

Factoren die van invloed zijn op de kwaliteit van de beslissingen van
praktijkexaminatoren tijdens de proeve van bekwaamheid in het mbo

Rianne Banning

9935320

Themagebied: leren in organisaties

Universiteit Utrecht

Begeleidster: Anne Khaled

Tweede beoordelaar: Eghe Osagie

Juni 2015

Samenvatting

De proeve van bekwaamheid is een exameninstrument dat in de praktijk competenties examineert van mbo-studenten. Veel mbo-instellingen worstelen met de kwaliteitsstandaarden, zoals *validiteit* en *betrouwbaarheid*, van deze vorm van competentiegerichte examinering (Tigelaar & Van der Vleuten, 2014). Onderwijsinstellingen moeten in het examenproces waarborgen dat de praktijkexaminatoren objectief, deskundig en betrouwbaar zijn (Inspectie van het Onderwijs, 2014). Het doel van dit onderzoek was inzicht krijgen in de factoren die de kwaliteit van de beslissingen door de praktijkexaminatoren tijdens de proeve beïnvloeden. Aan de hand van een conceptueel model, ontwikkeld vanuit de literatuur, is onderzocht door welke factoren de beslissing werd beïnvloed. Uit de analyse van 12 semigestructureerde interviews met praktijkexaminatoren, kan geconcludeerd worden dat de conceptuele factoren van invloed zijn op de beslissing van de deelnemers. Anders dan het conceptueel model liet zien, werden verbanden tussen de factoren gevonden. Het exameninstrument met de interpretatie en de toepassing van de criteria, stond centraal in elke beslissing en onderhield een relatie met andere factoren. Factoren als sociale relatie met de student, beoordelaars- en beroepservaring, reflectie en relatie met onderwijsinstelling hadden invloed op de interpretatie en de toepasbaarheid van de criteria. En daarmee ook invloed op de beslissing.

In de discussie worden mogelijke verklaren voor de bevindingen besproken en concrete implicaties voor de praktijk gegeven. Door het beperkt aantal deelnemers in huidig onderzoek is vervolgonderzoek noodzakelijk om de resultaten te kunnen generaliseren.

Sleutelwoorden: validiteit, betrouwbaarheid, kwaliteit, proeve, praktijkexaminator, examenproces, mbo

Inleiding

De snel veranderende maatschappij vraagt om afgestudeerde mbo-studenten die voorbereid zijn op een flexibele arbeidsmarkt (Schaap, De Bruijn, Van der Schaaf, Baartman, & Kirschner, 2011). Door competentiegerichte kwalificatiedossiers te ontwikkelen, wordt de brug geslagen tussen onderwijs en arbeidsmarkt. Het onderwijs zorgt in zijn onderwijsprogramma voor deze competentieontwikkeling door de integratie en ontwikkeling van kennis, houding en vaardigheden (De Vries, 2009; Schaap et al., 2011). Hiervoor is de participatie van studenten in de beroepspraktijk belangrijk, zodat studenten deze competenties kunnen oefenen en ontwikkelen in een reële situatie. Volgens De Bruijn en Leeman (2011) en Willemse en Denissen (2012) wordt er idealiter op de werkplek geëxamineerd. De afgelopen jaren zijn er verschillende vormen ontwikkeld om competenties te examineren op de werkplek, zoals praktijkopdrachten en proeven van bekwaamheid (Tigelaar & Van der Vleuten, 2014). Praktijkexaminatoren van praktijkopdrachten beoordelen de bewijslast, het product. De proeve van bekwaamheid is een authentiek competentiegericht examen, waarbij een student zijn competenties demonstreert in een realistische situatie zoals de beroepspraktijk. Tijdens de proeve van bekwaamheid (proeve)¹ observeren, interpreteren en beoordelen praktijkexaminatoren het product en het proces met behulp van een observatielijst met standaarden waaraan moet worden voldaan (Examenwerk, 2010; Van 't Spijker, 2012).

Alle proeven moeten net als alle andere examenvormen, voldoen aan kwaliteitsstandaarden van de onderwijsinspectie om accreditatie te ontvangen voor de certificering van studenten (Baartman, Bastiaens, Kirschner, & Van der Vleuten, 2007b; Willemse & Denissen, 2012). De onderwijsinspectie let hierbij o.a. op het *examenproces* en het *exameninstrument*. De examenstandaarden die zijn opgesteld door de inspectie, komen

¹ Niet alle mbo-opleidingen noemen dit type examen proeve van bekwaamheid. In huidig onderzoek voldoen de afgenomen examens inhoudelijk aan deze gehanteerde definitie van proeve van bekwaamheid en worden daardoor “proeve” genoemd.

overeen met de klassieke kwaliteitscriteria zoals *validiteit* en *betrouwbaarheid* (Van der Vleuten & Schwirth, 2005; Willemse & Denissen, 2012). *Validiteit* geeft aan of het examen meet wat het moet meten, *betrouwbaarheid* richt zich op de herhaalbaarheid (Van der Vleuten & Schwirth, 2005).

Om te voldoen aan de kwaliteitstandaarden voor het exameninstrument, kopen onderwijsinstellingen examens, waaronder de proeve, vaak in bij gecertificeerde examenleveranciers (Baartman et al., 2007b). Het garanderen van de kwaliteit van het examenproces, de afname en beoordeling, is voor onderwijsinstellingen vaak een worsteling (Tigelaar & Van der Vleuten, 2014; Willemse & Denissen, 2012). De reden voor deze worsteling is dat onderwijsinstellingen de objectiviteit en kwaliteit van de mensen die de examens beoordelen, de praktijkexaminatoren, moeten waarborgen en de deskundigheid moeten garanderen (Inspectie van het Onderwijs, 2014). Onderwijsverslagen van 2011 en 2012 laten zien dat onderwijsinstellingen te veel op afstand staan en onvoldoende zicht hebben op de kwaliteit van het examenproces (Servicepunt Examinering, 2014). De onderwijsinstellingen kunnen de deskundigheid en objectiviteit van de examinatoren niet garanderen, waardoor de kwaliteit van het examenproces in de reële praktijk onvoldoende is (Servicepunt Examinering, 2014).

Om de kwaliteit van het examenproces, de beoordeling van de proeve, te garanderen, lijkt het zinvol om het beslissingsproces van praktijkexaminatoren te onderzoeken. Dit kwalitatief onderzoek richt zich op de vraag door welke factoren de beslissingen van de praktijkexaminatoren worden beïnvloed tijdens de beoordeling van de proeve en wat dit zegt over de kwaliteit van de beslissingen, zoals validiteit en betrouwbaarheid. Door kennis te hebben van deze factoren, kunnen onderwijsinstellingen beter zicht houden op de beoordelingen in de praktijk tijdens de proeve. Zij kunnen met de uitkomsten richtlijnen opstellen voor het examineren van proeven in de praktijk. Doordat hiermee de kwaliteit van

de beoordelingsbeslissing verhoogd wordt, neemt ook de kwaliteit van het examenproces toe, omdat de beoordeling onderdeel uitmaakt van het examenproces (Willemse & Denissen, 2012).

Het bestuderen van het beslissingsproces van de praktijkexaminatoren levert daarnaast ook nieuwe wetenschappelijke kennis op. In het is mbo weinig onderzoek uitgevoerd naar de beoordelingsbeslissingen van praktijkexaminatoren waarbij de proeve als exameninstrument wordt gebruikt in de reële praktijk. Er zijn diverse onderzoeken verricht naar de kwaliteit van theoretische examens en portfolio's als exameninstrumenten (Greatorex, 2005; Harth & Hemker, 2013; Moss, Schutz & Collins, 1998). In onderzoeken naar examenprocessen gaat het meestal om docenten die beoordelen of om gesimuleerde praktijkexamens of theoretische examens. Er is echter weinig onderzoek gedaan naar de kwaliteit van de beoordeling van de proeve die op de werkplek wordt uitgevoerd door praktijkexaminatoren (Tigelaar & Van der Vleuten, 2014). Deze praktijkexaminatoren hebben vaak hetzelfde mbo-niveau als de studenten en zijn niet of in zeer gering mate pedagogisch of didactisch geschoold. Zij beoordelen studenten, door middel van observatie, in examensituaties die zeer divers en complex van aard zijn (Harth & Hemker, 2013). Onderzoeken die wel zijn gericht op praktijkexaminering, vinden meestal plaats in het hbo of universitaire onderwijs en richten zich op praktijkopdrachten (Baartman, Gulikers & Dijkstra, 2013; Baume, Yorke & Coffey, 2004; Moss, Schutz & Collins, 1998; Tigelaar & Van der Vleuten, 2014). De betrouwbaarheid en validiteit van de beoordeling van praktijkopdrachten is minder complex dan de beoordeling van proeve, omdat de proeve door middel van observatie wordt beoordeeld (Tigelaar & Van der Vleuten, 2014).

Theoretisch kader

Het theoretisch deel bestaat uit drie delen. Allereest zijn de concepten betreffende het examineren binnen de mbo-instellingen beschreven. Ten tweede is de kwaliteit van de

beslissingen van de praktijkbeoordelaars vanuit verschillende perspectieven benaderd.

Vervolgens is er een conceptueel model gepresenteerd met daarin de factoren die van invloed zijn op de beslissingen die de examiner van de proeve in de praktijk neemt.

Theoretische achtergrond

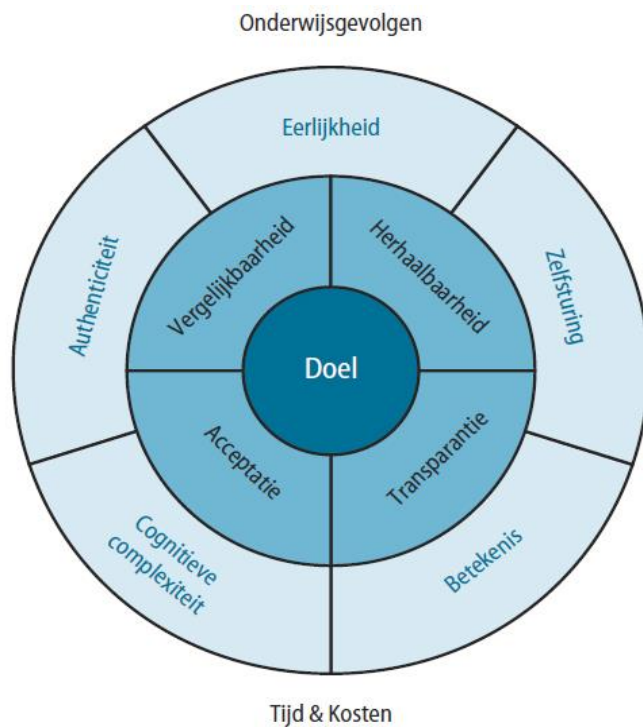
Het kwalificatiedossier van opleidingen binnen het mbo is opgebouwd uit kerntaken, werkprocessen, competenties en prestatie-indicatoren. De proeve is een kwalificerende examenvorm waarbij studenten in een reële beroepssituatie een set van samenhangende beroepsactiviteiten uitvoeren: de werkprocessen (Examenwerk, 2010). Voor een goed resultaat moeten studenten in verschillende situaties kennis, vaardigheden en houding in combinatie bewust kunnen toepassen (Examenwerk, 2010). Tijdens de proeve wordt door deskundige en onafhankelijke beoordelaars (praktijkexaminatoren) vastgesteld, of de deelnemer de werkprocessen van een kerntaak voldoende beheerst en weet toe te passen. De beoordeling in de proeve is gericht op het waarneembare gedrag. De praktijkexaminator observeert en interpreteert dit gedrag. Doordat vooraf de setting en de complexiteit van de opdrachten worden bepaald, is er sprake van redelijk gecontroleerde omstandigheden (Examenwerk, 2010; Willemse & Denissen, 2012).

