

Excellentieportfolio's beoordeeld: Waarop baseren examinatoren in het hoger beroepsonderwijs hun oordelen?

Jantine Huizing & Roeland van der Rijst

Samenvatting: Het doel van deze studie is inzicht verkrijgen hoe excellente prestaties, door studenten beschreven in een portfolio, worden beoordeeld. Dit inzicht kan docenten helpen bij een excellentieportfolio-beoordeling. In deze studie ligt de nadruk op de praktijkkennis van beoordelaars. Dataverzameling vond plaats door 36 beoordelingsformulieren en negen interviewtranscripten te analyseren vanuit een grounded theory benadering. In de analyse stond de samenhang tussen bewijslast, interpretaties en beoordelingscriteria centraal. Met de thema's die werden gevonden is een model ontwikkeld om de samenhang te beschrijven. Dit model geeft de relaties aan tussen vijf factoren die interacteren bij het beoordelen van een excellentieportfolio. De factoren zijn 1) context, 2) kenmerken van de portfolio-inhoud, 3) studentkarakteristieken, 4) gemeenschappelijke normvinding en 5) persoonlijke constructen van beoordelaars. Beoordelaars blijken vooral te letten op een combinatie van portfolio-inhoud en studentkarakteristieken (bijvoorbeeld initiatiefrijk zijn). De resultaten uit deze studie worden gerelateerd aan eerdere studies naar portfolio-beoordelingen. Implicaties voor de beoordelingspraktijk binnen excellentieprogramma's in het hoger onderwijs worden beschreven.

Trefwoorden: Excellentieonderwijs, honoursprogramma, portfolio-beoordeling, studentprestaties, vignettenonderzoek.

Auteurs: J.L. Huizing MSc. (jantine.huizing@hu.nl) is werkzaam bij de Faculteit Gezondheidszorg, Bachelor Medische Hulpverlening, Hogeschool Utrecht.

Dr. R.M. van der Rijst (rrijst@iclon.leidenuniv.nl) is werkzaam bij de Afdeling Hoger Onderwijs van het ICLON, Universiteit Leiden.

The assessment of portfolios in an honours program: on what grounds do assessors in higher vocational education base their judgments?

Summary: In this study we explored how excellent performances described by students in a portfolio are assessed. Results from this study can help both novice and experienced lecturers to better prepare for portfolio assessment of their students. We focus on the practical knowledge of assessors. The data consists of 36 student portfolios and transcripts from nine interviews with assessors. Both were analyzed using a grounded theory approach. Interviews were conducted by using vignettes of excerpts from existing student portfolios. In the analysis the relationship between proof of excellence, assessors' interpretations and assessment criteria were examined. Based on the results we developed a model to describe the relationships between five factors informing assessors' judgments; 1) context, 2) characteristics of portfolio content, 3) student characteristics, 4) accepted standards of assessment and 5) assessors' cognitive representations. Assessors put emphasis on a combination of portfolio content (e.g. what additional efforts students have made) and students' characteristics (e.g. students' initiatives). The results from this study are related to previous studies on portfolio assessment. Implications for assessment practices within honours programs in higher education are described.

Keywords: Gifted education, honours program, portfolio assessment, student performance, vignette study.

Inleiding

In het hoger onderwijs is een toegenomen waardering en aandacht voor getalenteerde studenten (Reis & Renzulli, 2010; Scager, Akkerman, Pilot & Wubbels, 2014; Wolfensberger, 2015). Om individuele leertrajecten van studenten in excellentieprogramma's in kaart te brengen worden vaak studentportfolio's gebruikt. Diverse studies naar het gebruik van portfolio's in het hoger onderwijs geven aan dat portfolio's een waardevol inzicht geven in het leren van studenten, maar dat beoordeling een complexe taak is (Driessen & van Tartwijk, 2006; van Tartwijk, Driessen, van der Vleuten & Stokking, 2007; van Tartwijk, van Rijswijk, Tuithof & Driessen, 2008). Om docenten inzicht te geven in het werken met portfolio's van excellente studenten is in deze studie onderzoek gedaan hoe excellentieportfolio's worden beoordeeld.

Portfolio's geven de mogelijkheid om zowel leerproces als leerprestatie in onderlinge samenhang te beoordelen. Daarom is een portfolio valide, vooral bij longitudinaal

verzamelde bewijslasten van meerdere bronnen (Gregori-Giralt & Menendez-Varela, 2015; Moss, 1994; Schuwirth & van der Vleuten, 2006). Daarnaast biedt een portfolio de mogelijkheid een individueel toegesneden beoordeling te geven (Driessen & van Tartwijk, 2006). Omdat de informatie vaak kwalitatief van aard is vanuit verschillende bronnen, raden Tichelaar et al. (2005) aan om een interpretatieve benadering gecombineerd met het gebruik van kwaliteitscriteria (beoordelingscriteria) te hanteren om geloofwaardigheid van de presentatie en het portfolioproces te bewijzen.

