlingen die in jouw klas aan het (bek)vechten zijn?".

In het toetshuis van opleidingen is vastgelegd hoe de student de kritische vaardigheden op de werkplek moeten tonen en beoordelen.

Toetsing van competenties en het daarbij behorende hbo-niveau

De competenties worden getoetst via portfolio-assessments. Een portfolio-assessment maakt het mogelijk om de competentieontwikkeling en het hbo-niveau van een student te beoordelen. Dit gebeurt op basis van informatie uit verschillende bronnen met gebruikmaking van (beroeps)situaties en over een langere tijd. Dit is één van de grote krachten van deze manier van toetsen: voor een goede beoordeling van competenties zijn namelijk meer momenten en metingen nodig. In het portfolio zijn de verschillende bewijzen opgenomen. Van de student wordt verwacht dat hij terugkijkend op zijn leerproces die onderdelen selecteert als bewijs en bij elkaar zet in zijn portfolio die met elkaar het gevraagde competentieniveau aantonen. Dit zijn precies de hogere toetsniveaus van Bloom: analyse en synthese. Maar ook het toetsniveau 'evaluatie' wordt gevraagd, de student dient immers in reflectie te verantwoorden waarom hij bepaalde keuzes heeft gemaakt voor bijvoorbeeld een specifieke methodiek.

2.2 Competentietoetsing binnen de lerarenopleidingen

Binnen de lerarenopleidingen speelt competentiebeoordeling een belangrijke rol (Geschiktheidsonderzoek voor zij-instromers in de praktijk, Stoas Onderzoek 2004 en later ook Assessments in de lerarenopleiding, ADEF 2010). Doel van deze competentiebeoordeling is om op integratieve wijze te kunnen beoordelen of de kandidaat de beoogde competenties beheerst. Dit impliceert het inzetten van kennis, vaardigheden en attitude in complexe situaties waarin beroepsmatig gehandeld moet worden.

Competentietoetsing wordt veelal gebruikt om uitspraken te doen over de integratieve bekwaamheid van de kandidaat voor toelating tot de (verkorte) opleiding, voor een specifieke fase van de opleiding (LIO) of voor de afronding van de opleiding.

De uitspraken die door de beoordelaars bij een dergelijke competentietoetsing gedaan worden hebben een belangrijk civiel effect in termen van toelating tot de opleiding, tot de LIO-fase of tot het beroep. Om die reden is betrouwbaarheid en validiteit in de uitspraken van de assessoren van groot belang, zowel naar de kandidaten toe, als richting de maatschappelijke omgeving van de lerarenopleidingen.

Taken van de assessor

In het assessorenprofiel dat is opgesteld door het EVC Centrum van de HvA en het landelijke ADEF netwerk Assessment staan de taken van de assessor beschreven.

De assessoren hebben de taak om in een korte tijd zicht te krijgen op de bekwaamheid van de kandidaat en tegelijk de student te laten zien waar hij staat ten opzichte van de gevraagde beoordelingscriteria (feedback). De assessoren dienen een adequate en verantwoorde beoordeling te maken, die recht doet aan de tot dan toe verworven competenties en aan het vereiste hbo-niveau. Een belangrijk onderdeel van het portfolio-assessment is het criteriumgericht interview. Centraal staat de kwaliteit van de reflectie door de student op zijn eigen ontwikkeling in het licht van de gestelde doelen. De uitkomst hiervan kan zijn dat hij inderdaad aannemelijk heeft gemaakt en getoond dat de vereiste competenties op het gewenste niveau verworven zijn. De beoordelaars kunnen echter ook om aanvullend materiaal vragen, of om een aanvullend gedragsassessment als de competenties en het bijbehorende hbo-niveau onvoldoende zijn bewezen door de student.

Voor alle assessoren geldt dezelfde assessmentsituatie:

Individueel beoordelen van het portfolio;

Bepalen waar de focus op ligt tijdens het CGI ofwel: wat moet de assessor nog te weten komen om een verantwoorde en kwalitatieve beoordeling te geven? Dit gebeurt in het assessorenoverleg voorafgaand aan het CGI;

Het uitvoeren van het CGI;

Komen tot de beoordeling. Onderling overleg met medeassessor na afloop van het CGI ;

Het geven van de mondelinge beoordelingsrapportage;

Het schrijven van de beoordelingsrapportage.

Het CGI

De criteriumgerichte interview (CGI) vindt plaats nadat de assessoren het portfolio hebben bestudeerd en een eerste individuele beoordeling is gemaakt. In het vooroverleg brengen beide assessoren hun bevindingen samen en komen tot een beperkt aantal kernvragen, die ze in het CGI willen uitvragen.

Het CGI is dan ook het assessmentonderdeel waarbij de assessor onduidelijkheden met betrekking tot bewijsvoering alsmede vragen die voortkomen uit de bestudering van het portfolio verder onderzoekt. De assessor heeft hiervoor een aantal gedragsindicatoren tot zijn beschikking om inhoudelijk en op professionele wijze te beoordelen of de student voldoet aan het vereiste competentieniveau. De vaardigheden die een bekwame assessor hanteert hebben betrekking op de volgende aspecten:

- Gespreksstructuur
- Gesprekstechniek (START, turbo, LSD, metacommuniceren, etc.)
- Relatie bewijsstukken, beroepshandelingen en beoordelingscriteria
- Beoordelingstechniek
- Feedback geven
- Omgaan met bezwaren (indien van toepassing)

Het CGI duurt tussen de dertig en negentig minuten. Over het algemeen duren de opleidings-CGI's rond het half uur à drie kwartier en een EVC-assessment duurt rond de negentig minuten. Alle CGI's hebben tot doel om het competentie niveau van de student zo precies mogelijk te bepalen. Hiervoor passen de assessoren verschillende vraag- en gesprekstechnieken toe. Een drietal gesprekmethodieken staat in het assessorenprofiel en in de trainingen centraal. Dit zijn de START-methodiek, de LSD en TURBO-vragen. Hieronder worden ze kort toegelicht.

De START-methodiek wordt gebruikt bij één van de gesprekskernen waarbij de assessor extra doorvraagt en de kern gedegen bloot legt. De assessor komt hierdoor te weten waar de student zich op de beoordelingsstandaard bevindt. Met globale oppervlakkige antwoorden wordt geen genoegen genomen. Bijvoorbeeld "Hoe leer jij ze met elkaar communiceren?" Student antwoordt in algemene zin "We hebben op school regels zoals naar elkaar luisteren." Hier vraagt de assessor op door "Op welke manier pas je de regels dan toe met je klas, geef eens een concreet voorbeeld waarin je dit" In de voorbereiding en de beoordeling van het portfolio kiest de assessor een paar essentiële onderwerpen uit die hij gericht wil bevragen en benoemt deze aan het begin van het CGI. Bij deze onderwerpen wordt de START-methodiek toegepast.

- Startvraag: Beschrijf eens een **S**ituatie meegemaakt waarin je ...?
- Wat was de opdracht / **T**aak / wat werd ervan jou verwacht?
- Wat heb je precies gedaan / wat is er gebeurd? (**A**ctie)
- Wat was het effect / hoe is het afgelopen? (**R**esultaat)
- Hoe zou je het bij een andere klas of ...? (**T**ransfer)

In een CGI van drie kwartier à zestig minuten past de assessor hooguit drie keer een START toe. De kwaliteit van de START-methodiek wordt bepaald door het feit dat de assessor grip krijgt en zicht heeft op het actuele niveau van de student. Het gaat er dan ook niet alleen om dat alle letters worden aangestipt en uitgevraagd maar dat de diepte van het kennis- en vaardigheidsniveau gevonden wordt. Een gesprekskern kan al snel een minuut of 10 à 15 duren.

Tijdens de START past de assessor **L**uisteren, **S**amenvatten en **D**oorvragen (*LSD*) toe. Dit structureert het gesprek voor alle betrokkenen en geeft de student een beeld of de assessor het juist gehoord heeft.

Daarnaast heeft de assessor 'TURBO-vragen' in zijn arsenaal zitten. TURBO-vragen past de assessor onder andere toe wanneer er nog weinig tijd over is en hij toch nog een aantal zaken wil weten of als hij via de "Stel dat ... vraag" de verdieping in de beschreven situatie zoekt.

Er zijn vier typen turbovragen:

1. Vergelijking: Als je jezelf vergelijkt met je collega's, wie vind je dan het beste in het analyseren van incidenten?
2. Rangorde: Op welke plaats zou je zelf staan als je een rangorde moet aanbrengen m.b.t.
3. Mening van anderen: Wat vinden je collega's van jouw manier van samenwerken?
4. Verdieping van de Situatie: Stel nou eens dat je wel alle vrijheid had gehad om volgens de Dalton principes de les in te vullen Hoe zou je dan te werk gaan?

Dit laatste type "Stel dat ... vraag" is geen oproep tot vrije speculatie wat wel eens gedacht wordt. Het is bedoeld om tot een nieuwe verdieping in een situatie te komen doordat de nieuwe situatie scherp is aangegeven. De assessor dient via strakke regie en focus voering de student in de nieuwe situatie te houden en door te vragen totdat helder wordt waar de student zich bevindt met betrekking tot het actuele denk- en handelingsniveau.

Een voorbeeld hiervan is een student die moeite heeft om zelf met een concrete situatie te komen en in de globale beschrijvingen blijft hangen. De assessor geeft dan een situatieomschrijving en houdt vervolgens de student 'vast' in de beschreven situatie (focus en regie houden). Een andere reden om de "Stel dat... vraag" toe te passen is wanneer dat de student voortdurend met te eenvoudige situaties als voorbeeld komt.

Als een gesprekskern is uitgevraagd, geeft de assessor een korte concluderende samenvatting en checkt of hij het juist verwoordt. Hiermee benoemt de assessor aan de student aan wat hij precies gehoord heeft en dat dit onderdeel van het CGI wordt afgesloten. Het uitvoeren van het CGI gebeurt als het goed is binnen het speelveld van de competentiebeoordeling. De assessor houdt de focus op de gespreklijn vast en heeft de regie in handen om de ervaringen, werkwijze en bewijzen die de student inbrengt te relateren aan de relevante competentiecriteria en het bijbehorende niveau.

Het hierboven beschreven beoordelingstraject van assessoren met het bijbehorende assessorenprofiel maakt onderdeel uit van het in juli 2012 ontvangen Hobéon Keurmerk voor Kwaliteit voor assessmentcentra⁴ van het EVC Centrum HvA.

⁴ Certificering HvA EVC Centrum uitgevoerd door Hobéon Certificering BV

Assessorenprofiel

Op basis van het bovenstaande is het profiel opgesteld voor assessoren (EVC Centrum, 2009) Het assessorenprofiel wordt bij alle assessorengroepen (ook andere HvA domeinen) als beoordelingsstandaard gebruikt.

Profiel Assessor⁵

Kerntaken assessor

beoordeelt of een portfolio van een kandidaat voldoet om het assessmentgesprek in te kunnen gaan; voert een criteriumgericht interview (= assessmentgesprek) n.a.v. het portfolio om het competentieniveau te bepalen; stelt het competentieniveau van de kandidaat vast (en deelt dit mee); geeft zo nodig ontwikkelingsgerichte feedback; legt bevindingen en eindoordeel vast op daarvoor bestemde formulieren t.b.v. de eindrapportage.