De complexiteit van het valide en betrouwbaar examineren door de praktijkexaminatoren is bij de proeve hoger dan bij praktijkopdrachten, omdat in de proeve de praktijkvaardigheden zijn gereduceerd tot werkprocessen en deze door middel van observatie worden beoordeeld (Tigelaar & Van der Vleuten, 2014). Beslissingen die de praktijkexaminatoren nemen, worden beïnvloed door meer dan alleen de criteria waarop moet worden beoordeeld (Baartman, 2008). Deze beïnvloeding van de beslissing wordt vanuit in de literatuur vanuit verschillende perspectieven benaderd, namelijk volgens het *competentieassessment-wiel*, het *subjectief perspectief* en het *beïnvloedingsfactorenmodel*.

Perspectief vanuit competentieassessment-wiel. Baartman (2008) heeft een competentie-assessmentwiel ontwikkeld om competentieassessment te kunnen evalueren vanuit de kwaliteitscriteria validiteit en betrouwbaarheid, zie Figuur 1.

Het competentie-assessmentwiel is in verschillende onderzoeken (Baartman et al., 2007b; Baartman et al., 2013; Willemse & Denissen, 2012) gebruikt om individuele examenonderdelen en volledige examenprogramma's te evalueren. In dit onderzoek kan een aantal kwaliteitscriteria uit het competentie-assessmentwiel worden toegepast om de beoordeling van de proeve door de praktijkexaminator te evalueren. Het criterium *transparantie* is een indicatie voor de interpretatie van het exameninstrument door de praktijkexaminatoren. Als handleidingen, criteria en doelen goed zijn beschreven, wordt het nemen van beslissingen makkelijker (Baartman, 2008; Trede & Smith, 2014). De criteria *vergelijkbaarheid* en *herhaalbaarheid* evalueren de verschillen tussen beoordelingen van de praktijkexaminatoren op andere momenten en in andere situaties. Praktijkexaminatoren beoordelen studenten in het bedrijf waar zij werkzaam zijn. *Authenticiteit* verschilt tussen de sectoren en beroepen waar studenten voor leren. Dat dit invloed zal hebben op de beoordeling is evident. De praktijkexaminatoren beoordelen studenten vanuit hun eigen praktijksituatie en zij passen hier de beoordelingscriteria op aan. Eisen die bijvoorbeeld gesteld worden aan studenten "receptionist", zijn in een vijfsterrenhotel anders dan die op een boerencamping, terwijl de beoordelingscriteria hetzelfde zijn.

Doordat examineren plaatsvindt in de beroepssituatie, is het criterium *tijd en kosten* mogelijk van invloed op de beoordeling van de proeve (Baartman, Bastiaens, Kirschner & Van der Vleuten, 2007a). Praktijkexaminatoren beoordelen de studenten tijdens een periode waarin zij zelf ook werkzaamheden verrichten. Tijdsdruk die zij ervaren tijdens het nemen van beslissingen zou van invloed kunnen zijn.



Figuur 1. Competentie-assessmentwiel, Baartman (2008).

Door in dit onderzoek de criteria van Baartman te gebruiken, krijgen de criteria *betrouwbaarheid* en *validiteit* een duidelijke betekenis. In Tabel 1 zijn de criteria geclusterd volgens de klassieke criteria *validiteit*, *betrouwbaarheid*, *transparantie* en *uitvoerbaarheid* (Willemse & Denissen, 2012). Examinatoren van onderwijsinstellingen zijn zich vaak bewust van deze criteria in het examenproces. Of praktijkexaminatoren zich ook bewust zijn van validiteit en betrouwbaarheid in het examenproces is nog niet onderzocht. Er is nog geen onderzoek naar de beoordelingen van praktijkexaminatoren verricht waarbij gebruik is gemaakt van het competentie-assessmentwiel. Doordat in het mbo de proeve vaak wordt afgenomen in de reële praktijk zouden zij als, belangrijke stakeholder in het examenproces, zich bewust dienen te zijn van de criteria (Baartman et al., 2007b; Baartman, 2008; Baartman et al., 2013). Verschillende criteria van het competentie-assessmentwiel lijken toepasbaar in dit onderzoek en zullen worden meegenomen in het conceptueel model.

Tabel 1

De clustering van de kwaliteitscriteria van Baartman

Cluster criterium	Kwaliteitscriteria van Baartman
Validiteit	Authenticiteit Betekenisvolheid Cognitieve complexiteit Ontwikkeling van zelfsturend leren Onderwijsgevolgen
Betrouwbaarheid	Objectiviteit Acceptatie Herhaalbaarheid van beslissingen Vergelijkbaarheid
Transparantie	Eerlijkheid Geschiktheid voor onderwijsdoelen
Uitvoerbaarheid	Tijd en kosten

Noot. Uit "Examineren op de werkplek, De ontwikkeling van een assessmentmodel voor het mbo" door P. Willemse en N. Denissen, p.25. Copyright 2012 door Eindhoven School of Education & IVA beleidsonderzoek en advies

Subjectief perspectief. Trede en Smith (2014) beargumenteren in hun onderzoek dat het te smal gedacht is om praktijkexamens in termen van competentiestandaarden te examineren. Praktijkexaminatoren examineren immers niet in een cognitief rationele omgeving. Gebaseerd op empirisch onderzoek concluderen zij dat deze vorm van examineren een subjectief proces is. Dit wordt bevestigd door diverse onderzoeken (Harth & Hemker, 2013; Greatorex, 2005; Schaap, De Bruijn, Van der Schaaf, & Kirschner, 2009; Gulikers, Bastiaens, & Kirschner, 2004). Trede en Smith (2014) delen de ervaringen van praktijkexaminatoren in vijf hoofdthema's in: *ervaring, materiaal, beslissing, relationeel, zelfreflectie*. Vanuit het onderzoek van Trede en Smith (2014) vullen de thema's *relationeel* en *ervaring* dit onderzoek aan. Volgens hen zijn praktijkexaminatoren zich tijdens beslissingsmomenten bewust van de *relationele* elementen en subjectiviteit. Opvallend is dat zij vinden dat de beoordeling het best kan worden uitgevoerd door praktijkexaminatoren die de studenten kennen. Juist door de subjectieve en *relationele* factor zien zij zich beter in staat om de kwaliteiten en ontwikkeling van een individu te beoordelen. *Ervaring* in het beoordelen

van examens zorgt ervoor dat de praktijkexaminatoren minder emotioneel en zelfverzekerder zijn.

Naast de thema's *ervaring* en *relationeel* zijn de andere drie thema's *materiaal*, *beslissing* en *zelfreflectie* ook toepasbaar in dit onderzoek. Zij worden ondersteund door het onderzoek van Wyatt-Smith en Castleton (2007) die gebruikmaken van het *beïnvloedingsfactorenmodel*.

Beïnvloedingsfactorenmodel. Wyatt-Smith en Castleton (2007) hebben een model ontwikkeld dat kan worden gebruikt om tijdens beoordelingen beslissingen te nemen. Het model helpt daarnaast de beoordelaar te beredeneren hoe de beslissing is genomen. Figuur 2 laat relaties tussen de verschillende factoren zien. Het model van Wyatt-Smith en Castleton (2007) is ontwikkeld vanuit onderzoek naar het beoordelen door docenten. Doordat de verschillende elementen terugkomen in diverse onderzoeken naar de beoordelingsbeslissingen van examinatoren, is het een model dat bruikbaar is voor dit onderzoek.



Figuur 2. Factoren die de beslissing van beoordelaars beïnvloeden (Wyatt-Smit & Castleton, 2007).

Aanvullend in dit model van Wyatt en Castleton is de nadrukkelijke aanwezigheid van de *context*. Dit wordt ondersteund in het onderzoek van Sandal, Smith en Wangenstein (2014). De *context* wordt benaderd vanuit verschillende invalshoeken. Allereerst de reële situatie in relatie tot de onderwijsinstelling. De beslissingen genomen door de praktijkexaminatoren vinden plaats in een samenwerkingscontext tussen school en werkplek. De verwachtingen van de praktijkexaminatoren komen niet altijd overeen met de verwachtingen van de onderwijsinstellingen en de studenten en dit heeft invloed op de beoordeling.

De tweede *context* waarin de praktijkexaminatoren beoordelen, is die van beoordelaars zelf. Ervaringen delen en discussiëren over de betrouwbaarheid zorgt voor consistente beoordelingen en een hogere interbeoordelaarsbetrouwbaarheid (Harth & Hemker, 2013). Als praktijkexaminatoren functioneren binnen een groep van examinatoren waarbij zij getraind kunnen worden en ervaringen kunnen uitwisselen, zal dit de beoordelingsbeslissing beïnvloeden. Wyatt-Smith en Castleton (2007) leggen hiermee ook een direct verband met de *expertise* en *scholing* van de praktijkexaminatoren (Baume et al., 2004). Het blijkt dat examinatoren vaak niet goed de criteria en de uitslagen begrijpen. Daarnaast zijn ze het vaak niet eens met de criteria en uitkomsten. Als examinatoren getraind zijn, neemt de *betrouwbaarheid* van de beoordeling toe. *Ervaring* en *uitwisseling* zijn factoren die in dit onderzoek onderzocht kunnen worden.

Conceptueel model

Vanuit de verschillende perspectieven is beschreven welke factoren invloed hebben op de beslissingen die worden genomen door de examinatoren. De literatuur toonde aan dat er overeenkomsten te vinden zijn tussen deze factoren. De factoren uit verschillende onderzoeken werden naast elkaar gelegd en samengebracht tot zes factoren. Deze zijn terug te vinden in Tabel 2.