Basiskwalificaties assessor⁶

beschikt over relevante en brede expertise en praktijkervaring op het domein van de opleiding; is werkzaam in de opleiding (interne assessor) of in een relevante beroepspraktijk (externe assessor); heeft ruime ervaring met het begeleiden van hbo-stagiairs of startende collega's; heeft een hbo-diploma of hoger behaald; is klantgericht en oprecht geïnteresseerd in de professionele ontwikkeling van kandidaten; werkt efficiënt, accuraat en komt afspraken na; communiceert op toegankelijke en overtuigende wijze, zowel mondeling als schriftelijk; is leergericht: is bereid te investeren in zichzelf.

Kwaliteiten assessor⁷

De assessor is in staat om inhoudelijk en op professionele wijze te beoordelen of de kandidaat voldoet aan de vereiste competenties. Omdat de ontwikkeling en beoordeling van competenties aan de hand van zeer diverse praktijksituaties kan plaatsvinden, en omdat in de praktijk meerdere wegen naar Rome leiden, is het vooral van belang dat een assessor blijk geeft 'de wereld achter de indicatoren' te kennen. D.w.z. dat hij/zij competenties kan relateren aan uiteenlopende praktijksituaties en gedrag.

Gedragsindicatoren:

- *brengt structuur aan in het assessmentgesprek;*
 - *creëert een sfeer waarin de kandidaat volledig tot zijn recht kan komen;*
 - *past verschillende vraag- en gesprekstechnieken op het juiste moment toe om:*
 - *het competentieniveau van de kandidaat doelgericht te achterhalen;*
 - *regie te houden in gesprek;*
 - *relateert ervaringen, werkwijze en bewijzen die de kandidaat inbrengt, aan relevante competentiecriteria;*
 - *herkent (niet in het zelfevaluatie model genoemde) producten als bruikbaar bewijs;*
 - *komt tot een onderbouwd eindoordeel en brengt dit op overtuigende en constructieve wijze over;*
 - *geeft ontwikkelingsgerichte feedback;*
 - *legt het oordeel op toegankelijke wijze vast in een rapportage;*
 - *gaat adequaat en klantgericht om met eventuele bezwaren tegen het eindoordeel.*
-

⁵ Profiel is opgesteld door het HvA EVC Centrum en als goed beoordeeld door Inspectie en Hobéon.

⁶ De basiskwalificaties zijn opgesteld ten behoeve van de selectie van assessoren.

⁷ Ontwikkeling van de genoemde kwaliteiten staat centraal in de assessorentraining. Bij eventuele certificering van assessoren worden de genoemde gedragscriteria als beoordelingscriteria gehanteerd.

2.3 *Expertiseontwikkeling criteriumgericht beoordelen*

Alle hier bovenstaande toetsvormen dienen zo objectief mogelijk door beoordelaars beoordeeld te worden. Dit vereist bekwaame beoordelaars die op een uniforme wijze de beoordelingsstandaard toepassen en op een transparante manier de beoordeling verantwoorden en beargumenteren. Die competentie bezitten opleiders niet zondermeer.

Dat roept de vraag op hoe de assessor zich van een beginnende tot een expert assessor ontwikkelt. Dat brengt ons bij onderzoek naar expertiseontwikkeling. In Nederland is er het onderzoek van Adriaan D. de Groot (1946) die eind jaren dertig van de vorige eeuw zijn "hardop-denk-onderzoek" deed bij schakers. De Groot deed zelf mee aan de schaakolympiade in Buenos Aires in 1939 en was vanuit die situatie in de gelegenheid om zeer gerenommeerde schakers bij zijn onderzoek te betrekken. Hij had dan ook toegang tot proefpersonen van zeer uiteenlopend expertiseniveau. Wat De Groot ontdekte was dat schaakexperts niet dieper nadenken dan beginners, maar dat ze beter nadenken over betere zetten. Tevens deed hij een geheugenexperiment waaruit bleek dat de schaakexperts veel meer schaakstellingen onthouden dan beginners. In later onderzoek door Chase en Simon (1973) bleek dat wanneer je schaakexperts random posities laat zien dit voordeel grotendeels verdwijnt. Uit deze onderzoeksresultaten werd de conclusie getrokken dat de ontwikkeling van kennis, die het resultaat is van 10.000-den uren (Ericsson 2006) schaken, discussiëren, analyseren en studeren, de kern is van inzicht en expertise.

Het onderzoek naar expertise van schakers is niet zondermeer te transfereren naar de professionele expertiseontwikkeling van bijvoorbeeld beoordelaars binnen het hbo.

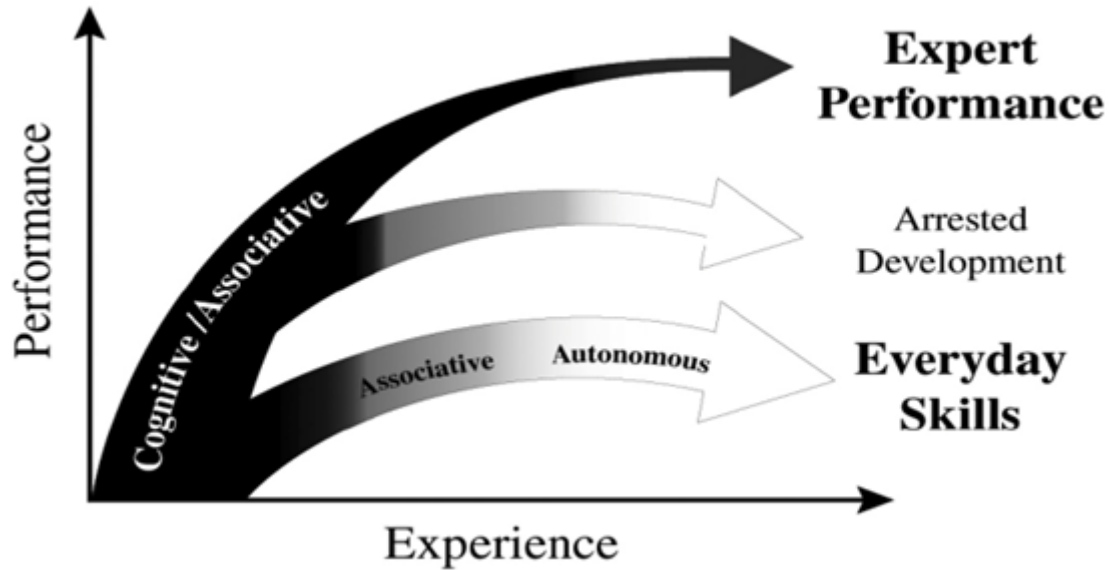
Professionele expertiseontwikkeling begint meestal in een formele opleiding. Deze opleiding wordt gevolgd door een transitie van school naar werk.

Volgens Mieg (2009) zijn er twee aspecten te onderscheiden aan expertiseontwikkeling, namelijk 'excellentie' en 'professionalisme'. Excellentie slaat op de kern van het beroep: Een leraar die verschillende leerstijlen weet aan te spreken door te differentiëren in werkvormen waardoor er betekenisvol geleerd wordt door de leerlingen. Professionalisme richt zich op de deelname aan verschillende professionaliseringsactiviteiten, lid zijn van professionele verenigingen, het lezen van vakbladen, het bezoeken van conferenties etc. Deze activiteiten hebben allemaal tot doel om enerzijds het beroep te beschermen en anderzijds de kwaliteit van de uitoefening van het vak te garanderen en verder te ontwikkelen. Beide aspecten kunnen samengaan in één persoon maar dat hoeft niet.

Experts worden geacht ergens heel goed in te zijn. Een belangrijke hedendaagse onderzoeker naar expertiseontwikkeling is K. Anders Ericsson. Zijn uitgangspunt is dat je alleen maar van expertise in een bepaald domein kunt spreken, wanneer er taken zijn aan te wijzen waarin experts het consistent en consequent beter doen dan gevorderden en/of beginners.

Ericsson's onderzoek bevestigt de 10.000 uur-regel, maar met een kanttekening: 10.000 uur pure ervaring is niet genoeg. In het artikel "The making of an Expert" laat Ericsson (2006) zien dat expertise het product is van jaren van 'deliberate practice'. Deliberate practice is het doelgericht uitvoeren van activiteiten, die gericht zijn op het verbeteren van de (werk)prestatie. Kenmerkend voor deliberate practice is het lange tijd oefenen van een taak, waarbij men zich concentreert op de belangrijke aspecten van de taak en daar feedback op krijgt. Het herhaald uitvoeren van de taak na feedback, die in het onderwijs plaats kan vinden in de vorm van reacties van leerlingen, toetsgegevens, enzovoorts, leidt stap voor stap tot verfijning van de prestatie.

Onderstaande figuur laat zien wat er gebeurt wanneer je bewust en gericht gaat oefenen, nadat de feedback op de vaardigheid is ontvangen (zwarte pijl). Hierdoor wordt het tevens een cognitieve activiteit.



Figuur 1: De relatie tussen performance en experience. (Uit: Ericsson, 2006)

Expertiseontwikkeling van de assessor ontstaat niet alleen door formele training, maar juist ook door leerervaringen in de toetsituaties, collegiale feedback van mede-assessor(en), het gezamenlijk zoeken naar oplossingen in lastige beoordelingssituaties en het ontvangen van gerichte feedback van experts.

Het bewust inoefenen van vaardigheden heeft zowel met leren als ontwikkeling te maken. Leren is gericht op het systematisch verwerven van kennis en vaardigheden in de vorm van (praktijk)regels en op het automatiseren van de vaardigheden door herhaaldelijk toepassen. Het verwerven van kennis en vaardigheden kan bijdragen aan cognitieve ontwikkeling. Bij cognitieve ontwikkeling is er sprake van een gefaseerde, kwalitatieve omslag in het denken en het handelingsrepertoire. Je bent er bewust mee bezig.

Het ontwikkelen van expertise in een specifieke vaardigheid die gepaard gaat met collegiale feedback. Het zoeken naar oplossingen voor problemen in de werksituatie kan leiden tot cognitieve ontwikkeling en daarmee tot expertiseontwikkeling, wat fasegewijs plaatsvindt. Het verschil in probleemoplossing tussen novices en experts toont het kwalitatieve verschil in denken en handelen. Bij novices is het leren gericht op kennisverwerving, kennisintegratie, kennisvalidatie en het controleren van de juistheid van kennis. Dit doet een groot beroep op het werkgeheugen en geeft een hoge cognitieve belasting. Een expert is in staat om complexe problemen op te lossen door uitgebreide domeinspecifieke kennis van zijn vakgebied, geautomatiseerde routines en het herkennen van patronen uit eerdere ervaringen. Geautomatiseerde routines verlagen het beroep wat op cognitieve capaciteiten gedaan moet worden, waardoor er meer aandacht is voor de specifieke kenmerken van een complex probleem (Berliner, 2004; Boshuizen, 2003).

In de hier boven beschreven theorieën ligt de legitimatie van de assessorentraining, het certificeringstraject en de uiteindelijke certificering.

3. Opzet en methodische verantwoording onderzoek

3.1 Type onderzoek

Om de onderzoeksvragen te beantwoorden is gekozen voor praktijkgericht onderzoek. Praktijkgericht onderzoek wordt hierbij gedefinieerd als:

“Praktijkgericht onderzoek is het onderzoek aan hogescholen dat geworteld is in de beroepspraktijk en dat bijdraagt aan de verbetering en innovatie van die beroepspraktijk. Dit vindt plaats door het genereren van kennis en inzichten, maar ook door het leveren van toepasbare producten en ontwerpen en concrete oplossingen voor praktijkproblemen. Daarbij is het onderzoek doorgaans multi- en of transdisciplinair van aard en ingebed in een scala van interne en externe organisatorische verbanden. Dit met behoud van de wetenschappelijke betrouwbaarheid en validiteit van het onderzoek zelf.”⁸

Vertaald naar het onderzoek “Professionalisering van de assessor” betekent dit dat dit onderzoek dienst bij te dragen aan de verbetering van de kwaliteit van de assessor. Het onderzoek maakt transparant dat er verschillende niveaus in de kwaliteit van het criteriumgericht beoordelen bij assessoren te onderscheiden zijn. Aan de hand daarvan is het minimum aan beoordelingsvaardigheden dat aan opleiders verwacht mag worden te bepalen en ook de gewenste groei tot gecertificeerd assessor (expert assessor).

Via publicatie en presentatie van de onderzoeksresultaten worden de kennis en actuele inzichten met betrekking tot de kwaliteit van assessoren verspreid.

Het onderzoek levert een kwaliteitsrubriek van interventies op waardoor het mogelijk wordt structureel vorm te geven aan het professionaliseringsproces van opleiders tot expert assessor op het gebied van criteriumgericht beoordelen.

In de onderstaande paragraaf wordt de opzet van het onderzoek en de wijze waarop de wetenschappelijke betrouwbaarheid en validiteit in dit onderzoek zijn gewaarborgd beschreven.

3.2 Opzet van het onderzoek

Om de onderzoeksvragen te kunnen beantwoorden is het noodzakelijk om de assessoren te observeren bij het afnemen van assessments. De onderzoeker heeft daarom gekozen voor observaties van ongetrainde, getrainde en expert assessoren. Dit was mogelijk doordat de onderzoeker vanuit haar assessorentrainingspraktijk direct toegang heeft tot het observeren van alle drie de onderzoeksgroepen ongetraind – ervaren - expert. De observaties waren deels ‘live’ en deels observaties van op DVD opgenomen assessments.

In dit onderzoek wordt aangesloten bij de gebruikte methodieken en instrumenten uit de assessorentrainingen, waar beoordelingen een vast onderdeel vormen. In de assessorentraining wordt gebruikt gemaakt van observatiemethodieken (WAKKER en het gedragsobservatieformulier, zie tekst hieronder) om zo objectief mogelijk te beoordelen. Daartoe zijn de zes onderdelen uit het assessorenprofiel in een gedragsobservatie schema gezet. Het gedragsobservatie formulier in de beoordelingspraktijk van assessoren ontwikkeld, uitgetoetst, besproken en gelegitimeerd met peerbeoordelaars.

De onderzoeker maakt gebruik van hetzelfde instrument bij de geobserveerde assessoren van onderstaand formulier om daarna de stap te maken naar de rapportage . Het gedragsobservatie formulier toon ik hier , aangezien de live- en dvd gesprekken met behulp van dit formulier zijn geanalyseerd.

⁸ Definitie van Validatiecommissie Kwaliteitszorg Onderzoek (geïnstalleerd door de HBO raad voor de periode 2009-2015. De kerntaak van de VKO is het valideren van de kwaliteitszorgsystemen van hogescholen, voor zover de kwaliteitszorg betrekking heeft op het *onderzoek* van de hogeschool.

Indicatoren	Gezien	Kwaliteit	Blijkt uit
	1 of 0	+ 0 -	
1. Relatie opbouwen, Kennismaken			
Stelt zich voor			
Stelt kandidaat op z'n gemak (unfreezing)			
2. Structuur aanbrengen in het assessmentgesprek			
Geeft doel, agenda en procedure aan (Opening)			
Verzamelt de vereiste informatie (Midden)			
Zorgt afsluiting van het gesprek, oordeel, feedback en vervolgprocedure (Eind)			
Geeft vervolgprocedure aan			
3. Vraag- en gesprekstechnieken			
Maakt gebruik van START-methodiek			
Maakt gebruik van Turbovraagstelling (rangorde, vergelijking, stel dat etc.)			
Houdt regie en past LSD toe (Luisteren, Samenvatten, Doorvragen)			
Geeft ruimte voor vragen kandidaat			
4. Relatie bewijsstukken, beroepshandelingen en beoordelingscriteria			
Assessor weet wat hij wil weten: Zowel persoonlijk functioneren, handelingen als beoordelingscriteria doorlopen			
5. Beoordelen			
Formuleert de beoordelingen in termen van competentie-indicatoren			
Herkent producten als bruikbaar bewijs			
Beoordeelt in 3 stappen: observeert, registreert en beoordeelt			
Komt tot een onderbouwd eindoordeel			
6. Geven van Feedback			
Brengt feedback op overtuigende wijze en geeft concrete vb'den uit het gesprek			
Geeft ontwikkelingsgerichte feedback			
7. Omgaan met bezwaren			
Totaal			

Figuur 2: Scoreformulier Gedragsobservatie assessmentgesprek

Tijdens de observaties in het kader van dit onderzoek is door de observatoren gebruik gemaakt van de WAK⁹KER-methodiek (NOA, 2002). Deze methodiek ondersteunt de onderzoeker/observator in het zo objectief en feitelijk mogelijk beoordelen.

De letters staat voor:

- **W**aarnemen
- **A**antekeningen maken
- **K**lasseren
- **K**walificeren
- **E**valueren
- **R**apporteren

Door je als observator strikt te houden aan de WAKKER-volgorde wordt het eigen subjectieve oordeel uitgesteld en wordt de observator gedwongen om het concrete feitelijke gedrag te observeren en dit te registreren. Dit gebeurt via de aantekeningen die de observator maakt tijdens het observeren van het CGI "live" of van de DVD opnamen. Vervolgens worden de feitelijke observaties geklasseerd door de beschreven gedragingen te koppelen aan de gedragsindicatoren. Er wordt een "1" ingevuld als het gedrag geobserveerd is en een "0" wanneer niet. Daarna volgt de kwalificatie en krijgt het gedrag een oordeel "+", wanneer het overeenkomt met de beschrijving uit het assessorenprofiel. Een "0" wanneer er onderdelen missen, onderdeel onvolledig uitgevraagd en/of onderdeel niet diepgaand wordt uitgevraagd ten opzichte van de beschrijving assessorenprofiel. Een "-" wordt gecodeerd als het uitvragen niet volgens de methodiek uit het assessorenprofiel plaatsvindt en de assessor het geheel op zijn eigen wijze doet. Een voorbeeld van het laatste is bijvoorbeeld als de assessor aan het begeleiden is in plaats van aan het beoordelen.

Als het hele CGI geanalyseerd is, wordt de samenvatting van de observatie vastgelegd in de rapportage.

Dit alles met als doel om antwoord te kunnen geven op de drie gestelde onderzoeksvragen:

1. Is er verschil te observeren bij de 3 onderzoeksgroepen (ongetraind, getraind en expert) in het gebruik van de gedragsindicatoren die de assessoren bij het criteriumgericht beoordelen hanteren tijdens de afname van het CGI. Het onderzoek richt zich vooral op de eerste vijf gedragsindicatoren (zie noot 2).
2. Als blijkt dat er verschil is in het gebruik van de gedragsindicatoren zijn er dan verschillen te observeren en/of patronen waar te nemen in de interventies die de assessoren gebruiken?
3. Is er een kwaliteitsrubriek criteriumgericht beoordelen te beschrijven waarin de drie verschillende professionele niveaus (ongetraind, getraind en expert) qua interventies van criteriumgericht beoordelen zichtbaar zijn.

3.3 Respondenten

In het onderzoek worden drie onderzoeksgroepen onderscheiden. De eerste groep zijn de deelnemers in de basis assessorentraining. De opleiders komen vanuit de HvA, InHolland en de VU. Dit zijn alle ervaren opleiders vanuit lerarenopleidingen en/of docenten, interne begeleiders, directeuren en opleiders in de school, die gewend zijn om studenten, beginnende docenten en collega's te begeleiden. Ze zijn ervaren in het gebruik van verschillende gesprekstechnieken. Wat nieuw voor ze is, is de criteriumgerichte beoordelingsmethodiek die de gesprekskwaliteiten en mensenkennis die men heeft, inzet bij integratieve toetsing zoals een portfolio assessment. Ze worden voor dit onderzoek geobserveerd tijdens de afname van het CGI met een acteur (dagdeel 4 van de training).

De assessoren uit de tweede groep zijn alle assessoren, die hebben deelgenomen aan een basis - assessorentraining en die nu bezig zijn met het certificeringstraject. Het is een ervaren groep in die zin dat ze regelmatig worden ingezet als assessor binnen de eigen opleiding. De geobserveerde DVD's komen van verschillende educatieve faculteiten uit de Randstad (Rotterdam, Utrecht, VU, InHolland en de HvA). In dit traject oefenen ze en zijn ze gericht op de assessorenkwaliteiten uit het assessorenprofiel, waarop ze in het certificeringsassessment beoordeeld worden.

⁹ De C en K in WAKKER wil volgens de spellingsregels wel eens wisselen. Ik kies hier voor de K van het woord WAKKER.

De derde groep zijn de gecertificeerde assessoren, die positief beoordeeld zijn op de assessorkwaliteiten uit het assessorprofiel door beoordelaars van het EVC Centrum HvA of dat van ADEF.

De drie onderzoeksgroepen zijn van samenstelling alle heteroog waar het gaat om leeftijd en geslacht. In het onderzoek zijn 46 assessoren geobserveerd die in de basis assessorentraining zaten. 12 assessoren uit het certificeringstraject zijn geobserveerd aan de hand van DVD opnames dan wel tijdens de afname van een opleidingsassessment.

De 5 gecertificeerde assessoren zijn geobserveerd aan de hand van DVD opnames of tijdens de afname van een opleidings- en/of ZiB-assessment. Deze laatste groep is beperkt in aantal. Dit komt doordat er relatief nog weinig assessoren gecertificeerd zijn. In het gehele domein Onderwijs en Opvoeding van de HvA waren dit er in 2012 veertien in totaal.

3.4 *Dataverwerking*

De methode van onderzoek is kwalitatief. Gekozen is voor de observatie als waarnemingstechniek om via de directe waarneming (zien en horen) de gedragsinterventies die de assessor pleegt tijdens het criteriumgericht interview (CGI) te observeren en te analyseren. Dit gebeurt in de natuurlijke context van de assessmentafname. De onderzoeker gebruikt het assessorenprofiel als kijkkader om de CGI-gesprekken en DVD's te analyseren.

Om de betrouwbaarheid te verhogen zijn van de op DVD opgenomen CGI's drie willekeurig gekozen CGIs door drie andere gecertificeerde assessoren (waarvan één tevens beoordelaar is bij certificering van assessoren) geanalyseerd. De analyse van de drie gecertificeerden is vergeleken met de analyse van de onderzoeker (peer examination, LeComte en Goetz 1982). De 46 CGI's van beginnende assessoren zijn geobserveerd en gescoord tijdens de basistraining, waarbij naast de onderzoeker ook een acteur (gecertificeerde assessor) aanwezig was die mee observeerde. In beide gevallen bleek de interbeoordelaarsbetrouwbaarheid van de analyses hoog.