Tabel 2

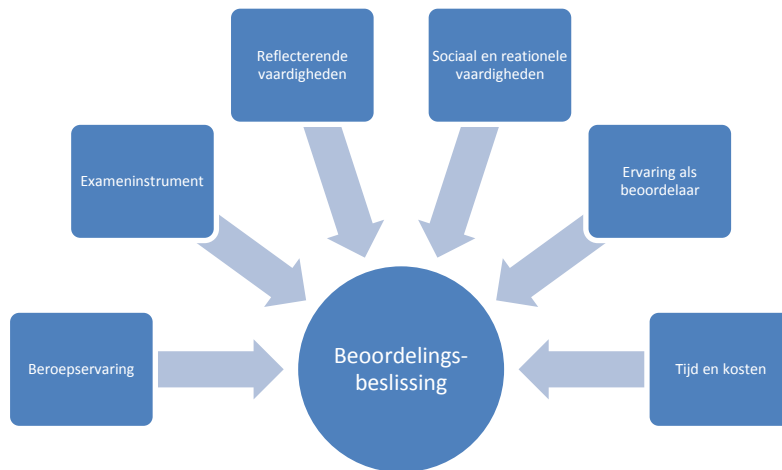
Factoren die van invloed zijn op de kwaliteit van de beslissingen van praktijkexaminatoren tijdens de proeve van bekwaamheid

Factoren	Trede en Smith (2014)	Wyatt-Smith & Castleton (2007)	Baartman (2008)
Beroepservaring	Ervaring	Kennis van pedagogische context Ervaring Kennis van context	Acceptatie Authenticiteit
Exameninstrument	Materiaal	Assessment criteria en standaarden	Authenticiteit Betekenisvolheid
Reflecterende vaardigheden	Beslissing	Redelijkheid	Herhaalbaarheid Vergelijkbaarheid
Sociale en relationele vaardigheden	Relationeel	Observatie van de student Student kennen	Objectiviteit Eerlijkheid
Ervaring als beoordelaar	Zelfreflectie	Ervaring als beoordelaar	Eerlijkheid
Tijd en kosten			Tijd en kosten

Daarnaast konden we op basis van de literatuur aannemen dat deze factoren van invloed zouden zijn op de kwaliteit van de beslissingen die werden genomen tijdens examens. Zoals eerder vermeld, hebben de voorgaande studies zich niet gericht op de beslissingen tijdens de proeve in de reële praktijk in het mbo. In dit onderzoek zijn de in de literatuur gevonden factoren gebruikt voor het ontwikkelen van een conceptueel model (Figuur 3). Dit conceptueel model is gebruikt om de beslissingen van de praktijkexaminatoren tijdens beoordeling van de proeve in de reële praktijk te onderzoeken.

Door inzicht te krijgen in de beïnvloedingsfactoren, kan de onderwijsinstelling richtlijnen formuleren, zodat de kwaliteit van de beslissingen door praktijkexaminatoren tijdens de proeve in de reële praktijk verhoogd kan worden. Hiermee kunnen de

kwaliteitsstandaarden die zijn opgesteld door de onderwijsinspectie om het examenproces te garanderen worden beïnvloed (Inspectie van het Onderwijs, 2014).



Figuur 3. Conceptueel model van factoren die van invloed zijn op de kwaliteit van de beoordelingsbeslissing genomen door de praktijkexaminator in de reële praktijk in het mbo.

Onderzoeksvraag

Dit onderzoek is uitgevoerd aan de hand van de vraagstelling: *Welke factoren zijn van invloed op de beslissingen die praktijkexaminatoren nemen tijdens de proeve van bekwaamheid van mbo-studenten en wat zegt dit over de kwaliteit van deze beslissingen?*

Methode

Onderzoeksopzet

Huidig onderzoek is een beschrijvend kwalitatief onderzoek. Om een beeld te krijgen van hoe en waarom praktijkexaminatoren hun beslissingen namen, diende men de factoren die de kwaliteit van de beslissing beïnvloeden te analyseren. De invloeden op de beslissingen die praktijkexaminatoren namen, zijn complex. Daarom is er gekozen voor semigestructureerde interviews. Semigestructureerde interviews boden ruimte om dieper in te gaan op antwoorden.

Deelnemers

Aan dit onderzoek namen 12 praktijkexaminatoren deel (6 mannen, 6 vrouwen). Een extra deelnemer participeerde in de pre-test. De deelnemers zijn werkzaam als examinator in een bedrijf of instelling. Zij examineerden mbo-studenten in de recreatie-, zorg-, commerciële dienstverlenings-, sport en bewegen- of handelssector, omdat deze sectoren examineerden gedurende de onderzoeksperiode. De kenmerken van de deelnemers lieten diversiteit zien in leeftijd, onderwijssector, aantal afgenomen examens, aantal jaren beroepservaring en deelname aan beoordelaarscursussen, zie Tabel 3.

Tabel 3

Beschrijvende gegevens deelnemers

<i>n</i>	sector	leeftijd in jaren	beroepservaring in jaren	aantal afgenomen examens	cursus gevolgd op gebied van examineren
1	recreatie	45-55	≥ 10 jaar	> 20	langer dan 10 jaar geleden
2	recreatie	20-30	< 10 jaar	< 10	in afgelopen 5 jaar
3	zorg	20-30	< 10 jaar	< 10	in afgelopen 5 jaar
4	zorg	45-55	≥ 10 jaar	> 20	in afgelopen 5 jaar
5	zorg	55-67	≥ 10 jaar	> 20	in afgelopen 5 jaar
6	com dien	45-55	≥ 10 jaar	<10	geen
7	com dien	45-55	≥ 10 jaar	10-20	geen
8	handel	55-67	≥ 10 jaar	> 20	geen
9	handel	55-67	≥ 10 jaar	> 20	langer dan 10 jaar geleden
10	handel	45-55	≥ 10 jaar	> 20	langer dan 10 jaar geleden
11	sport	30-45	≥ 10 jaar	< 10	geen
12	sport	30-45	≥ 10 jaar	> 20	in afgelopen 5 jaar

Noot. N=12

Instrumenten

Ten behoeve van de semigestructureerde interviews werd een interviewprotocol ontwikkeld (zie Bijlage 1 voor het protocol). De vragen zijn gebaseerd op de vragenlijst van Trede en Smith (2014), Baume en collega's (2004) en het conceptueel model (Figuur 3). In het protocol zijn de vragen ingedeeld in thema's: algemene vragen, inleidende vragen, exameninstrument, sociale en relationele vaardigheden, beoordelaarservaring,

beroepservaring, reflecterende vaardigheden en tijd en kosten. Een aantal van deze thema's komt overeen met de thema's waarin Tred en Smith (2014) hun interviewvragen hebben geclusterd. Net als vragen uit het onderzoek van Baume en collega's (2014) werden er ook vragen opgesteld over mogelijke problemen die optreden tijdens de examinering, zoals onzekerheid van de examinerator en de onduidelijkheid van de criteria.

De interviewvragen hadden allemaal betrekking op het laatst afgenomen examen van de praktijkexaminator. Door middel van deze concrete situatie kon er worden doorgevraagd op de factoren. Ook eventuele ervaringen uit voorgaande examens waar de praktijkexaminatoren op ingingen werden meegenomen in het interview, omdat dit relevant zou kunnen zijn voor het onderzoek. Tijdens het interview werd na elk thema en aan het einde van het interview door de interviewer samengevat. Dit om te controleren of er nog aanvullingen nodig waren en of de input van de deelnemer goed werd weergegeven.

Procedure

Het onderzoek werd mondeling voorgelegd aan een directeur van een roc in het oosten van Nederland. Na afstemming over de uitvoering werden met behulp van examencoördinatoren van de onderwijssectoren, zeven praktijkexaminatoren geselecteerd. Daarnaast werden twee medewerkers van een willekeurige zorginstelling en basisschool benaderd. Deze twee medewerkers selecteerden in totaal vijf praktijkexaminatoren, die voldeden aan de deelnemerscriteria.

De praktijkexaminatoren maakten allen gebruik van een proeve opgesteld door een gecertificeerde examenleverancier. De selectie werd gemaakt op basis van de onderwijssector, het aantal praktijkbeoordelingen en het wel of niet hebben gevolgd van examineringscursussen. Diversiteit van het aantal praktijkbeoordelingen en het wel of niet hebben gevolgd van examineringscursussen was noodzakelijk, omdat dit van invloed zou kunnen zijn op de resultaten. Deelnemers hadden in de periode van drie weken voor het

interview een of meerdere examenstudenten beoordeeld met de proeve als exameninstrument. Examenresultaten werden voorafgaand aan het contact tussen de interviewer en de deelnemer vastgesteld door de onderwijsinstelling, zodat het interview geen invloed zou kunnen hebben op het examen.

Dertien praktijkexaminatoren werden gebeld met de vraag of zij mee wilden werken aan het onderzoek. Bij een positief antwoord werd een afspraak gemaakt voor het interview. Bij een negatief antwoord werd een nieuwe deelnemer geselecteerd die voldeed aan de selectiecriteria. Gegevens werden alleen voor het onderzoek gebruikt. De anonimiteit van deelnemers werd gewaarborgd doordat er geen namen van deelnemers, bedrijven en onderwijsinstellingen werden gebruikt in het onderzoek.

Na de introductie werden de interviews afgenomen op de werklocatie van de praktijkexaminatoren en auditief opgenomen, met het oog op het transcriberen van de interviews. Tijdens het interview werden memo's geschreven om indrukken en gedachten van de interviewer systematisch bij te houden. De interviews duurden ongeveer een uur. In een pilot werd gecontroleerd of de vragen werden begrepen door de praktijkexaminator. Er heeft alleen een verandering plaatsgevonden in de volgorde van de vragen. De factor "instrument" werd naar voren gehaald. Gedurende de periode van afname van de interviews werden de vragen aangepast in diepte en de breedte, om zo het iteratieve analyseproces te ondersteunen.

Analyse

De interviews werden integraal getranscribeerd op basis van geluidsopnames. De non-verbale informatie werd, indien mogelijk relevant, meegenomen in de transcripties. Elk getranscribeerde interview werd zorgvuldig gelezen, waarbij gelet werd op het onderscheid tussen wel of niet relevant om de onderzoeksvraag te beantwoorden (Baarda et al., 2013). Delen die niet relevant waren voor het onderzoek, werden verwijderd. Dit waren bijvoorbeeld delen waarin de deelnemers spraken over de diverse werkzaamheden in de bedrijven. De

overgebleven delen werden in fragmenten gedeeld op basis van betekenisvolle eenheden. Daarna werden de fragmenten in het softwareprogramma NVivo-10 ingevoerd ten behoeve van de analyse met behulp van dit programma.

Er werd geanalyseerd vanuit een combinatie van deductieve en inductieve methode (Boeije, 2012). Volgens de deductieve methode werd vanuit de conceptuele factoren uit de literatuur een codeerschema ontwikkeld met werkdefinities. Zo kreeg de code “overleg collega” de werkdefinitie: “ de examinerator overlegt met collega(s) voordat de beslissing wordt gemaakt”. Dit codeerschema werd samen met het eerste gefragmenteerde interview besproken met een onderwijskundige om overeenstemming te bereiken in de interpretatie van de werkdefinities (Chi, 1997). Na overeenstemming werden de fragmenten van de eerste drie interviews gecodeerd door de onderzoeker. Tijdens dit proces werden het getranscribeerde interview en de memo’s met elkaar vergeleken. Fragmenten werden geïnterpreteerd en gecodeerd volgens het codeerschema. Niet alle fragmenten konden gericht met het codeerschema gecodeerd worden. Deze fragmenten werden in vivo gecodeerd (Baarda et al., 2013) Zo werden “indirect invloed van onderwijsinstelling”, “leeftijd” en “levenservaring” toegevoegd aan het codeerschema. Deze inductief ontwikkelde codes leken in verband te kunnen worden gebracht met het beantwoorden van de onderzoeksvraag.

Na de eerste codeerronde van drie van de twaalf interviews werden twee interviews aan de hand van de codeboom door een tweede onderwijskundige gecodeerd. De interbeoordelaarsbetrouwbaarheid werd hiermee getoetst. De Cohen’s kappa werd vastgesteld op .62. De interbeoordelaarsbetrouwbaarheid was hiermee voldoende (Landis & Koch, 1977). Vervolgens werden de volgende negen interviews gecodeerd. Na deze tweede fase van codering werd verzadiging bereikt (Baarda et al, 2013; Boeije, 2012) . Er waren geen nieuwe thema’s en codes nodig. De codelijst bestond uit zeven thema’s en 28 codes (zie Bijlage 2 voor de codelijst).

Door vervolgens de gegevens axiaal te coderen kwam er een structuur in de bevindingen naar voren (Baarda et al, 2013). Tijdens deze fase van axiaal coderen werden fragmenten voortdurend vergeleken. Door steeds op zoek te gaan naar vergelijkingen werden overeenkomsten en verschillen tussen de codes ontdekt. Hiervoor werden twee manieren gebruikt. Er werden categorieën gevormd, doordat thema's werden afgebakend en beschreven. Daarnaast ontstond er een gelaagde ordening, een boomstructuur (Baarda et al, 2013). Deze boomstructuur met hoofd- en subcategorieën ontstond door codes te splitsen (top-downwerkwijze) en door codes samen te voegen (bottom-upwerkwijze) (Baarda et al., 2013). Zo werd bijvoorbeeld de code “eerste indruk” samengevoegd met de “code persoonlijke ontwikkeling”.

Tenslotte werd er tijdens de fase van selectieve codering gekeken naar de hiërarchie en de verbanden tussen de hoofd- en subcategorieën die tijdens de fase van axiaal coderen waren ontstaan (zie Tabel 4). De belangrijkste categorieën werden bepaald en een kerncategorie werd geformuleerd. Deze kerncategorie stond centraal, omdat veel andere categorieën ermee verbonden waren (Boeije, 2014). De verbanden tussen de hoofd- en subcategorieën werden geanalyseerd door de gecodeerde fragmenten kritisch met elkaar in verband te brengen in NVivo-10 (Baarda et al., 2013). Een voorbeeld van zo'n fragment was “Ik denk dat mijn ervaring een rol speelt. Voor jonge beoordelaars zijn de criteria een goede kapstok, maar ik beoordeel op mijn gevoel” (deelnemer 6). In dit voorbeeld werd een relatie zichtbaar tussen beoordelaarservaring en het toepassen van de criteria. NVivo-10 toonde de hiërarchie van de relaties door het aantal keren dat de gecodeerde relatie zichtbaar was en het aantal bronnen waarin het voorkwam (zie Bijlage 3 het overzicht van de relaties).

Tabel 4

Hiërarchie van categorieën

Kerncategorie	Hoofdcategorie	Subcategorie	Bronnen	Referenties
Exameninstrument			12	155
		Interpretatie en toepassen criteria	12	134
		Handleiding	8	21
		Sociale relatie	12	170
		Persoonlijke ontwikkeling student	11	64
		Persoonlijke vaardigheden student	12	59
		Relatie met student	11	37
		Reflecterend	12	169
		Overleg collega's	12	52
		Persoonlijke vaardigheden	12	48
		Overleg schoolexaminator tijdens examen	9	13
		Aantekeningen	9	13
		Gebruik methode	6	9
		Beoordelaars-ervaring	12	72
		Ervaring	12	44
		Training	12	28
		Relatie met roc	10	41
		Beroepservaring	12	28
		Werkervaring	12	17
		Kosten en tijd	11	24

Noot. Referenties geeft het aantal gedoeerde fragmenten aan. Bronnen geeft aan uit hoeveel interviews deze fragmenten afkomstig zijn (maximum = 12)

Resultaten

Toelichting categorieën en relaties. De resultaten worden op basis van categorieën en relaties beschreven, waarbij de deelnemerskenmerken en de verschillende sectoren worden meegenomen. Allereerst wordt de kerncategorie beschreven. Vervolgens zullen de belangrijkste categorieën en relaties beschreven worden. Verbanden die het meest prominent aanwezig waren d.w.z. in de meeste interviews voorkwamen en de meeste aantal referenties aangaven komen aan bod.

Exameninstrument. Het exameninstrument bestond uit het toepassen van de criteria, de hanteerbaarheid en de handleiding. De handleiding was bij 11 deelnemers aanwezig. Er

waren vier ervaren beoordelaars die de handleiding niet gebruikten, omdat zij op de hoogte waren van de inhoud uit eerder afgenomen proeven. De beslissing werd niet beïnvloed door de handleiding. De beslissing van de deelnemers werd door alle deelnemers beïnvloed door het toepassen van de criteria. Om de criteria te kunnen toepassen werden door alle deelnemers de criteria geïnterpreteerd. Alle deelnemers gaven aan dat in de interpretatie van de criteria de voor iedereen anders is. “Iedereen interpreteert de criteria anders. Bijvoorbeeld vanochtend nog. Toevallig hebben we dat net in de pauze besproken met elkaar” (deelnemer 3).

Tijdens de analyse van de data werd duidelijk dat de categorie “interpretatie en toepassen criteria” de kernfactor was die de beslissing van de beoordelaar beïnvloedde. Alle overige thema’s en categorieën konden in verband worden gebracht met de interpretatie en toepassing van de criteria.

Sociale relatie. Figuur 4 laat het verband zien tussen de sociale relatie en de genomen beslissing tijdens de beoordeling van de proeve.



Figuur 4. Verband tussen persoonlijke vaardigheden en beslissing.

Tien deelnemers hadden een sociale relatie opgebouwd met de student, omdat de student voorafgaand aan de proeve stage had gelopen of omdat de student tijdens de stage verschillende proeven aflegde. Een deelnemer kende de examenstudent niet ten tijde van beoordelen en een deelnemer kende de examenstudent alleen uit verhalen van collega’s. Van 10 deelnemers die een sociale relatie met de student hadden opgebouwd gaven acht deelnemers aan dat zij persoonlijke vaardigheden uit de stageperiode mee lieten spelen in de beoordeling, ook al ontbraken de persoonlijke vaardigheden als expliciete criteria in de

meeste proeven. Een voorbeeld hiervan is te zien in het citaat: “Ik kijk niet puur naar de criteria, het ligt aan het persoontje” (deelnemer 4). Drie deelnemers probeerden de beslissing te nemen ten tijde van het examenmoment en niet de stage ervoor te laten meewegen in de beslissing. Zij gaven echter aan dat de persoonlijke vaardigheden die de student tijdens de stageperiode liet zien, onbewust mee zouden kunnen spelen tijdens het nemen van de beslissing.

De beslissing van de deelnemer werd ook beïnvloed door de persoonlijke ontwikkeling die de student tijdens de stageperiode liet zien, zie Figuur 5. Zeven deelnemers



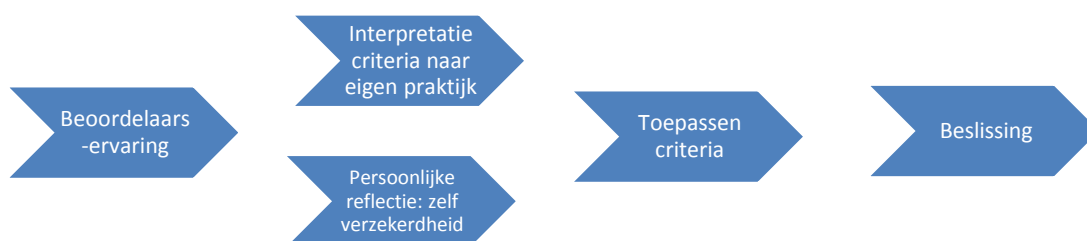
Figuur 5. Verband tussen persoonlijke ontwikkeling en beslissing.

gaven aan dat zij tijdens de beslissing werden beïnvloed door de persoonlijke ontwikkeling die de student doormaakte tijdens de stageperiode. Hiermee lieten zij de criteria van de proeve beïnvloeden door het proces voorafgaand aan het examenmoment. Een voorbeeld hiervan komt uit de sport- en bewegensector: “Als het examen onvoldoende is dan kijk ik naar de stage, het proces. Als het zo onvoldoende is dat het niet kan bijdragen dan moet het examen herkanst worden” (deelnemer 12). Een deelnemer uit de handelssector: “Iedereen komt op een ander niveau binnen en maakt een andere groei door. Ja, dat beoordeel ik anders. Ik kijk naar het startpunt” (deelnemer 8). Zes deelnemers gaven aan dat zij het als voordeel zagen om deze persoonlijke ontwikkeling de beslissing te laten beïnvloeden. “Want het kan zijn dat net tijdens het examen de bewoner misschien op dat moment boos of agressief wordt. Dan moet je tegen een leerling zeggen: “je hebt het niet gehaald”, terwijl ik weet dat ze het wel kan” (deelnemer 3).

In de zorgsector is een hoge mate van variatie in leeftijd van de studenten (tussen 16 en 55 jaar). Alle drie deelnemers uit deze sector gaven aan dat deze persoonlijke ontwikkeling van studenten niet mee zou mogen spelen in de examenbeslissing. Zij gaven echter tegelijk aan dat zij de levenservaring van oudere studenten mee lieten spelen in het ontwikkelingsproces en mogelijk ook in de beoordeling.

Twee examinatoren beoordeelden alleen de proeve en hadden geen sociale relatie opgebouwd met de student. Beide gaven aan dat zij ervaring hadden in het beoordelen van studenten die voorafgaand aan de proeve wel stage hadden gelopen. Op de vraag aan een van de deelnemers of de beslissing door de sociale relatie anders werd gemaakt, reageerde zij: “Nee en dat verbaast me nog. Ik doe het met evenveel plezier. En eh, ik kan het zo beoordelen zoals ik vind dat ik het moet doen, of ik nu iemand wel ken of niet ken” (deelnemer 1).

Beoordelaarservaring. Zeven van de 12 deelnemers hadden elk meer dan 20 mbo-proeven afgenomen. Deze beoordelaarservaring kon in verband worden gebracht met het interpretatie van criteria en zelfverzekerdheid (persoonlijke reflectie) van de beoordelaar, zie Figuur 6.



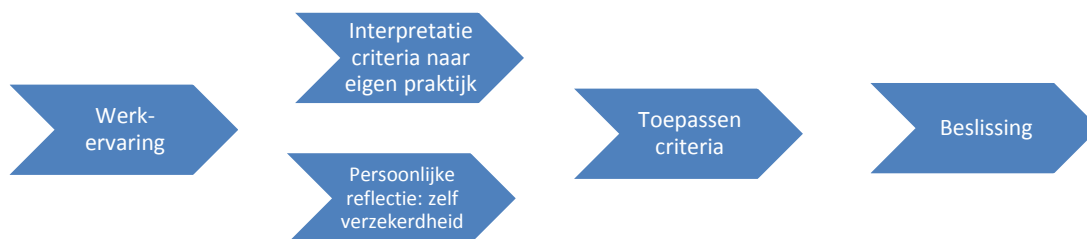
Figuur 6. Verband tussen beoordelaarservaring en beslissing.