Het assessorenprofiel en de WAKKER –methodiek is samengebracht in het gedragsobservatie formulier. Het vormt hierdoor een betrouwbaar instrument voor gestructureerde dataverzameling tijdens het CGI.

Naast de hierboven geschetste peer examination is de validiteit van een aantal van de analyses nog gecontroleerd door middel van een member-check. Bij de member-check gaat het om de vraag in hoeverre de geobserveerde assessoren zich herkennen in de analyse. In het onderzoek krijgt dit vorm doordat de 12 assessoren uit de getrainde onderzoeksgroep, de samenvatting van de observatiescore ontvingen met de vraag of ze zichzelf hierin konden herkennen. Dit bleek in alle situaties het geval. Bij de terugkoppeling geven de geobserveerde assessoren aan dat ze zich in hoge mate herkennen in de mondelinge en/of schriftelijke feedback.

Op basis van de uitkomst van de observaties is een kwaliteitsrubriek criteriumgericht beoordelen opgesteld.

4. Resultaat kwaliteitsrubriek criteriumgericht beoordelen

4.1 Verwerking ruwe onderzoeksgegevens

Om te komen tot de kwaliteitsrubriek criteriumgericht beoordelen zijn de interventies van de verschillende CGI gesprekken geanalyseerd. De verwerking is in onderstaand schema weergegeven.

Indicatoren	Gezien	Kwaliteit ¹⁰	Ongetraind	Getraind	Gecertificeerd
	1 of 0	+ 0 -	N=46	N=12	N=5
1. Relatie opbouwen, Kennismaken					
Stelt zich voor			46 keer 1 ; 46 +	12 keer 1 ; 12 +	5 keer 1 ; 5 +
Stelt kandidaat op z'n gemak (unfreezing)			43 keer 1 ; 43 +	12 keer 1 ; 12 +	5 keer 1 ; 5 +
2. Structuur aanbrengen in het assessmentgesprek					
Geeft doel, agenda en procedure aan (Opening)			31 keer 1; 5 +, 26 0, 15 -	10 keer 1; 2 + ; 9 0	5 keer 1 ; 5 +
Verzamelt de vereiste informatie (Midden)			46 keer 1; 46 0/-	12 keer 1; 4 + ; 7 0 1 0/-	5 keer 1 ; 5 +
Zorgt afsluiting van het gesprek, oordeel, feedback en vervolgprocedure (Eind)			40 keer 1; 40 0/-	12 keer 1; 4 + ; 8 0	5 keer 1 ; 5 +
Geeft vervolgprocedure aan			22 keer 1; 18 +, 4 0	12 keer 1; 12 +	5 keer 1 ; 5 +
3. Vraag- en gesprekstechnieken					
Maakt gebruik van START-methodiek			14 keer 1; 14 0/-	10 keer 1; 4 + ; 6 0	5 keer 1 ; 5 +
Maakt gebruik van Turbovraagstelling (rangorde, vergelijking, stel dat etc.)			12 keer 1; 12 0	7 keer 1; 4 + ; 3 0	4 keer 1 ; 4 +
Houdt regie en past LSD toe (Luisteren, Samenvatten, Doorvragen)			13 keer 1; 2 +; 11 0	11 keer 1; 6+; 5 0	5 keer 1 ; 5 +
Geeft ruimte voor vragen kandidaat			42 keer 1; 42 +	12 keer 1; 12 +	5 keer 1; 5 +

¹⁰ + volledig/bijna volledig, voldoende ; 0 missen onderdelen, onderdeel onvolledig, onderdeel niet diepgaand ; - niet volgens methodiek; 0/- beide gelden

Indicatoren	Gezien 1 of 0	Kwaliteit + 0 -	Ongetraind N=46	Getraind N=12	Gecertificeerd N=5
4. Relatie bewijsstukken, beroepshandelingen en beoordelingscriteria					
Assessor weet wat hij wil weten: Zowel persoonlijk functioneren, handelingen als beoordelingscriteria doorlopen			26 keer 1; 26 keer 0/-	12 keer 1; 7 +; 5 keer 0	5 keer 1; 5 +
5. Beoordelen					
Formuleert de beoordelingen in termen van competentie-indicatoren			36 keer 1; 36 0	12 keer 1; 12 +	5 keer 1; 5 +
Herkent producten als bruikbaar bewijs				12 keer 1; 12 +	5 keer 1; 5 +
Beoordeelt in 3 stappen: observeert, registreert en beoordeelt				12 keer 1; 8 +; 4 keer 0	5 keer 1; 5 +
Komt tot een onderbouwd eindoordeel			36 keer 1; 36 keer 0/-	12 keer 1; 8 +; 4 keer 0	5 keer 1; 5 +
6. Geven van Feedback					
Brengt feedback op overtuigende wijze en geeft concrete voorbeelden uit het gesprek			46 keer 1; 46 keer 0/-	12 keer 1; 12 +	5 keer 1; 5 +
Geeft ontwikkelingsgerichte feedback			42 keer 1; 12 keer +	12 keer 1; 12 +	5 keer 1; 5 +
7. Omgaan met bezwaren					
Assessor doet dit klantgericht met gebruik van ADSIC Aanmoedigen Doorvragen Samenvatten Informereren Controleren			30 keer 1; 2 +; 28 0/-	1 keer 1; 1 0	

Figuur 3: Overzicht geobserveerde interventies

4.2 Kwaliteitsrubriek criteriumgericht beoordelen

Uit het overzicht van de geobserveerde interventies bij de drie verschillende groepen assessoren blijkt dat er duidelijk verschil is in het gebruik van interventies die de assessoren inzetten bij het CGI (onderzoeksvraag 1). Tevens is er een patroon zichtbaar waarbij er sprake is van een opbouw in de compleetheid en correctheid van de inzet van interventies van ongetraind assessor tot expert assessor (onderzoeksvraag 2).

Op basis hiervan is het mogelijk om te komen tot een kwaliteitsrubriek criteriumgericht beoordelen (onderzoeksvraag 3).

Om te komen tot de kwaliteitsrubriek criteriumgericht beoordelen zijn de onderzoeksdata geanalyseerd met gebruikmaking van het eerder genoemde assessorenprofiel, de WAKKER - methodiek en het gedragsobservatie formulier. De interventies zijn geclassificeerd aan de hand van de verschillende gedragsindicatoren van het assessorenprofiel en in een kwaliteitsrubriek gezet. Dit maakt de expertiseontwikkeling van assessoren zichtbaar van ongetraind naar expertassessor. Dit heeft geleid tot de kwaliteitsrubriek criteriumgericht beoordelen assessoren domein educatie.

De kwaliteitsrubriek criteriumgericht beoordelen geeft het handelingsgedrag weer van de interventies die assessoren doorlopen van ongetraind, getraind naar expert. De kwaliteitsrubriek maakt zichtbaar of en in welke mate de beoordelaar dan wel (werkplek) begeleider op een transparante wijze en zo objectief mogelijk tot zijn beoordeling komt. De gedragsindicatoren van het assessorenprofiel (EVC Centrum HvA) en de observaties uit het onderzoek zijn als beoordelingsstandaard toegepast in de kwaliteitsrubriek. De rubriek is zo neutraal mogelijk opgesteld vanuit het perspectief waar de assessor staat en wat hij doet.

Daarmee kan de kwaliteitsrubriek dus tweeledig gebruikt worden: als beoordelingsgericht instrument of als ontwikkelingsgericht instrument. Als *beoordelingsgericht instrument* biedt de kwaliteitsrubriek een normatief kader: De assessor op het expertniveau weet de juiste technieken en procedures op een eenduidige wijze toe te passen en de beoordeling te beargumenteren met verwijzing naar bewijsstukken en beroepshandelingen in relatie tot de beoordelingscriteria. Als dit op boven beschreven wijze gebeurt dan kan de student de genomen stappen volgen en ziet hij op welke manier de beoordelingscriteria door de beoordelaar worden toegepast. De student herkent dat er is recht gedaan en dat hij zo objectief mogelijk is beoordeeld. De student weet na de criteriumgerichte beoordeling waar hij staat (zijn actuele ontwikkelingsniveau) ten opzichte van de beoordelingscriteria (hbo- en competentieniveau) en heeft zicht gekregen op wat er nog ontwikkeld dient te worden voor een volgend niveau (hoofdfase bekwaam, leraar in opleiding (LIO)-/afstudeerbekwaam en startbekwaam).

De kwaliteitsrubriek kan tevens als *ontwikkelingsgericht instrument* gebruikt worden. Om deze ontwikkeling te structureren zijn reflectievragen toegevoegd en bewerkt die afkomstig zijn uit het profiel docent Basiskwalificatie Onderwijs (BKO) van de HvA. De assessor scoort zich op de gedragsindicatoren van de rubriek en stelt zich aan de hand van de reflectievragen assessoren criteriumgericht beoordelen een aantal reflectievragen.

Zoals bij elke niveau indeling kunnen assessoren zich versneld door de kwaliteitsrubriek bewegen richting expert assessor. Ook hier gelden EVC's en aanleg.

De kwaliteitsrubriek is ter validering voorgelegd aan drie assessorentrainers en twee beoordelaars van assessoren. Dit heeft ertoe geleid dat de kwaliteitsrubriek criteriumgericht beoordelen vijf verschillende versies heeft gekend en steeds verder is bijgeschaafd.

Kwaliteitsrubriek Criteriumgericht Beoordelen
Assessoren domein Educatie
(doorlopende leerlijn kwalitatief beoordelen)

Gedragingen Niveaus	Ongetraind	Getraind	Expert
Gedragsindicatoren			
1. Relatie opbouwen, Kennismaken	<p>De assessor is vriendelijk, stelt zichzelf voor. Stelt de student op zijn gemak.</p> <p>Resultaat: De student voelt zich op zijn gemak.</p>	<p>Stelt de student op zijn gemak Heeft een uitnodigende houding. Is alert op de aanwezigheid van eventuele weerstanden, angst of onzekerheid en is in staat om die op empathische wijze te benoemen.</p> <p>Resultaat: De student voelt zich op zijn gemak en is in staat een open gesprek te voeren met de assessoren.</p>	<p>Stelt de student op zijn gemak Doorbreekt, neutraliseert eventuele weerstanden of irritaties in het gesprek.</p> <p>Resultaat: Student voelt zich op zijn gemak en is in staat een open gesprek te voeren met de assessoren en de vereiste informatie te verstrekken.</p>
2. Gespreksstructuur	<p>1. Opening De assessor is vriendelijk maar onvolledig over de procedure van het CGI en het beoordelende karakter daarvan. De inhoudelijke bevindingen van de assessor over het portfolio worden niet met de student gedeeld. De bewijzen die de student nog moet aandragen om aan de criteria te voldoen worden niet besproken. Niet alle onderdelen van opening worden benoemd (voorstellen, doel gesprek, uitleg procedure, agenda, rolverdeling assessoren, ruimte voor vragen student).</p> <p>Resultaat: De student weet niet wat er wordt verwacht, heeft als het ware te maken met een 'verborgen agenda' van de assessor. De focus tijdens het CGI ligt blijft onduidelijk.</p> <p>2. Midden De overgang van opening naar het vraaggedeelte (middengedeelte) van het CGI wordt niet duidelijk gemarkeerd. De assessor luistert vriendelijk en reageert op de student vanuit zijn eigen referentiekader.</p>	<p>1. Opening De assessor legt het doel van het gesprek uit en bespreekt de procedure, de taak- en rolverdeling van de assessoren. Er is ruimte voor vragen van de student. De inhoudelijke bevindingen van de assessor over het portfolio worden met de student gedeeld. De bewijzen die de student nog moet aandragen om aan de criteria te voldoen worden niet besproken.</p> <p>Resultaat: De student kent de procedure van het CGI maar heeft inhoudelijk onvoldoende informatie om te begrijpen waar de assessor naar op zoek is. Het gewenste niveau en de criteria worden te globaal en te algemeen getoetst. Student weet niet precies waar de assessor naar op zoek is.</p> <p>2. Midden De assessor houdt zich aan de structuur van het gesprek en geeft aan wanneer het vraaggedeelte start (middengedeelte) van het CGI. De assessor is bekend met de</p>	<p>1. Opening Voorstellen, unfreezing/relaxation, doel gesprek uitleg procedure, agenda, rolverdeling assessoren, ruimte voor vragen student. De student/ student krijgt precies te horen wat de assessor nog te weten wil komen. Het "assessment-speelveld" is voor student en assessor transparant.</p> <p>Resultaat: De student kent de procedure van het CGI en weet exact waar de assessoren naar op zoek zijn.</p> <p>2. Midden Verzamelt de vereiste informatie m.b.t. de te beoordelen competenties. Benoemt telkens het onderwerp van gesprek, vraagt door en rondt af door concluderend samen te vatten.</p>