Ervaren beoordelaars interpreteerden de criteria vanuit de eigen beroepscontext waarin zij ook veel ervaring hadden (werkten langer dan 10 jaar in de branche). Een van de deelnemers zei over zijn zoon die net was gestart met examineren: “Ik heb tegen hem gezegd: Ga daar maar eens mee aan de slag, maar hij zal zich puur aan de kapstok (de criteria)

vasthouden. Ik zie de kapstok als hartstikke leuk, maar ik beoordeel jou als hoe ik jou vind of hoe jouw prestaties zijn tijdens de gehele stageperiode” (deelnemer 6). Dit voorbeeld was tekenend voor ervaren beoordelaars; hoewel zij zich bewust waren van de criteria ervoeren zij het examineren als *fingerspitzengefühl*. Zij gebruikten hiervoor ervaring uit eerdere afgenomen proeven, eigen interpretatie van de criteria en de praktijkervaring.

Ervaren beoordelaars waren ook zelfverzekerder tijdens het nemen van de beslissing. Zij twijfelden minder over hun oordeel: “Onervaren beoordelaars tekenen volgens mij heel snel af. Ik denk dat ik dat in het begin ook wel eens heb gedaan. Die verantwoordelijkheid is ook wel wat. Dat je niet de discussie aan wilt gaan” (deelnemer 4). Drie deelnemers gaf aan dat zij vroeger studenten hadden laten slagen, terwijl zij dit nu niet meer zouden doen.

Werkervaring. In het onderzoek waren er 10 deelnemers die meer dan 10 jaar werkervaring hadden. Werkervaring had invloed op de interpretatie van criteria en op de



Figuur 7. Verband tussen werkervaring en beslissing.

zelfverzekerdheid in de persoonlijke reflectie, zie Figuur 7. “Ik denk dat werkervaring van invloed is op de beslissing. Door de jaren bouw je dat op. Je weet waar je voor staat, waar een student aan moet voldoen. De criteria interpreteer je daarop” (deelnemer 1). De deelnemer die voor het eerst een proeve sport en bewegen had afgenomen, maar die al wel examenervaring had binnen andere sectoren, reflecteert: “Een beginner zal anders beoordelen. Minder strak, omdat ze zelf ook nog niet goed weten wat goed en fout is” (deelnemer 11).

Reflecterende vaardigheden. Op de vraag wat de sterke en zwakke punten zijn van de deelnemers waardoor de beslissing zou kunnen worden beïnvloed, werden persoonlijke vaardigheden genoemd. Tijdens de analyse werden deze ingedeeld in zes vaardigheden, zie Bijlage 4. De beoordelaars- en werkervaring speelden hier een rol in, zoals al eerder is beschreven. In de laatst afgenomen proeve speelde zelfverzekerdheid een rol bij 10 deelnemers. Er werd niet getwijfeld over de beslissing. Uit het onderzoek bleek dat in eerder afgenomen proeven wel eens getwijfeld werd of dat er behoefte was aan bevestiging van de beslissing. Negen deelnemers gaven aan graag te overleggen met collega's in geval van twijfel bij het nemen van de beslissing, zie Figuur 8. In de zorgsector en de recreatiesector werd de beslissing nooit alleen genomen en was er altijd reflectie met een collega.

Alle deelnemers vonden het lastig om studenten in de laatste fase van de opleiding te laten zakken. De redenen hiervoor liepen uiteen. Vier deelnemers zagen dit als een persoonlijke nederlaag. Zij waren niet in staat geweest om een student zo te begeleiden dat de student slaagde.



Figuur 8. Verband tussen twijfel en genomen beslissing.

Overleg schoolexaminator. De deelnemers kenden de schoolexaminator verschillende rollen toe. Tabel 5 geeft een overzicht van de rollen van de schoolexaminator tijdens de afname van de proeve. Deze diverse rollen van de schoolexaminator waren verschillend van invloed op de beslissing van de deelnemer. Twee proeven werden tevens beoordeeld door de schoolexaminator. De deelnemer en de schoolexaminator kwamen hierbij samen tot een beslissing. Beide deelnemers ervoeren de subjectiviteit van de schoolexaminator als storend

tijdens het nemen van de beslissing. “Ik beoordeelde met twee docenten. De ene docent was veel strikter als de ander” (deelnemer 1). En een deelnemer over een eerder afgenomen proeve: “Zij zitten tegenover iemand waar zij al vier jaar lang aan hebben trekken bijvoorbeeld” (deelnemer 12).

Tabel 5

De rol van de schoolexaminator tijdens het examen in relatie tot sector

Rol van schoolexaminator	Aantal interviews	Sector
Beoordelende rol	2	Recreatie Sport en bewegen
Controlerende rol	6	Commerciële dienstverlening Recreatie Handel
Niet direct betrokken	4	Zorg Sport en bewegen

Noot. N=12

In de zorgsector hadden de deelnemers geen overleg met de schoolexaminator. Dit overleg met de onderwijsinstelling werd vanuit een centraal aanspreekpunt, de opleidingscoördinator van de zorginstelling, gevoerd. De schoolexaminator had dus geen invloed op de beslissing.

Een derde rol die de schoolexaminator aannam, was de controlerende rol. Hierbij werd tijdens een gesprek de beoordeling van de examiner met de student besproken en gecontroleerd. De student werd geacht de werkzaamheden onderliggend aan de beoordelingscriteria toe te lichten. De beslissing van de deelnemer werd niet beïnvloed door dit gesprek. In eerder afgenomen proeven werd de beslissing tussen een “goed” of een “voldoende” een enkele keer aangepast door de onderwijsinstelling.

Drie deelnemers gaven aan zich wel eens onder druk te hebben voelen staan door de onderwijsinstelling tijdens het maken van de beslissing. Deze druk werd op twee manieren uitgelegd; indirect en direct. Indirect, doordat de deelnemers de onderwijsinstelling

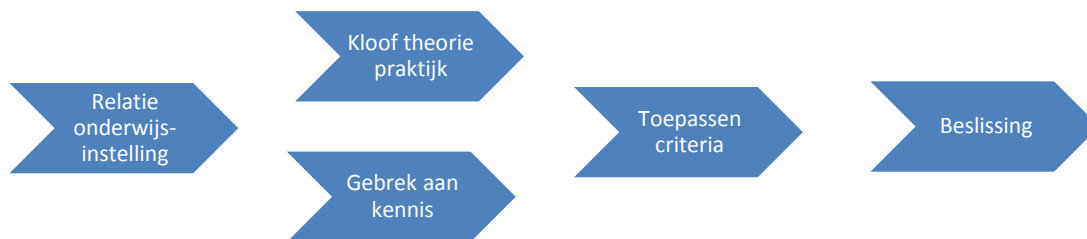
verantwoordelijk achtten voor het laten doorstromen van de student naar het einde van de opleiding. Indien de student het laatste examen (de proeve) niet zou halen, dan was de volledige onderwijstijd voor de student weggegooid en had hij geen diploma. De beslissing die de deelnemers namen, had dus grote gevolgen voor de student. Deze indirecte druk werd tijdens eerdere beoordelingen gevoeld door de twee deelnemers en dit had destijds invloed gehad op de beslissing.

De directe druk van de onderwijsinstelling werd door deelnemers gevoeld, doordat de schoolexaminator aangaf de student niet te willen laten zakken in verband met de bekostiging van de student voor de onderwijsinstelling. Twee deelnemers gaven aan dat hun negatieve beslissing in een eerder afgenomen proeven om deze reden werd overruled door de onderwijsinstelling.

Relatie met onderwijsinstelling. Twee kenmerken van de relatie met de onderwijsinstelling hadden invloed op de beslissing van de deelnemer, zie Figuur 9. Allereerst werd er een drie maal een kloof ervaren door de deelnemer tussen het onderwijsprogramma en de af te nemen proeve. De deelnemers gaven aan dat zij de examencriteria zelf dienden aan te passen op het bedrijf. Een voorbeeld hiervan was dat in de commerciële dienstverleningssector de criteria gericht waren op de verkoop van goederen, terwijl de student zijn proeve deed in een dienstverlenend bedrijf. De deelnemers gaven aan dat de relatie met de onderwijsinstelling verantwoordelijk was voor deze kloof. De onderwijsinstelling was te weinig op de hoogte van de praktijksituatie.

Daarnaast gaven zes deelnemers aan theoretische kennis te missen bij de studenten tijdens het uitvoeren van het examen. Ook dit gemis werd de onderwijsinstelling aangerekend. In de stageperiode voorafgaand aan de stage probeerden zij de kennis zelf over te brengen op de studenten. Tijdens het nemen van de examenbeslissing pasten zij de criteria aan door deze kennis als vragen toe te passen.

Twee examinatoren gaven aan niet meer met een onderwijsinstelling samen te werken om een van beide redenen: “Ze zijn onvoldoende competent in het opleiden van studenten” (deelnemer 1). “Als ik studenten van roc xxx krijg is het een grote bende” (deelnemer 6).



Figuur 9. De invloed van de relatie met de onderwijsinstelling op de beslissing.

Aantekeningen. Drie deelnemers maakten tijdens het examen aantekeningen die gebruikt werden tijdens het nemen van de beslissing. Deze aantekeningen werden apart genoteerd en niet opgeschreven of overgenomen in het examenverslag. Vijf examinatoren gebruiken tijdens de beslissing aantekeningen die werden gemaakt tijdens de stageperiode. Daarnaast maakte één deelnemer een persoonlijk verslag over het totale functioneren tijdens de stage én het examen (Figuur 10). Dit kwam overeen met de behoefte aan onderbouwing van de beslissing van de deelnemers. Een citaat van een deelnemer die aantekeningen maakt: “Wat je vindt, wat je ziet en wat je denkt en wat je hoort, dat moet je wel onderbouwen als je het meeneemt in je beoordeling” (deelnemer 4).



Figuur 10. De invloed van het maken van aantekeningen op de beslissing.

Training. Vijf deelnemers hadden in de afgelopen vijf jaar een training gevolgd op het gebied van het beoordelen van proeven. De vier deelnemers uit de zorgsector en de sport- en bewegensector waren positief over het nut van deze training. Tijdens deze trainingen werden praktijkvoorbeelden voorgelegd ter kalibratie van de criteria. De beslissing die werd genomen tijdens de proeve werd beïnvloed door deze trainingen, zie figuur 11. De training van de deelnemer uit de recreatieve sector was niet gericht op de kalibratie van criteria, omdat niet alleen examinatoren uit de recreatieve sector de training volgden. Doordat de inhoud van de training niet aansloot op de behoefte en verwachtingen van de deelnemer ervoer hij deze al nutteloos.



Figuur 11. De invloed van training op de genomen beslissing.

Methode. Alleen de examiner die in de sport- en bewegensector examineerde, gebruikte de WAKKER-methode². Zij gaf aan dat haar beslissing werd beïnvloed door de toepassing van deze methode.

Kosten en tijd. De beïnvloeding van kosten en tijd in de examenbeslissing werd door geen examiner ervaren. Wel werd de tijdsinvestering tijdens de stageperiode meegenomen in de beslissing of de student zijn stage mocht afronden inclusief de proeve.

Kwaliteit van de beslissing. De gevonden verbanden tussen de thema's en de categorieën werden in verband gebracht met de kwaliteit van de genomen beslissing door de deelnemer. Uit de interviews blijkt dat de kwaliteit van de beslissing onder druk staat. De

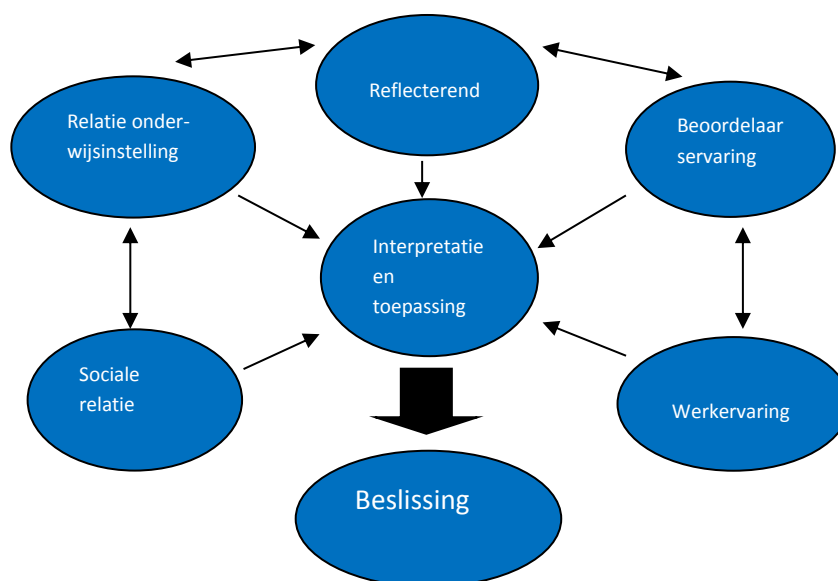
² WAKKER-methode: Waarnemen, Aantekeningen maken, Klassificeren, Kwantificeren, Evalueren, Rapporteren (Examenwerk, 2010). De WAKKER-methode wordt ook afgekort met WACKER-methode, waarbij de C Classificeren aangeeft (Cinop, 2015).

deelnemers interpreteerden de criteria anders. Doordat de proeve in een reële praktijksituatie werd uitgevoerd, werden de criteria geïnterpreteerd vanuit deze situatie. Daarnaast lieten zij de criteria beïnvloeden door de verschillende categorieën, zoals is beschreven.

De beslissing werd in de meeste proeven door een praktijkexaminator gemaakt. De praktijkexaminatoren gaven zelf aan dat de herhaalbaarheid van de proeve bij een andere examinator of in een andere praktijksituatie zou kunnen leiden tot een andere beoordeling.

Conclusie en discussie

Het doel van deze studie was inzicht te verkrijgen in de factoren die van invloed zijn de beslissing die praktijkexaminatoren nemen tijdens de proeve van bekwaamheid van mbo-studenten en wat dit zegt over de kwaliteit van deze beslissing. Daarvoor is eerst, op basis van literatuur, een conceptueel model gepresenteerd. Vervolgens zijn praktijkexaminatoren geïnterviewd over factoren die hun beslissing beïnvloeden. Dit onderzoek toont aan dat het conceptueel model grotendeels aansluit bij de invloed op de beslissingen die tijdens de proeve in de reële praktijk worden gemaakt. Toch zijn er enkele wijzigingen te maken naar aanleiding van dit onderzoek. Het onderzoek laat zien dat er een kernfactor is en dat er verbanden zijn tussen de factoren, zie figuur 12.



Figuur 12. Herontwerp conceptueel model.

Het eerste opvallende resultaat is dat de factor *exameninstrument* de grootste invloed heeft op de beslissing die de praktijkexaminator neemt tijdens de proeve in de reële praktijk. Het onderzoek toont aan dat de toepasbaarheid met onder meer de interpretatie van criteria, centraal staat in elke beslissing. In plaats van het exameninstrument als een van de factoren te zien, zoals in het conceptueel model, kan het exameninstrument als de centrale factor worden gezien die wordt beïnvloed door de andere factoren. Wolf (1995) schrijft in haar boek over examinering van competenties dat de criteria voor competentiegericht examineren juist zo beschreven zijn dat ze op meerdere manieren geïnterpreteerd kunnen worden. Hierdoor kunnen deze in verschillende beroepscontexten gebruikt worden. Dat alle deelnemers de criteria anders interpreteerden en de criteria anders toepasten is volgens haar inherent aan competentiegericht examineren.

Huidig onderzoek toont aan dat de interpretatie van het exameninstrument door meer factoren wordt beïnvloed. De interpretatie van criteria wordt in hoge mate beïnvloed door de sociale relatie die de examinerator heeft opgebouwd met de student. De Onderwijsinspectie (2014) adviseert om een scheiding te realiseren tussen stagebegeleiders en praktijkexaminatoren om de objectiviteit te waarborgen: formatieve beoordeling scheiden van summatieve. Het onderzoek toont aan dat dit in veel gevallen niet gebeurt: de stagebegeleiding en proevebeoordeling werden in tien van de twaalf proeven door eenzelfde persoon uitgevoerd. Een formatieve beoordeling is gericht op de ontwikkeling van de studenten waarbij het geven van feedback centraal staat (Dochy, 2009). De studenten ontvangen feedback om vaardigheden en kennis te ontwikkelen tijdens de werkzaamheden. De proeve, het summatieve examen, gaat om de beslissing “geslaagd” of “gezakt”. De onderbouwing van deze beslissing aan de student is hierbij niet relevant (Dochy, 2009). Huidig onderzoek laat zien dat de meeste examinatoren de proeve beoordelen op “formatieve wijze”. Zij gaven aan dat zij het juist als voordeel zagen om de stageperiode mee te laten

tellen in de proevebeoordeling, doordat er meer bewijsmateriaal ontstond. Daarnaast wilden zij graag hun beoordeling onderbouwen door middel van het geven van feedback. De vraag die hieruit voortvloeit is of de proeve niet beter kan worden ingezet als formatieve beoordeling. Dit sluit aan bij onderzoek van Sandal, Smith en Wangenstein (2014). Zij beargumenteren dat door de complexiteit van de proeve deze beter kan worden ingezet als formatief beoordelingsinstrument.

Uit het onderzoek blijkt dat de schoolexaminator in enkele gevallen samen met de praktijkbeoordelaar de proeven beoordeelt. Op deze manier houdt de onderwijsinstelling zicht op de examinering. Vanwege praktische en ethische redenen is het niet altijd mogelijk om een externe examiner op in de reële praktijk aanwezig te laten zijn (Willemse en Denissen, 2012). Huidig onderzoek toont aan dat de relatie met de onderwijsinstelling op verschillende manieren de beslissing beïnvloed. Opvallend is dat de relatie met de onderwijsinstelling mee werd genomen in de beslissing, doordat er een praktische of een theoretische kloof werd ervaren door de beoordelaar. Deze kloof tussen onderwijsinstelling en bedrijven wordt in meerdere onderzoeken binnen het mbo ervaren (De Bruijn & Leeman, 2011). De relatie tussen de onderwijsinstelling en de praktijkopleiders dient ervoor om praktijk en theorie goed met elkaar af te stemmen (De Bruijn & Leeman, 2011). Een intensief dialoog met alle betrokkenen over de inhoud en de afname van de proeve is hiervoor noodzakelijk.

Beperkingen

Voor de betrouwbaarheid van dit onderzoek werd een interviewprotocol ontwikkeld. Deze is gecontroleerd door de docent en aangepast na de pilot. De vragen waren gebaseerd op het conceptueel model en de literatuur. Voor de validiteit van het onderzoek is gebruik gemaakt van controlevragen in het interview (*member check*) en werden de relaties en verbanden gestaafd aan de interviewdata en memo's met behulp van het programma NVivo-

10. Bovendien werden collega-studenten en onderwijskundigen betrokken in het onderzoek (*peer debriefing*) (Baarda et. al, 2013) .

Toch kent huidig onderzoek een aantal mogelijke beperkingen. Ten eerste heeft het onderzoek een beperkt aantal deelnemers uit een beperkt aantal branches waar het mbo studenten voor opleidt. Dit zorgt ervoor dat de generaliseerbaarheid van het onderzoek ter discussie staat. Het is belangrijk om meer data te verzamelen om een gegronde uitspraak te kunnen doen en de bevindingen mogelijk te kunnen generaliseren naar alle praktijkexaminatoren in het mbo. Vervolgonderzoek zou hiervoor noodzakelijk zijn.

Ten tweede zijn de interviews afgenomen na de beoordeling, om de beslissing niet te beïnvloeden. De deelnemers werden binnen drie weken na de proeve geïnterviewd, waardoor ervaringen en kennis kunnen zijn weggezaakt. Een andere onderzoeksmethode, zoals observatie tijdens de beoordeling door de onderzoeker, of *stimulated recall* zou gehanteerd kunnen worden in vervolgonderzoek (Huijts, De Bruijn & Schaap, 2011).

Ten derde zijn de interviews afgenomen zonder het exameninstrument als hulpmiddel te gebruiken. Het door de beoordelaar ingevulde exameninstrument zou ondersteunend kunnen zijn, om zo concrete informatie te kunnen meenemen in het onderzoek. Door vanuit de verschillende sectoren de exameninstrumenten te gebruiken tijdens het onderzoek wordt duidelijk welke criteria het meest worden beïnvloed tijdens de beslissing.

Aanbevelingen

Als de proeve wordt gebruikt als summatief exameninstrument kan de kwaliteit van de beslissing die wordt genomen door de praktijkexaminator op verschillende manieren verhoogd worden. Allereerst zal de onderwijsinstelling de criteria moeten kalibreren met de beoordelaars. Dit kan tijdens een training aan de hand van praktijkcasus. Hierdoor kunnen de criteria eenduidiger worden geïnterpreteerd en worden toegepast in de eigen praktijkcontext. De kern van de beslissing ligt in de toepasbaarheid van de criteria. Als de interpretatie van de

criteria helder en duidelijk is voor de beoordelaar zal deze minder worden beïnvloed door de overige factoren, zoals onzekerheid, sociale relatie en ervaring (Tigelaar en Van der Vleuten, 2014). Door deze training te laten plaatsvinden op het roc met diverse praktijkbeoordelaars uit de branche kan de relatie met de onderwijsinstelling versterkt worden wat de kloof tussen onderwijs en praktijk verkleint (De Bruijn & Leeman, 2011).

Daarnaast moet het exameninstrument “proeve” gezien worden als een van de examenvormen binnen de examenmix waarop de student gekwalificeerd wordt (Baartman, 2007b, Dochy, 2009). De betrouwbaarheid en validiteit van examens is met de komst van het competentiegerichte onderwijs in een nieuw daglicht komen te staan wat heeft geleid tot nieuwe eisen aan examens. Belangrijk hierin is de noodzaak om examenbewijsmateriaal vanuit verschillende perspectieven te gebruiken (Baartman, 2008; Dochy, 2009; Willemsse & Denissen, 2012). Hiervoor kunnen verschillende examenvormen worden gebruikt, zoals theorie examens, criterium gerichte interviews of proeven afgenomen in andere praktijksituaties of simulaties. De kwaliteit van de examenmix is belangrijker dan de kwaliteit van een enkel onderdeel zoals de proeve (Dochy, 2009).