	<p>De assessor zit vaker in de begeleidende- dan in beoordelende rol. De assessor houdt niet de regie over het gesprek. De structuur lijkt eerder op een 'gewoon' gesprek waarin gesprekspartners beurtelings het initiatief nemen en volgen. Het gesprek blijft daardoor algemeen beschrijvend. De bewijzen die nodig zijn om tot beoordeling te komen, worden niet verzameld.</p> <p>Resultaat: De student komt niet tot reflectie over de aanpak van een beroepssituatie en de leeropbrengst daarvan.</p> <p>3. Eind De assessor heeft niet voldoende bewijzen kunnen verzamelen en beëindigt het gesprek met een subjectief getint oordeel.</p> <p>De vervolprocedure wordt niet of onvoldoende aangegeven.</p> <p>Resultaat: De student is wel of niet tevreden over de uitslag en gaat eventueel zelf op zoek naar manieren om bezwaar aan te tekenen.</p>	<p>gesprekstechnieken zoals START en LSD en past deze technieken geheel of gedeeltelijk toe om een indruk te krijgen van het de competentie- en hbo niveau van de student..</p> <p>Resultaat: De assessor gebruikt START en LSD technisch gezien wel maar focust nog niet of niet voldoende op verdieping en reflectie. De student krijgt onvoldoende informatie over zijn kunnen en de assessor heeft onvoldoende informatie om te beoordelen.</p> <p>3. Eind De assessor ontdekt vaak pas aan het eind welke informatie nog ontbreekt om tot een onderbouwd oordeel te komen. Het oordeel kan daardoor nog steeds gevoelsmatig en subjectief op de student overkomen.</p> <p>De assessor geeft de termijn aan waarbinnen de student de rapportage ontvangt en - indien van toepassing - hoe de student bezwaar kan aantekenen.. De vervolprocedure wordt aangegeven.</p> <p>Resultaat: Zowel student als assessoren kennen de procedure van het CGI. Student kan het oordeel van de assessor als subjectief ervaren.</p>	<p>Resultaat: De assessor heeft de technieken geïnternaliseerd en gebruikt het vloeiend. De regie en focus op wat men te weten wil komen is goed, waardoor er verdieping en reflectie ter plekke plaatsvindt.</p> <p>3. Eind Beëindigt het gesprek met onderbouwd oordeel , geeft termijn aan waarbinnen student rapportage ontvangt en - indien van toepassing - procedure van bezwaar aan.</p> <p>Resultaat: Zowel student als assessoren weten op elk moment 'waar' ze zich bevinden in het gesprek en welk onderwerp aan de orde is. De student reflecteert over de aanpak van een beroepssituatie en de leeropbrengst daarvan. Het is voor de student duidelijk welk niveau en welke criteria getoetst worden.</p>
--	--	--	---

	Ongetraind	Getraind	Expert
3. Gesprekstechniek (START, turbo, LSD, metacommuniceren, etc.)	<p>De assessor is onbekend of onbewust bekwaam in het hanteren van gesprekstechnieken zoals START, LSD die hij nodig heeft om de ontbrekende bewijzen op tafel te krijgen</p> <p>De S van Situatie komt soms aan bod, soms ook de A van Aanpak.</p> <p>Metacommunicatie wordt niet bewust ingezet als middel om tot beoordeling/verdieping te komen. Hetzelfde geldt voor Turbovragen.</p> <p>Resultaat: Student weet niet of hij op de goede weg is. Er komt geen reflectie op gang over het leerproces en er komt geen samenvatting als bewijs voor de beoordeling van de competentie.</p>	<p>De assessor luistert actief, vraagt door en vat samen (LSD).</p> <p>De assessor probeert d.m.v. vragen, samenvattingen of onderbrekingen de regie in het gesprek te houden.</p> <p>De assessor is bekend met methodieken als START en Turbo en past die regelmatig toe om door te vragen op het handelingsrepertoire van de student in een concrete beroepssituatie.</p> <p>De assessor oefent met de rol van examiner/beoordelaar maar valt nog vaak terug in de meer volgzaam rol van begeleider.</p> <p>Nog niet alle elementen van START komen systematisch en consequent aan bod. Met name naar de R van Reflectie en/of Resultaat en naar de T van Transfer worden overgeslagen.</p> <p>Resultaat: De student krijgt veel ruimte voor uitweidingen. De assessor onderbreekt de student zelden of niet en de tijd wordt voor een groot deel gevuld met het bespreken van zaken die al in het portfolio stonden of die irrelevant zijn voor de beoordeling.</p> <p>De student weet onvoldoende over zijn beheersingsniveau en over de kennis of vaardigheden, attitudes die bij het competentieniveau horen en die nog ontbreken.</p>	<p>Luistert actief</p> <p>Hanteert geëigende gesprekstechnieken zoals START om door te vragen naar concreet gedrag en naar denk- en beslisprocessen.</p> <p>Verkent de reikwijdte van competenties (transfer)</p> <p>Houdt regie in het gesprek (Luisteren-Samenvatten-Doorvragen)</p> <p>Past, indien nodig, de techniek van metacommuniceren toe op niveau van relatie en inhoud.</p> <p>Sluit aan bij student voor wat betreft taalgebruik en niveau van communiceren.</p> <p>Resultaat: Een gesprek met diepgang: de vereiste informatie voor competentiebeoordeling komt op tafel.</p>
4. Relatie bewijsstukken, beroepshandelingen en beoordelingscriteria	<p>De assessor is onbekend met de VRAAK criteria voor de beoordeling van de bewijzen. (authenticiteit, relevantie, actualiteit, variatie en kwantiteit).</p> <p>De assessor is onbekend met de beoordelingsdriehoek (portfolio/bewijzen – beroepshandelingen – criteria).</p>	<p>De assessor is bekend met de VRAAK criteria (authenticiteit, relevantie, actualiteit, variatie en kwantiteit) om bewijzen mee te beoordelen.</p> <p>De assessor is bekend met de beoordelingsdriehoek waarin bewijzen, beroepshandelingen en beoordelingscriteria structureel met elkaar verbonden worden.</p>	<p>Beoordeelt bewijzen op authenticiteit, relevantie, actualiteit, variatie en kwantiteit.</p> <p>Hanteert de beoordelingsdriehoek waarin bewijzen, beroepshandelingen en beoordelingscriteria structureel met elkaar verbonden worden.</p> <p>Herkent (niet in het portfolio) genoemde bewijzen als bruikbaar bewijs.</p>

	<p>Resultaat: Competentiebeoordeling is subjectief.</p>	<p>Resultaat: De assessor hanteert VRAAK en de beoordelingsdriehoek als instrumenten om de bewijzen van de student op bruikbaarheid te toetsen.</p>	<p>Relateert bewijzen aan de juiste competenties Legt indien nodig, verbanden tussen competenties</p> <p>Resultaat: de vereiste informatie voor onderbouwde competentiebeoordeling komt op tafel.</p>
	Ongetraind	Getraind	Expert
5. Beoordelingstechniek	<p>Omdat de assessor onbekend is met de beoordelingsinstrumenten is het niet mogelijk om tot een transparante beoordeling te komen.</p> <p>Resultaat: De beoordeling is subjectief en kan sterk afwijken van het oordeel van een volgende assessor.</p>	<p>De assessor benoemt het handelen van de student in termen van de competentie-indicatoren van de matrix of Kijkwijzer(PABO). De assessor beoordeelt in 3 stappen: observeren, registreren en beoordelen. De assessor onderbouwt de beoordeling met voorbeelden uit het portfolio, praktijkopdracht en CGI.</p> <p>De assessor komt tot een eindbeoordeling en stemt dit af met de co-assessor. De assessor informeert de student op toegankelijke wijze over het oordeel en onderbouwing.</p> <p>Resultaat: Eenduidige en transparante werkwijze van assessoren onderling.</p>	<p>Past de geëigende beoordelingstechnieken gestructureerd toe en hanteert hierbij de beoordelingsdriehoek. Onderbouwt per competentie het oordeel met voorbeelden van competent gedrag uit portfolio, praktijkopdracht en gesprek gerelateerd aan de beoordelingsstandaard. Komt tot een oordeel en stemt dit af met de co-assessor voor een intersubjectief oordeel. Bij verschil van mening over de beoordeling wordt met gebruik van feiten uit portfolio, praktijkopdracht en gesprek consensus bereikt. Informeert student op toegankelijke wijze over het oordeel en geeft onderbouwing.</p> <p>Resultaat: Eenduidige en transparante werkwijze waarop het oordeel tot stand komt met als eindresultaat een intersubjectief oordeel op de beoordelingsstandaard.</p>
6. Feedback geven	<p>De assessor in de rol van begeleider/coach kent de feedbackregels en kan die goed toepassen.</p> <p>Resultaat: De student heeft feedback ontvangen en het is in grote lijnen duidelijk welke competenties en bijbehorend hbo-niveau zijn aangetoond, welke niet of gedeeltelijk. Hoe het oordeel tot stand is gekomen is niet transparant waardoor gemakkelijk discussie ontstaat over het oordeel.</p>	<p>De assessor kent de feedbackregels en kan die goed toepassen.</p> <p>Geeft feedback op maat gekoppeld aan concrete voorbeelden uit het portfolio of CGI..</p> <p>Resultaat: De student heeft feedback ontvangen en het is duidelijk welke competenties en bijbehorend hbo-niveau zijn aangetoond, welke niet of gedeeltelijk.</p>	<p>Past het model en de regels van feedback geven toe.</p> <p>Resultaat: Student heeft feedback ontvangen en het is duidelijk welke competenties en bijbehorend hbo-niveau zijn aangetoond, welke niet of gedeeltelijk. Tevens is (indien nodig) helder gemaakt waaraan hij nog moet werken om het vereiste niveau te behalen.</p>