Het neemt niet weg, dat de proevebeslissing van de beoordelaar weloverwogen en zo objectief mogelijk gemaakt dient te worden aan de hand van de criteria. Echter het is een illusie dat alle beïnvloedende factoren moeten worden uitgeschakeld om de proeve kwalitatief goed te beoordelen. Juist door de diversiteit aan contexten en de authenticiteit van de proeve is de proeve een goed beoordelingsinstrument: formatief of summatief als een van de onderdelen van de examenmix.

Literatuurlijst

- Baarda, B. Bakker, E., Fischer, T., Julsing, M., De Goede, M., Peters, V., & Van der Velden, T. (2013). *Basisboek kwalitatief onderzoek. Handleiding voor het opzetten en uitvoeren van kwalitatief onderzoek*. Groningen/Houten: Noordhoff Uitgevers.
- Baartman, L. K. J. (2008). *Assessing the assessment: Development and use of quality criteria for Competence Assessment Programmes* (Promotieonderzoek). Verkregen van <http://dspace.learningnetworks.org/bitstream/1820/1555/1/Liesbeth%20Baartman.pdf>.
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2007a). Evaluating assessment quality in competence-based education: A qualitative comparison of two frameworks. *Educational Research Review*, 2, 114-129. doi:10.1016/j.edurev.2007.06.001
- Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2007b). Teachers' opinions on quality criteria for Competency Assessment Programs. *Teacher and Teacher Education*, 23, 857-867. doi:10.106/j.tate.2006.04.43
- Baartman, L., Gulikers, J., & Dijkstra, A. (2013). Factors influencing assessment quality in higher vocational education. *Assessment & Evaluation in Higher Education*, 38(8), 978-997. doi:10.1080/02602938.2013.771133
- Baume, D., Yorke, M., & Coffey, M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment & Evaluation in Higher Education*, 29(4), 451-477. doi:10.080/02602930310001689037
- Boeije, H. (2012). *Analysis in Qualitative Research*. Londen: Sage.
- Boeije, H. (2014). *Analyseren in kwalitatief onderzoek*. Den Haag: Boom Lemma Uitgevers.
- De Bruijn, E., & Leeman, Y. (2011). Authentic and self-directed learning in vocational education: Challenges to vocational educators. *Teaching and Teacher Education*, 27(4), 694-702. doi:10.1016/j.tate.2010.11.007

Chi, M. T. H., (1997). Quantifying Qualitative Analyse of Verbal Data: A Practical Guide.

The journal of the learning sciences, 6(3), 271-315. doi: 10.1207/s15327809jls0603_1

Cinop. (2015). Training voor assessoren. Verkregen op 24 mei 2015 van:

http://www.cinop.nl/1_3229_Training_voor_assessoren.aspx.

De Vries, G.J. (2009). *Onderzoek naar de invoering van competentiegerichte*

kwalificatiedossiers in het mbo. Deelonderzoek A: het beleidsproces. Bureau

Onderzoek en Rijksuitgaven. Verkregen van <http://euparl.net/9353000/1/j4nvgs5kjg>

[27kof_j9vvhskmycle0vf/vi2kavy3zkw0/f=/blg18933.pdf](http://euparl.net/9353000/1/j4nvgs5kjg27kof_j9vvhskmycle0vf/vi2kavy3zkw0/f=/blg18933.pdf).

Dochy, F. (2009). The Edumetric Quality of New Modes of Assessment: Some Issues and

Prospects. In *Assessment, Learning and Judgement in Higher Education*, 1-30

doi:10.1007/978-1-4020-8905-3_6

Examenwerk (2010). *Examenwerk. Handboek training competentiegericht beoordelen*.

Uitgegeven door de stichting Examenwerk, Zoetermeer.

Greatorex, J. (2005). Assessing the Evidence: different types of NVQ evidence and their

impact on reliability and fairness. *Journal of Vocational Education and Training*,

57(2).

Gulikers, J. T. M., Bastiaens, T. J., Kirschner, P. A. (2004). A Five-Dimensional Framework

for Authentic Assessment. *Educational Technology Research and Development*,

52(3), 67-86.

Inspectie van het Onderwijs (2014). *Veel gestelde vragen toezicht mbo*. Ministerie van

Onderwijs, Cultuur en Wetenschap. Verkregen van

<http://www.onderwijsinspectie.nl/binaries/content/assets/Documents+algemeen/2014/>

[veelgestelde-vragen-toezicht-mbo-1-december-2014.pdf](http://www.onderwijsinspectie.nl/binaries/content/assets/Documents+algemeen/2014/veelgestelde-vragen-toezicht-mbo-1-december-2014.pdf).

- Harth, H., & Hemker, B. T. (2013). On the reliability of vocational workplace-based certifications. *Research Papers in Education*, 28(1), 75-90. doi:10.1080/02671522.2012.754228
- Huijts, P.M., De Bruijn, E., & Schaap, H. (2010). Revealing personal professional theories. An explorative methodological study. *Qual Quant*, 45, 783-800. doi:10.1007/s11135-010-9322-z
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1) 159-174. doi: 10.2307/2529310
- Moss, P. A., Schutz, A. M., Collins, K. M. (1998). An Integrative Approach to Portfolio Evaluation for Teacher Licensure. *Journal of Personnel Evaluation in Education*, 12(2), 139-161. doi: 10.1023/A:1008081027322
- Sandal, A. K., Smith, K., & Wangensteen, R. (2014). Vocational Students Experiences with Assessment in Workplace Learning. *Vocations and Learning*, 7, 241-261. doi:10.1007/s12186-014-9114-z
- Schaap, H., De Bruijn, E., Van der Schaaf, M. F., & Kirschner, P. A. (2009). Students' personal professional theories in competence-based vocational education: the construction of personal knowledge through internalisation and socialisation. *Journal of Vocational Education & Training*, 61(4), 481-494. doi: 10.1080/13636820903230999
- Schaap, H., De Bruijn, E., Van der Schaaf, M. F., Baartman, L. K. J., & Kirschner, P. A. (2011). Explicating students' personal professional theories in vocational education through multi-method triangulation. *Scandinavian Journal of Educational Research*, 55(6), 567-586. doi:10.1080/00313831.2011.555922
- Servicepunt examinering mbo (2014). *Examinering in de beroepspraktijk. Beoordeling van de beroepspraktijkvorming de hoofdlijnen*. Verkregen van

<http://www.examineringmbo.nl/examineringmbo/download/servicepunt/20140829-examinering-in-de-beroepspraktijk-op-hoofdlijnen.pdf>.

- Tigelaar, D. E. H., & Van der Vleuten, C. P. M. (2014). Assessment of Professional Competence. *International Handbook of Research in Professional and Practice-based Learning*, 1237-1270. Springer Netherlands.
- Trede, F., & Smith, M. (2014). Workplace educators' interpretations of their assessment practices: a view through a critical practice lens. *Assessment & Evaluation in Higher Education*, 39(2), 154-167. doi:10.1080/02602938.2013.803026
- Van der Vleuten, C. P. M. & Schwirth, L. W. T. (2005). Assessing professional competence: from methods to programmes. *Medical Education*, 39(3), 309–317. doi: 10.1111/j.1365-2929.2005.02094.x
- Van 't Spijker, J.G.M. (2012). *De praktijkopleider als beoordelaar. De bereidheid van betrokkenen tot inzet van de praktijkopleider als beoordelaar bij summatieve beoordeling in de beroepspraktijk* (Masterthesis). Verkregen van: http://www.stagewerkt.nl/wp-content/uploads/2012/04/Def_Masterthesis_JGMvantSpijker_gew_layout.pdf.
- Wolf, A. (1995). *Competence Based Assessment*. Buckingham: Open University Press.
- Willemse, P., & Denissen, N. (2012). *Examineren op de werkplek. De ontwikkeling van een assessmentmodel voor het mbo*. Eindhoven: Eindhoven School of Education, Technische universiteit Eindhoven & Tilburg: IVA beleidsonderzoek en advies.
- Wyatt-Smit, C., & Castleton, G. (2005). Examining how teachers judge student writing: an Australian case study. *Journal of Curriculum Studies*, 37(2), 131-154. doi: 10.1080/0022027032000242887

Bijlage 1 Interview protocol

Vooraf

Meenemen notitieblok, pennen, audio apparatuur met volle batterijen, horloge.
Nummer respondent en datum noteren in de aantekeningen.

Introductie

Fijn dat u mij de gelegenheid biedt om een interview met u af te nemen. Mijn naam is Rianne Banning en ik ben student Onderwijskundig advies en ontwerp aan de Universiteit Utrecht. Voor mijn afstudeeronderzoek neem ik interviews af met diverse praktijkexaminatoren die mbo-studenten door middel van een proeve van bekwaamheid beoordelen in de praktijk. Het doel van het onderzoek is inzicht krijgen in de factoren die de beoordelingsbeslissing beïnvloeden. Graag zou ik met u tijdens dit interview uw laatste beoordeling willen uitpluizen om inzicht te krijgen in de factoren die uw beslissing hebben beïnvloed. Gegevens uit dit interview worden anoniem verwerkt en worden alleen gebruikt voor het onderzoek. Het interview duurt ongeveer een uur het wordt auditief opgenomen.
Heeft u nog vragen?

Inleidende vragen:

1. Kunt u iets vertellen over het laatste examen dat u heeft afgenomen.

Sub vragen:

- Wanneer is het examen afgenomen?
- Welk examen heeft u afgenomen?
- Kunt u iets vertellen over de inhoud van het examen?
- Hoe lang duurde het examen?

Exameninstrument

2. Kunt u iets vertellen over de criteria en de handleiding die u heeft moeten hanteren ter beoordeling van de werkzaamheden die de student uitvoert tijdens het examen.

Sub vragen:

- Waren de criteria en de richtlijnen voor u helder?
- Wat zijn de sterke en zwakke punten in de documenten en voorschriften die u heeft gebruikt tijdens de beoordeling?
- Op welke wijze en door wie is het exameninstrument aan u uitgelegd of toegelicht?
- Kunt u aangeven of u spanning heeft ervaren tijdens de beoordeling op basis van de criteria.
- Bent u afgeweken van de richtlijnen/criteria? Kunt u hier een voorbeeld van geven.

3. Kunt u iets vertellen over de aansluiting van het examen bij de werkzaamheden in uw bedrijf?

Sub vragen:

- Heeft u problemen ervaren met het examineren van onderdelen, omdat deze niet zijn voorgekomen tijdens de examenperiode?
- Mist u criteria waarop u in dit examen graag had beoordeeld? Kunt u dit toelichten?

4. De student wordt geëxamineerd voor een beroep dat ook in andere bedrijven in uw sector kan worden uitgevoerd. Heeft u hier rekening mee gehouden en zo ja op welke manier?