	Ongetraind	Getraind	Expert
7. Vastleggen oordeel	<p>De assessor legt bevindingen vast, eventueel na afstemming met de co-assessor op het daarvoor bestemde rapportageformat. Dit gebeurt naar beste weten en volgens eigen inzicht van de assessor. De assessor is nog onbekend met de uniforme regels die voor rapportages gelden.</p> <p>Resultaat: De kwaliteit van de rapportage is afhankelijk van de kwaliteit van schriftelijke verslaglegging van de assessor.</p>	<p>Legt bevindingen, na afstemming met de co-assessor, toegankelijk vast op het daarvoor bestemde rapportageformat. Beschrijft de wijze waarop het oordeel tot stand is gekomen.</p> <p>Resultaat: In de rapportage is op toegankelijke wijze vastgelegd op welke manier het oordeel tot stand is gekomen. Het oordeel is onderbouwd met bevindingen uit portfolio, praktijkopdracht en gesprek en geeft een totaaloverzicht van de beoordeelde competenties en hbo-niveau.</p>	<p>Legt bevindingen, na afstemming met de co-assessor, toegankelijk vast op het daarvoor bestemde rapportageformat. Beschrijft de wijze waarop het oordeel tot stand is gekomen. Onderbouwt dit oordeel met bevindingen uit portfolio, praktijkopdracht en gesprek (toelichting op bewijs, gedrag, denk- en beslisprocessen, reflectie en transfer).</p> <p>Resultaat: In de rapportage is op toegankelijke wijze vastgelegd: het intersubjectieve oordeel is per competentie gerelateerd aan het hbo-niveau en de beoordelingsstandaard. Dit oordeel is onderbouwd met bevindingen uit portfolio, praktijkopdracht en gesprek en geeft een totaaloverzicht van de beoordeelde competenties en hbo-niveau.</p>
8. Omgaan met bezwaren (indien van toepassing)	<p>De assessor hoort bezwaren aan.</p> <p>Resultaat: De student is teleurgesteld over de uitslag maar weet dat er bezwaar aangetekend kan worden bij de Examencommissie.</p>	<p>Assessor hoort bezwaren aan en verzamelt de verkregen informatie.</p> <p>Assessor geeft aan waar student de procedure van bezwaar kan vinden.</p> <p>Resultaat: Er is naar de student geluisterd en hij heeft de informatie verkregen waar hij de procedure van bezwaar kan vinden en welke stappen te nemen.</p>	<p>Assessor hoort bezwaren aan en doet dit studentgericht met gebruik van ADSIC. (Aanmoedigen, Doorvragen, Samenvatten, Informeren, Controleren).</p> <p>Resultaat: De student voelt zich gehoord in zijn bezwaar. Er is naar hem geluisterd en waar mogelijk informatie gegeven om onduidelijkheden, misverstanden weg te nemen. Hij weet waar hij de procedure van bezwaar kan vinden en welke stappen te nemen.</p>
	Ongetraind	Getraind	Expert

Figuur 4: Kwaliteitsrubriek criteriumgericht beoordelen

4.3 Kwaliteitsrubriek als ontwikkelingsgericht instrument

Het zou jammer zijn als de kwaliteitsrubriek alleen als beoordelingsinstrument gebruikt zou worden en als een soort afvinklijst gaat dienen. Om dit te voorkomen zijn reflectievragen¹¹ toegevoegd, zodat assessoren zelf en binnen collegiale consultatieachtige settings gestructureerd de rubriek toepassen als ontwikkelingsgericht instrument. Hiervoor is gebruik gemaakt van de reflectievragen die zijn toegevoegd aan het profiel Basiskwalificatie Onderwijs (BKO) van de HvA. Uit het profiel BKO zijn die reflectievragen uitgezocht en aangepast die passend zijn op de vaardigheid criteriumgericht beoordelen en begeleiden.

Reflectievragen assessoren Criteriumgericht Beoordelen (ontleend aan profiel docent Basiskwalificatie Onderwijs (BKO) – HvA en toepasbaar gemaakt voor het stellen van gerichte ontwikkelingsvragen passend bij de kwaliteitsrubriek criteriumgericht beoordelen).

Ongetraind	Getraind	Expert
<ul style="list-style-type: none"> • Beheers ik de methodieken en vaardigheden om criteriumgericht te beoordelen? • Zorg ik voor voldoende veiligheid en eerlijkheid tijdens de beoordeling van de student? • Kijk ik mee hoe collega-beoordelaars hun criteriumgericht beoordelen optimaliseren? • Sta ik open voor en zoek ik inter-collegiale ondersteuning bij criteriumgericht beoordelen? • Onderzoek ik hoe ik de specifieke kennis en ervaring die ik meeneem kan benutten in mijn beoordelingen? 	<ul style="list-style-type: none"> • Voer ik de toetsvorm criteriumgericht beoordelen uit zoals afgesproken? • Gebruik ik de student als bron voor feedback bij het criteriumgericht beoordelen? • Bouw ik expertise op die bruikbaar is in mijn rol als beoordelaar? • Boor ik relevante toetsbronnen aan, die bruikbaar zijn voor mijn criteriumgericht beoordelen? • Zet ik eigen praktijkervaring, actuele kennis over het werkveld en contacten in het werkveld in bij het criteriumgericht beoordelen? • Wissel ik uit, zoek ik intercollegiale consultatie bij integratieve toetsvormen? • Ben ik me bewust van mijn eigen subjectiviteit? Lukt het me om zo open en doelgericht mogelijk criteriumgericht te beoordelen? 	<ul style="list-style-type: none"> • Voer ik in meerdere contexten toetsing vloeiend uit en stem ik de interventies af op omstandigheden? • Ben ik bezig met de vraag hoe studenten leren en zorg ik tijdens het criteriumgericht beoordelen voor een toetsomgeving waarbij studenten hun ontwikkeling tonen en verantwoorden? • Heb ik oog voor diversiteit (o.a. sekse, socialisatie, etniciteit, geardeerdheid en leeftijd) tijdens het criteriumgericht beoordelen? • Optimaliseer ik samen met collega's integratieve toetsvormen in aansluiting op beoogde leerresultaten, de beginsituatie van studenten en de actualiteit in de beroepspraktijk? • Investeer ik in eigen toetservaring, actuele kennis en contacten in het werkveld om deze zodanig in het onderwijs te integreren dat dit waarde toevoegt aan het beoordelen van studenten? • Verfijn ik de inzet van mijn toetsexpertise door te reflecteren op mijn ervaringen en gebruik te maken van inzichten uit assessorenbijeenkomsten en onderzoek? • Geef ik zelf vorm aan intercollegiale consultatie op het gebied van criteriumgericht beoordelen?

Figuur 5: Reflectievragen assessoren Criteriumgericht Beoordelen

¹¹ Afkomstig en bewerkt uit het profiel docent Basiskwalificatie Onderwijs (BKO) van de HvA.

5. Conclusies en aanbevelingen

5.1 Antwoord op de onderzoeksvragen

De onderzoeksvragen van dit praktijkonderzoek naar de professionalisering van de assessor herhaal ik hieronder en ik zal direct antwoord geven vanuit het gedane onderzoek.

1. Is er verschil te observeren bij de 3 onderzoeksgroepen (ongetraind, getraind en expert) in het gebruik van de gedragsindicatoren die de assessoren bij het criteriumgericht beoordelen hanteren tijdens de afname van het CGI. Het onderzoek richt zich vooral op de eerste vijf gedragsindicatoren (zie noot 1).

Ja, er is duidelijk verschil te observeren tussen de 3 onderzoeksgroepen in het gebruik van de gedragsindicatoren bij het criteriumgericht beoordelen tijdens de CGI afname (zie 4.1).

2. Als blijkt dat er verschil is in het gebruik van de gedragsindicatoren zijn er dan verschillen te observeren en/of patronen waar te nemen in de interventies die de assessoren gebruiken?

Ja, er zijn verschillen te observeren en patronen waar te nemen in de interventies die de assessoren tijdens het CGI gebruiken (zie 4.1).

3. Is er een kwaliteitsrubriek criteriumgericht beoordelen te beschrijven waarin de drie verschillende professionele niveaus (ongetraind, getraind en expert) qua interventies van criteriumgericht beoordelen zichtbaar zijn.

Ja, er is op basis van de gevonden patronen een kwaliteitsrubriek criteriumgericht beoordelen te beschrijven die zich richt op de assessoren uit het domein Educatie¹², waarin de drie verschillende professionele niveaus (ongetraind, getraind en expert) qua interventies van criteriumgericht beoordelen zichtbaar zijn (zie 4.2).

5.2 Drie niveau's in assessoren

Op basis van de observaties die in het kader van dit onderzoek zijn uitgevoerd, kunnen drie prototypische assessorenprofielen beschreven worden, die onderling verschillen in de mate waarin ze systematisch gebruik maken van gesprekstechnieken, in staat zijn om tot objectieve en transparante oordelen te komen, zich bewust zijn van de kwaliteit van hun handelen en gericht zijn op het verbeteren daarvan.

Het gaat hier om prototypische beschrijvingen: individuele assessoren kunnen ook zonder formele training tot de categorie expert assessor horen, en assessoren kunnen kenmerken uit verschillende prototypen hebben.

- Een **ongetrainde assessor** hanteert bepaalde, eenvoudige gesprekstechnieken die hij ooit geleerd heeft in zijn opleiding tot opleider. Vanuit die basiskennis weet hij onvoldoende de focus van het criteriumgericht interview (CGI) duidelijk te maken aan de student. Hij is voorts onvoldoende bekend in het hanteren van gesprekstechnieken, zoals START, LSD en TURBO, die ingezet worden om ontbrekende bewijzen helder door de student te laten tonen. Hierdoor is de beoordeling niet transparant voor anderen. De student kan de beoordeling als subjectief ervaren; hij is afhankelijk van niet expliciet gemaakte criteria die de beoordelaar hanteert.
Concluderend: Beoordeling is niet transparant, sterk subjectief getint en afhankelijk van de assessor. Er is geen uniformiteit in procedures en technieken zichtbaar.
- De **getrainde assessor** kent en hanteert de juiste procedures en gewenste technieken om te komen tot criteriumgerichte beoordeling. De assessor gebruikt de technieken (START, LSD en TURBO) echter op een technische wijze en/of focust zich in beperkte mate op verdieping en reflectie. De assessor "vinkt af" en vraagt niet verder. Hierdoor krijgt niet alleen de assessor maar ook de student onvoldoende informatie. De student heeft niet het gevoel "alles" te hebben getoond van wat hij weet en kan. In feite heeft de assessor onvoldoende informatie tot zijn beschikking om tot kwalitatieve en transparante beoordeling te komen. De student

¹² Interessant zou zijn voor vervolgonderzoek of de kwaliteitsrubriek herkenbaar en bruikbaar is ook voor assessoren/beoordelaars uit andere domeinen.

ervaart hierdoor het oordeel van de assessor nog te vaak als subjectief, doordat het niet transparant is. Het blijft onduidelijk hoe bewijsvoering van de student gerelateerd is aan beoordelingscriteria en bijbehorend hbo- en competentieniveau. De assessoren onderling komen wel tot een eenduidige en transparante manier van werken en beoordelen, maar missen een beslissende slag in hun eindoordeel.

Concluderend: De juiste procedures en gewenste technieken zijn in beeld maar worden nog te technisch toegepast. Tussen de assessoren is de uniformiteit op procedures en technieken groot. De student kan de beoordeling nog altijd als “subjectief” ervaren, doordat hij in zijn ogen onvoldoende heeft kunnen tonen wat volgens de complexe context wordt bedoeld. De assessoren zijn altijd verantwoordelijk dat er voldoende diepgang is bereikt en op dit punt faalt een getrainde assessor nog te vaak.

- De **expert assessor** weet de juiste procedures en technieken toe te passen: het doel om te komen tot een transparante beoordeling die zo maximaal mogelijk objectief is wordt bereikt. De student en de assessoren weten op elk moment van het CGI ‘waar’ ze zich bevinden en welk onderwerp aan de orde is. De assessor let voortdurend op de juiste focus en houdt de regie op dynamische (natuurlijke) manier in handen. Het gesprek heeft de gewenste diepgang waardoor informatie betreffende de vereiste competenties wordt verkregen. De beoordeling geschiedt op concrete feiten en aantekeningen. De student heeft aan het einde van het assessment het gevoel “alles” te hebben kunnen zeggen en kan in het nagesprek “begrijpen” waarom bepaalde bewijzen die hij heeft gegeven, wel of niet “als voldoende” worden gehonoreerd. De wijze van beoordelen gebeurt eenduidig en transparant, zodat het eindresultaat een intersubjectief oordeel is, passend bij de gehanteerde beoordelingsstandaard. De assessor heeft het assessorenprofiel geïnternaliseerd waardoor het geheel hem/haar eigen is geworden. Voor de student is de beoordeling herkenbaar en het is duidelijk welk vereist niveau men dient te behalen

Concluderend: student voelt zich transparant, kundig en zo objectief mogelijk beoordeeld.

5.3 Vervolg vragen naar aanleiding van het onderzoek

Het onderzoek zoals hier beschreven kan nog verder geoptimaliseerd worden.

Een belangrijk aspect is het aantal respondenten uit de onderzoeksgroepen getraind en gecertificeerd nog beperkt was. Vergroting van die aantallen kan helpen bij het verder valideren van de kwaliteitsrubriek.

Een tweede aspect is de verbreding in de richting van criteriumgerichte beoordelaars van andere domeinen dan educatie.

Het onderzoek prikkelt tevens tot het stellen van nieuwe onderzoeksvragen:

1. Ervaren de assessoren de kwaliteitsrubriek criteriumgericht beoordelen als een zinvol en waardevol instrument?
2. Herkennen de assessoren zich voldoende in de verschillende gedragsindicatoren behorend bij de verschillende niveaus?
3. Willen de assessoren de kwaliteitsrubriek criteriumgericht beoordelen als ontwikkelingsgericht instrument gebruiken?
4. Ervaren studenten daadwerkelijk verschil in **transparantie** van het beoordelingsproces tussen ongetrainde, getrainde en expert assessoren?
5. Worden de beoordelingen van expert assessoren door de studenten ervaren als **meer transparant** en daarmee **minder subjectief**?
6. Zijn de verschillen in gebruik van gedragsinterventies tussen de groepen groter dan binnen dezelfde groep?
7. Bovenstaande vragen leiden ook tot onderzoek of het uitgangspunt van Ericsson (2006) juist is **“dat je alleen maar van expertise in een bepaald domein kunt spreken, wanneer er taken zijn aan te wijzen waarin experts het consistent en consequent beter doen, dan gevorderden en/of beginners.”**

5.4 Aanbevelingen

Op basis van het beschreven onderzoek en de beschreven niveaus komt de vraag op wat als een minimumniveau van assessoren in het HBO geëist zou moeten worden..

Wetend dat het takenpakket van opleiders in het HBO voor een groot deel bestaat uit het begeleiden en beoordelen van studenten is het van belang dat dit op een juiste, transparante en zo objectief mogelijke wijze gebeurt. Het studiesucces van studenten is immers mede afhankelijk van verkregen feedback en beoordelingen.

Dit is met name het geval bij het voeren van stage- en/of begeleidingsgesprekken waarbij er criteriumgericht beoordeeld wordt en waarbij de student dient te reflecteren op zijn leerproces en op zijn ontwikkeling in het licht van de gestelde doelen, en aan de hand van gestelde beoordelingscriteria en hbo-niveau dient zichtbaar te maken waar hij zich bevindt. Het weten wat je als student nog moet ontwikkelen, welke vragen je moet beantwoorden en welke leercontext opgezocht dient te worden om het gewenste gedrag te kunnen ontwikkelen is van groot belang.

De kwaliteitsrubriek criteriumgericht beoordelen laat zien dat het minimumniveau hiervoor eigenlijk het niveau van de expert assessor is. In het onderzoek is echter zichtbaar geworden dat een groot deel van de huidige assessoren nog in de beginfase van hun ontwikkeling als assessor zit en het gewenste eindniveau nog niet bezit. Het professionaliseren van de eigen opleiders naar expert assessoren via training en certificering is dan ook een taak waarvoor iedere HBO opleiding voor staat.

Op basis van deze constatering komen we tot de volgende aanbevelingen:

Aanbeveling 1

Neem de assessorentraining op in de basiskwalificatie onderwijs (BKO) als eerste stap van het professionaliseringsproces naar volwaardig assessor. Het voeren van criteriumgerichte beoordelingsgesprekken is nu geen onderdeel van de BKO.

Aanbeveling 2

Stop niet na de assessorentraining maar laat opleiders doorgroeien naar het niveau van expert assessor via certificering, zodat in allerlei begeleidings- en beoordelingssituaties studenten transparant en zo objectief mogelijk beoordeeld worden.

Aanbeveling 3

Het hebben van expert assessoren is financieel aantrekkelijk. Het intersubjectief beoordelen (2 beoordelaars) wordt als een dure toetsvorm gezien. Echter wanneer de opleiding expert assessoren kan inzetten wordt het werken met één assessor weer mogelijk. De kwaliteit is dan immers via de persoon van de assessor geborgd, zeker als tijdens de criteriumgerichte beoordelingen een geluidsopname gemaakt wordt, zodat bij een lastige beoordelingssituatie of bij een klacht een tweede beoordelaar ingezet kan worden voor het intersubjectieve oordeel.

Opleidingen kunnen bewust met de inzet van expert assessoren omgaan.

Prettige bijkomstigheid is verder dat examencommissie(s) expert assessoren in kunnen zetten bij herkansingen van studenten of bij klachten.

6. Samenvatting

Toetsen is een vak apart en niet elke goede opleider is direct ook een goede beoordelaar. Toetsen en leren zijn onlosmakelijk met elkaar verbonden en voor het verkrijgen van beter onderwijs zullen onderwijs en toetsing steeds beter geïntegreerd en op elkaar afgestemd dienen te worden. Professionalisering van de beoordelaars is daarvoor noodzakelijk.

Het onderzoek dat in dit rapport beschreven wordt, wil antwoord geven op de vraag of er inderdaad verschil in interventies waar te nemen is tijdens het criteriumgericht interview (CGI) tussen de verschillende groepen van assessoren (ongetraind, getraind en expert), en op de vraag of er verschillende professionele niveaus van criteriumgericht beoordelen zijn te beschrijven in de vorm van een kwaliteitsrubric (rubric) criteriumgericht beoordelen? Dit leidt tot de onderstaande onderzoeksvragen:

1. Is er verschil te observeren bij de 3 onderzoeksgroepen (ongetraind, getraind en expert) in het gebruik van de interventies die de assessoren bij het criteriumgericht beoordelen hanteren tijdens de afname van het CGI? Het onderzoek richt zich vooral op de eerste vijf onderdelen van het assessorenprofiel.
2. Als blijkt dat er verschil is in de hantering van de onderdelen van het assessorenprofiel, zijn er dan patronen waar te nemen in de interventies die de assessoren gebruiken?
3. Is er een kwaliteitsrubric criteriumgericht beoordelen te beschrijven waarin de drie verschillende professionele niveaus (ongetraind, getraind en gecertificeerd) qua interventies met betrekking tot criteriumgericht beoordelen zichtbaar zijn.

Om de onderzoeksvragen te kunnen beantwoorden is het noodzakelijk om de assessoren te observeren bij het afnemen van assessments. De onderzoeker heeft daarom gekozen voor observaties van ongetrainde, getrainde en expert assessoren. Dit was mogelijk doordat de onderzoeker vanuit haar assessorentrainingspraktijk direct toegang heeft tot het observeren van alle drie de onderzoeksgroepen ongetraind – ervaren - expert. De observaties waren deels 'live' en deels observaties van op DVD opgenomen assessments.

De methode van onderzoek is kwalitatief. Gekozen is voor de observatie als waarnemingstechniek om via de directe waarneming (zien en horen) de gedragsinterventies die de assessor pleegt tijdens het criteriumgericht interview (CGI) te observeren en te analyseren. Dit gebeurt in de natuurlijke context van de assessmentafname. De onderzoeker gebruikt het assessorenprofiel als kijkkader om de CGI-gesprekken en DVD's te analyseren.

Op basis van de observaties die in het kader van dit onderzoek zijn uitgevoerd, kunnen drie prototypische assessorenprofielen beschreven worden, die onderling verschillen in de mate waarin ze systematisch gebruik maken van gesprekstechnieken, in staat zijn om tot objectieve en transparante oordelen te komen, zich bewust zijn van de kwaliteit van hun handelen en gericht zijn op het verbeteren daarvan.

Het gaat hier om prototypische beschrijvingen: individuele assessoren kunnen ook zonder formele training tot de categorie expert assessor horen, en assessoren kunnen kenmerken uit verschillende prototypen hebben.

Een **ongetrainde assessor** hanteert bepaalde, eenvoudige gesprekstechnieken die hij ooit geleerd heeft in zijn opleiding tot opleider. Vanuit die basiskennis weet hij onvoldoende de focus van het criteriumgericht interview (CGI) duidelijk te maken aan de student. Hij is voorts onvoldoende bekend in het hanteren van gesprekstechnieken, zoals START, LSD en TURBO, die ingezet worden om ontbrekende bewijzen helder door de student te laten tonen. Hierdoor is de beoordeling niet transparant voor anderen. De student kan de beoordeling als subjectief ervaren; hij is afhankelijk van niet expliciet gemaakte criteria die de beoordelaar hanteert.

Concluderend: Beoordeling is niet transparant, sterk subjectief getint en afhankelijk van de assessor. Er is geen uniformiteit in procedures en technieken zichtbaar.

De **getrainde assessor** kent en hanteert de juiste procedures en gewenste technieken om te komen tot criteriumgerichte beoordeling. De assessor gebruikt de technieken echter op een technische wijze en/of focust zich in beperkte mate op verdieping en reflectie. De assessor "vinkt af" en vraagt niet verder. Hierdoor krijgt niet alleen de assessor maar ook de student onvoldoende informatie. Het blijft onduidelijk hoe bewijsvoering van de student gerelateerd is aan beoordelingscriteria en bijbehorend hbo- en competentieniveau. De assessoren onderling komen wel tot een eenduidige en transparante manier van werken en beoordelen, maar missen een beslissende slag in hun eindoordeel.