Sub vragen:

- Verwacht u dat uw collega van een ander bedrijf de student hetzelfde zal beoordelen in dat andere bedrijf? (toelichten a.d.h.v. concreet voorbeeld uit de branche)

5. Hoe heeft u ervaren dat de student was voorbereid op het examen door de onderwijsinstelling?

Sub vragen:

- Was de student op de hoogte van de verwachte werkzaamheden en de beoordelingscriteria?
- Waaraan merkte u dit?

Samenvatten

Reflecterende vaardigheden

6. Hoe heeft u de definitieve beslissing over het optreden van de student gemaakt tijdens het examen? Kunt u dit beredeneren?

Sub vragen:

- Maakte u bijvoorbeeld aantekeningen?
- Heeft u gebruik gemaakt van het oordeel van anderen, bijvoorbeeld uw collega's, om tot de beslissing te komen?
- Als u dit heeft gedaan, kunt u dit toelichten.
- Verschilde u van mening?

7. Heeft u gebruik gemaakt van advies of kennisdeling van de onderwijsinstelling om tot de beslissing te komen?

Sub vragen:

- Kunt u dit toelichten waarom wel of niet.

Sub vragen:

8. Wat ziet u zelf als uw sterke punten en zwakke punten tijdens de beoordeling?

Sub vragen:

- Waar heeft u op gelet om ervoor te zorgen dat u goed beoordeelde?
- Welke onderdelen van de examinering zijn voor u het meest uitdagend?

Samenvatten

Sociale en relationele vaardigheden

9. Wat was de eerste indruk die u had van de student en komt dit overeen met de beoordeling die u uiteindelijk heeft gegeven?

Sub vragen:

- Merkte u in uw beoordeling dat u de eerste indruk of verwachting mee heeft laten spelen in de beoordeling?

10. Welke relatie had u met de student voordat het examen werd afgenomen?

Sub vragen:

- Hoelang kende u de student voordat het examen is afgenomen?
- Heeft de student bij u stage gelopen? Zo ja hoelang?

11. Kunt u aangeven of de sociale relatie die u heeft opgebouwd met een student, een rol heeft speelt tijdens de beoordeling?

Sub vragen:

- Speelde bij de beslissing, de individuele ontwikkeling van de student een rol? Zo ja, kunt u dit toelichten?
- Beoordeelt u elke student volgens dezelfde criteria?
- Op welke wijze zal de beoordeling van een collega die de student niet kent, afwijken van uw beoordeling?

12. In hoeverre vindt u sociale vaardigheden van de student doorslaggevend in het wel of niet laten slagen ook al maakt zou het geen onderdeel uitmaken van het beslissingsmodel?

Sub vragen:

- Kunt u een voorbeeld noemen van een sociale vaardigheid die u mee laat spelen?
Eventueel toelichten met voorbeelden: te laat komen, initiatief nemen, enthousiasme.

Samenvatten

Ervaring als beoordelaar

13. Hoe heeft u geleerd om studenten te examineren?

Sub vragen:

- Heeft u een cursus gevolgd? Wanneer heeft u die gevolgd?
- Wat heeft u tijdens de cursus geleerd?
- Hoe past u deze kennis toe tijdens het examineren?

14. Hoeveel examens heeft u afgenomen?

Sub vragen:

- Kunt u aangeven op welke wijze deze ervaring van invloed is op de huidige beoordeling?
- Op welke wijze zal uw laatste beoordelaarservaring van invloed zijn op uw volgende examenbeoordeling?

15. Mogelijke vraag indien vaker geëxamineerd: Heeft u wel eens een student laten zakken?

Sub vragen:

- Kunt u hier meer over vertellen?
- Zo niet, waarom niet? Heeft het met uzelf te maken?

Samenvatten

Ervaring in beroep

16. Hoelang bent u werkzaam in uw huidige functie?

17. Op welke manier kan beroepservaring van invloed zijn op de beoordeling van de student?

Sub vragen:

- In hoeverre zal de beoordeling door een junior beroepsbeoefenaar anders zijn dan door een senior?

Samenvatten

Tijd en kosten

18. Kunt u aangeven wat de tijd en kosten zijn geweest van de examinering?

Sub vragen:

- Wegen deze op tegen de voordelen die het u heeft opgeleverd?
- Wat zijn de voordelen van het examineren van studenten voor u zelf?
- Heeft u wel eens het gevoel dat uw werkzaamheden onder druk staan, omdat u een student moet beoordelen?

Samenvatten

Afsluiting

Samenvatten van het interview. Bedanken. Aangeven dat het onderzoek in juni klaar zal zijn en dat de resultaten gedeeld zullen worden.

Bijlage 2

Hoofdthema	Subthema	Code	Beschrijving
Reflecterend		RE	De examiner reflecteert voordat de beslissing wordt genomen
	Gebruik methode	RM	De examiner maakt gebruik van een methode, zoals de WAKKER-methode, de STAR-methode.
	Aantekeningen	RA	De examiner maakt de beslissing op basis van gemaakte aantekeningen tijdens het examen
	Overleg collega's	RC	De examiner overlegt met collega('s) voordat beslissing wordt gemaakt.
	Overleg schoolexaminator tijdens examen	RS	De examiner overlegt over de beslissing met de onderwijsinstelling of schoolexaminator voordat de definitieve oordeel wordt gegeven
	Persoonlijke vaardigheden	RP	De examiner benoemt zijn persoonlijke vaardigheden in zijn rol als examiner (sterke of zwakke punten)
Examenervaring		EX	De ervaring als examiner beïnvloed de beslissing tijdens het examineren
	Training	ET	De training die de examiner heeft gevolgd heeft invloed op de beslissing die de examiner neemt tijdens de beoordeling
	Autodidact	EA	De examiner heeft zichzelf het examineren aangeleerd
	Instructie collega's	EC	De examiner heeft van collega's geleerd hoe te moeten beoordelen
	Ervaring	EV	De mate van ervaring van de examiner in het afnemen van examens
Beroepservaring		BE	De beroepservaring in relatie tot de beslissing van de examiner
	Werkervaring	BW	De examiner laat zijn beslissing beïnvloeden door mate van werkervaring
	Levenservaring	BL	De examiner ervaart dat zijn levenservaring van invloed is op de beslissing
Kosten en tijd		KO	De invloed van kosten en tijd op de beslissing
	Kosten	KK	De kosten zijn van invloed op de beslissing
	Tijd	KT	De tijdsinvestering van de beoordeling is van invloed op de beslissing
Sociale relatie		SO	De mate waarop de relatie die examiner heeft met de student of/en school, van invloed is op de beslissing
	Eerste indruk	SI	De eerste indruk is van invloed op de beslissing
	Relatie student	SR	De beslissing wordt beïnvloed door de relatie die is opgebouwd gedurende periode
	Persoonlijke ontwikkeling student	SP	De invloed van de persoonlijke ontwikkeling die de student doormaakt tijdens het examen op de beslissing
	Leeftijd	SL	De invloed van de leeftijd op de sociale relatie met de student en hiermee op de beslissing
	Persoonlijke vaardigheden	SV	De invloed van de persoonlijke vaardigheden van de student op de beslissing (motivatie, inzet e.d.)
Indirect rol van onderwijsinstelling		OO	De onderwijsinstelling beïnvloed het examen door haar optreden voorafgaand en tijdens het examenproces. (dit is niet de rol van de schoolexaminator)
	Sociale relatie	OS	De examiner wordt beïnvloed door de sociale relatie met de onderwijsinstelling (aantal contacten, afspraken die gemaakt zijn)
	Kloof onderwijs	OK	De examiner ervaart een algehele kloof tussen onderwijs en praktijk
	Vorbereiding	OV	De examiner wordt beïnvloed door de voorbereiding van de student door de onderwijsinstelling
Exameninstrument		IN	De invloed van het exameninstrument op de beslissing genomen door de examiner
	Aanwezigheid handleiding	IH	De examiner is op de hoogte van de inhoud van de handleiding
	Inhoud handleiding	II	De examiner neemt de beslissing volgens de richtlijnen uit de handleiding
	Vorbereiding student	IV	De student is op de hoogte van de criteria waarop de examiner de beslissing neem
	Ontbrekende criteria	IO	De examiner mist criteria in de beoordeling

	Interpretatie criteria	IA	De beslissing van de examinerator wordt gemaakt door gebruik van de criteria.
	Bruikbaarheid van criteria	IB	De bruikbaarheid van het exameninstrument (criteria zijn toepasbaar in de praktijk)
	Instructie door school	IS	De instructie van het examen door de onderwijsinstelling

Bijlage 3

Verbanden tussen de thema's en codes die van invloed zijn op de beoordelingsbeslissing

Categorie draagt bij aan tweede categorie		
Sociale relatie: pers. vaardigheden student	→	Toepassen criteria
Sociale relatie: pers ontwikkeling student	→	Toepassen criteria
Beoordelaarservaring	→	Reflecterend: persoonlijke vaardigheden
Beoordelaarservaring	→	Toepassen criteria
Ervaring: werkervaring	→	Reflecterend: persoonlijke vaardigheden
Ervaring: werkervaring	→	Toepassen criteria
Reflecterend: pers. vaardigheden	→	Overleg collega's
Reflecterend: aantekeningen	→	Toepassen criteria
Reflecterend: overleg schoolbeoordelaar	→	Toepassen criteria
Relatie met onderwijsinstelling	→	Exameninstrument
Beoordelaarservaring: training	→	Toepassen criteria
Reflecterend: gebruik methode	→	Toepassen criteria

Noot: op volgorde van meest naar minst prominente verband op basis van aantal bronnen en gecodeerde fragmenten.

Bijlage 4

Voorbeelden van genoemde persoonlijke vaardigheden die van invloed waren op het nemen van de beslissing

Persoonlijke vaardigheid	Voorbeeld van citaat
Observeren	“Ik kan heel goed observeren” (deelnemer 1) “In een half uur kan ik zien wat ik moet zien” (deelnemer 12)
Interpreteren van criteria	“Er is geen duidelijke interpretatie” (deelnemer 11) “Ik weet wat school wil zien” (deelnemer 12)
Twijfelen	“Ik vraag aan leerlingen of ik niet te streng ben” (deelnemer 3) “Ik vind het lastig ouderejaars negatief te beoordelen” (deelnemer 12) “Ik twijfel over de interpretatie van de criteria” (deelnemer 3) “Als ze goede inzet hebben en dan toch laten zakken” (deelnemer 4)
Duidelijk zijn	“Het komt nooit als verassing als ze zakken” (deelnemer 12) “Ik ben heel direct en heb geen twijfel” (deelnemer 7) “Er zijn examinatoren die student niet aan durven spreken” (deelnemer 12) “Je moet moeilijke gesprekken aan durven gaan” (deelnemer 12) “Door mijn leeftijd heb ik meer afstand tot een leerling en durf ik een leerling aan te spreken” (deelnemer 1)
Objectief zijn	“Ik heb een open vizier” (deelnemer 9) “Het gevaar is het halo-effect” (deelnemer 12) “Ik denk dat je objectiever bent als je de student niet kent” (deelnemer 11)
Motiveren	“Ik wil de beoordeling onderbouwen” (deelnemer 5) “Ik heb voor xxx een heel lang verhaal gemaakt van hoe ik hem heb ervaren en gezien” (deelnemer 6)