Concluderend: De juiste procedures en gewenste technieken zijn in beeld maar worden nog te technisch toegepast. Tussen de assessoren is de uniformiteit op procedures en technieken groot. De student kan de beoordeling nog altijd als "subjectief" ervaren, doordat hij in zijn ogen onvoldoende heeft kunnen tonen wat volgens de complexe context wordt bedoeld

De **expert assessor** weet de juiste procedures en technieken toe te passen: het doel om te komen tot een transparante beoordeling die zo maximaal mogelijk objectief is wordt bereikt. De student en de assessoren weten op elk moment van het CGI 'waar' ze zich bevinden en welk onderwerp aan de orde is. De assessor let voortdurend op de juiste focus en houdt de regie op dynamische (natuurlijke) manier in handen. Het gesprek heeft de gewenste diepgang waardoor informatie betreffende de vereiste competenties wordt verkregen. De beoordeling geschiedt op concrete feiten en aantekeningen. De student heeft aan het einde van het assessment het gevoel "alles" te hebben kunnen zeggen en kan in het nagesprek "begrijpen" waarom bepaalde bewijzen die hij heeft gegeven, wel of niet "als voldoende" worden gehonoreerd. De wijze van beoordelen gebeurt eenduidig en transparant, zodat het eindresultaat een intersubjectief oordeel is, passend bij de gehanteerde beoordelingsstandaard.

Concluderend: student voelt zich transparant, kundig en zo objectief mogelijk beoordeeld.

Op basis van het beschreven onderzoek en de beschreven niveaus komt de vraag op wat als een minimumniveau van assessoren in het hbo geëist zou moeten worden. Wetend dat het takenpakket van opleiders in het hbo voor een groot deel bestaat uit het begeleiden en beoordelen van studenten is het van belang dat dit op een juiste, transparante en zo objectief mogelijke wijze gebeurt. Dit is met name het geval bij het voeren van stage- en/of begeleidingsgesprekken waarbij er criteriumgericht beoordeeld wordt en waarbij de student dient te reflecteren op zijn leerproces en op zijn ontwikkeling in het licht van de gestelde doelen, en aan de hand van gestelde beoordelingscriteria dient zichtbaar te maken waar hij zich bevindt. Het weten wat je als student nog moet ontwikkelen, welke vragen je moet beantwoorden en welke leercontext opgezocht dient te worden om het gewenste gedrag te kunnen ontwikkelen is van groot belang. Het studiesucces van studenten is immers mede afhankelijk van verkregen feedback en beoordelingen. De kwaliteitsrubriek criteriumgericht beoordelen laat zien dat het minimumniveau hiervoor eigenlijk het niveau van de expert assessor is. In het onderzoek is echter zichtbaar geworden dat een groot deel van de huidige assessoren nog in de beginfase van hun ontwikkeling als assessor zit en het gewenste eindniveau nog niet bezit. Het professionaliseren van de eigen opleiders naar expert assessoren via training en certificering is dan ook een taak waarvoor iedere hbo opleiding voor staat.

Inge Oudkerk Pool

7. Literatuur

- Baartman, L.K.J. (2008). *Assessing the assessment. Development and use of quality criteria for Competence Assessment Programmes*. Academisch proefschrift. Utrecht: Universiteit van Utrecht.
- Bekker-Holtland A., Brandenburg, C., de Bruin, S., Kösters, J., Reinink, M., Strijbis, J., Worp van der K.(2010). *Assessments in de lerarenopleiding*. ADEF projectgroep competentiegericht opleiden en beoordelen (COB).
- Berkel, van, A., Dinther, van M., Oudkerk Pool, I., & Speetjes, J., (2008). *Certificeren van assessoren in het hbo. Waarom en hoe?* Onderwijs Innovatie, nummer 2, pagina 17-25, juni 2008.
- Berliner, D. C. (2004). *Describing the behavior and documenting the accomplishments of expert teachers*. Bulletin of Science Technology & Society, 24(3), 200-212.
- Bloom, B.S., Hastings, J.T., & Madaus, G.F., (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill Book Company.
- Bloom, B.S., (1956). *Taxonomy of educational objectives, Handbook I – Cognitive Domain*. New York: David McKay Company, Inc.
- Boshuizen, H. P. A. (2003). *Expertise development: The transition between school and work*. Inaugural address. Heerlen: , Open Universiteit Nederland.
- Bruijns, V e.a. (1997). *Wie ben ik ,wat wil ik, wat kan ik ; een assessment als onderdeel van het hbo curriculum*, Hogeschool van Amsterdam, Amsterdam
- Commissie Bruijn (2012). *Vreemde ogen dwingen*. Eindrapport Commissie externe validering examenkwaliteit hoger beroepsonderwijs. Den Haag: HBO-raad.
- De Groot, A. D. (1946). *Het denken van den schaker*. Den Haag: Mouton.
- Driscoll, M. P. (2005). *Psychology of learning for instruction*. Boston: Allyn and Bacon.
- Eisner, E.W. (1982). *Cognition and Curriculum – A basis for deciding what to teach.*,New York/London: Longman.
- Eisner, E.W. (1991). *The Enlightened Eye – Qualitative Inquiry and the Enhancement of Educational Practice*. New York: Macmillan Publishing Company.
- Elshout-Mohr, M. e.a. (2004). *Functies van een elektronisch portfolio in een constructivistische leeromgeving*: Eindrapport. Amsterdam: SCO-Kohnstamm Instituut.
- Ericsson, K. A. (2009). *Enhancing the development of professional performance: Implications from the study of deliberate practice*. In K. A. Ericsson (Ed.), *The development of professional performance: Approaches to objective measurement and designed learning environments* (pp. 405-431). Cambridge: Cambridge University Press.
- Ericsson, K. A., Prietula, M.J., Cokely, E.T. (2007). *The making of an Expert*. Harvard Business Review (July-August 2007)
- Ericsson, K. A. (2006). *The influence of experience and deliberate practice on the development of superior expert performance*. In K. A. Ericsson, N. Charness, P. Feltovich, and R.R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685-706). Cambridge: Cambridge University Press.
- Haitink, A., & Oudkerk Pool, I. (2009) *Het één versterkt het ander – stages begeleiden en beoordelen*. In: *De Academische Opleidingsschool, Samen opleiden, onderzoeken en onderwijs ontwikkelen*. Amsterdam: De Hogeschool van Amsterdam en de Amsterdamse Stichtingen voor Katholiek Onderwijs.

- Heuvel van de R. (2012). *Adviesrapport digitale toetsproces HvA Amsterdam*: HvA DOO.
- Hutjes, J.M., & Buuren van, J.A. (2007). *De gevalsstudie – strategie van kwalitatief onderzoek*. Meppel/Amsterdam/Heerlen: Boom.
- Inspectie van het Onderwijs, Ministerie van Onderwijs, Cultuur en Wetenschap (2012): *Examencommissies & ervaringscertificaten – over EVC in het mbo en het hbo*. Utrecht: Inspectie voor het Onderwijs.
- Inspectie van het Onderwijs, Ministerie van Onderwijs, Cultuur en Wetenschap (2012): *De begeleiding van beginnende leraren in het voortgezet onderwijs*. Utrecht: Inspectie voor het Onderwijs.
- Inspectie van het Onderwijs, Ministerie van Onderwijs, Cultuur en Wetenschap (2011): *Zij-instroom in het Beroep, 2000-2009 – Geschiktheidsverklaringen in cijfers en feiten tot en met 2009*. Utrecht: Inspectie voor het Onderwijs.
- Kayzel, R., Nieweg M., & Kok M. (2006). *Toetsen en beoordelen; cahier 2*. Amsterdam: Hogeschool van Amsterdam.
- Lamers, L. Oudkerk Pool, I., Reinink, M., Scholte, W., & Zewald, M.(2009). *De assessment praktijk verbeteren door intercollegiale consultatie in de vorm van een audit in het hbo*. ADEF projectgroep competentiegericht opleiden en beoordelen.
- LeCompte, M.D., & Goetz, J.P. (1982). *Problems of reliability and validity in ethnographic research*, Review of Educational Research 52, 31 – 60.
- Luijten, A.J.M. & Wijnen W.H.F.W. (1995). *Toetsen in het studiehuis*. Tilburg: MesoConsult.
- Meijs, van der M., Smulders, H., Wee, ter E., (2004) *Het geschiktheidsonderzoek voor zij-instromers in de praktijk*. Stoas Onderzoek.
- Mieg, H. A. (2009). *Two factors of expertise? Excellence and professionalism of environmental experts*. High Ability Studies, 20(1), 91–115.
- Ministerie van OC&W (2012). *Kamerbrief hoger beroepsonderwijs*, 17 jan 2012. Den Haag:Ministerie van OCW.
- Molkenboer, H. & Soeting, J. (2011). *Toetsontwikkeling in de praktijk. Hoe maak ik goede vragen en toetsen?*, Wilp: Teelen Management bv.
- Oudkerk Pool, I., Reinink, M., Scholte, W., Klink, van der, M., & Haan, J.,(2008). *EVC in de lerarenopleidingen*, VELON Tijdschrift voor Lerarenopleiders 29(2) 14-21.
- Oudkerk Pool, Th. (1994). *Modulair Opleiden –Basisboek voor onderwijskunde*. Zutphen: Thieme.
- Prins, H.J. (1990). *Aanwijzingen voor stagebeoordelingen*. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.
- Rijst, H. van der, & Kok, J. (2011). *250 begrippen van het onderwijs*. Groningen: Noordhoff
- Riksen, D. (2010). *Digitaal toetsen*. SURF magazine, 2, 6-8.
- Ruijters, M., & Simons, R.J.(red.) (2012). *Canon van het leren*. Deventer: Kluwer
- Sanders, P. (red.) (2011): *Toetsen op school*. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling.
- Schön, D.A. (1983). *The reflective practitioner. How professionals think in action*. New York: Basis Books.

Sinke, G.P.J. e.a. (2006). *Aan de slag met assessment; toetsen en beoordelen in een competentiegerichte leeromgeving*. Nuenen: Onderwijsadviesbureau Dekkers.

Swennen, A. en M. van der Klink (Eds.) (2009). *Becoming a Teacher Educator. Theory and Practice for Novice Teacher Educators*. Dordrecht: Kluwer Academic Publishers Group.

Sluijsmans, D.M.A. (2008). *Duurzaam beoordelen in vraaggestuurd leren (lectorale rede)*. Nijmegen: Hogeschool Arnhem en Nijmegen.

Sluijsmans, D.M.A., Peeters, A., Jakobs, L., & Weijzen, S. (2012). *De kwaliteit van toetsing onder de loep*. *Onderwijs Innovatie*, 4, december 2012.

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

Tyler, R.W. (1949). *Basic principles of curriculum and instruction* Chicago/London: The University of Chicago Press.

Veen, T. van der, & Wal, J. van der (2012). *Van leertheorie naar onderwijspraktijk*. Groningen: Noordhoff Uitgevers.

Observatiemethodieken

NOA, (2002). *Beoordeling van praktijktoetsen m.b.v. de methode WAKKER*. In het instrument *Assessment voor opleidingen Personeel en Arbeid*.

Oudkerk Pool, I., (2010). *Gedragsobservatieformulier waar WAKKER methode en assessorenprofiel zijn samengevoegd ten behoeve van beoordeling assessorencertificering*, Hogeschool van Amsterdam.