De beoordelingscriteria voor het excellentieportfolio in deze studie zijn algemeen geformuleerd waarmee recht wordt gedaan aan de open benadering als kern van het excellentietraject. Echter, een algemene formulering geeft beoordelaars ruimte voor meerdere interpretaties en kunnen aldus een probleem geven bij zowel betrouwbaarheid als validiteit van de portfoliobeoordeling. Aangezien in een formele toetsprocedure een eerlijke beslissing moet worden genomen over de prestaties is inzicht in betrouwbaarheid en validiteit van de portfoliobeoordeling relevant (Cook & Beckman, 2006; Tigelaar, Dolmans, Wolfhagen & van der Vleuten, 2005; van der Schaaf, Stokking & Verloop, 2005). Validiteit verdient volgens van der Schaaf et al. (2005) prioriteit boven betrouwbaarheid bij portfoliotoetsing omdat objectieve waardering van het portfolio lastig te bereiken is. Door de rijkdom en uniciteit vraagt een portfolio om interpretatie en moet rekening worden gehouden met de samenhang voordat beoordeling plaatsvindt (Moniz, Arntfield, Miller, Lingard, Watling & Regehr, 2015; Qvortrupa & Keiding, 2015). Moss (1994) en Tigelaar et al. (2005) bevelen daarom een hermeneutische benadering aan. Door de inhoud van een portfolio, kwalitatieve informatie in verschillende contexten, kunnen beoordelaars voornamelijk interpreteren wat is bedoeld, rekening houdend met de context. Cook en Beckman (2006) beschrijven in dit kader constructvaliditeit: dat is de mate waarin een beoordeling kan worden geïnterpreteerd als representatief voor het samengestelde materiaal, verzameld uit verschillende bronnen.

Cognitieve representaties van beoordelaars blijken bij een portfoliobeoordeling een belangrijke rol te spelen: beoordelingen en interpretaties worden gekleurd door persoonlijke opvattingen en ervaringen. Beoordelingen kunnen grotendeels worden verklaard door cognitieve representaties op twee dimensies: abstracte-concrete evaluatie (bijvoorbeeld "ze is een beetje lui" (abstract) en "ze zegt hier"(concreet)) en positieve-negatieve evaluatie (bijvoorbeeld "dat is toch wel goed"). Ook zonder persoonlijk contact is sprake van mentale interactie; dit komt omdat de mentale structuur van iemand filtert wat wordt waargenomen en hoe dat wordt geïnterpreteerd en beoordeeld. Een voorwaarde voor valide beoordelingen is dat beoordelaars hun oordeel baseren op de portfolio-inhoud. Concrete representaties bevorderen dit. Om beoordelingen te kunnen generaliseren naar andere situaties zijn abstracte evaluaties nodig. Inzicht in cognitieve representaties van beoordelaars is nodig om de kwaliteit van de beoordelingen te kunnen begrijpen (van der Schaaf et al., 2005).

Doel en vraagstelling

In deze studie is onderzocht hoe beoordelaars excellentieportfolio's beoordelen. Dit inzicht kan gebruikt worden voor docentprofessionalisering, bijvoorbeeld om nieuwe docenten te helpen excellentieportfolio's te beoordelen. Alhoewel voor portfolio-beoordelingen bij lerarenopleidingen relevante literatuur voorhanden is (Tigelaar et al., 2005; van Tartwijk et al., 2008), is weinig onderzoek beschreven dat inzicht geeft in het beoordelen van excellentieportfolio's. In dit onderzoek is nagegaan waar beoordelaars in eenzelfde situatie op letten. De leidende onderzoeksvraag was: hoe beoordelen beoordelaars de bewijslasten van studenten bij het verlenen van excellentiepunten? De deelvragen zijn:

- 1) Wat is de relatie tussen de bewijslasten en de beoordeling en welke keuzes maken beoordelaars daarbij?
- 2) Wat is de invloed van cognitieve representaties op de beoordeling?
- 3) Wat is invloed op de beoordeling wanneer beoordelaars de student kennen of begeleider van de student zijn geweest?

Methode

Context

Het bestudeerde excellentietraject van de hogeronderwijsinstelling voor gezondheidszorg heeft een omvang van 30 EC bovenop de reguliere studieomvang van 240 EC. Excellentiestudenten werken met extra opdrachten aan een verdieping van competenties ten opzichte van het bachelorniveau. Het excellentietraject kent een open curriculum met ruimte voor eigen initiatief en verschillende soorten prestaties. Daarmee wordt voor elke student een leertraject geboden dat is toegesneden op de eigen talenten (Excellentiecommissie Faculteit Gezondheidszorg, 2013). Het excellentietraject loopt van het tweede jaar tot het einde van het vierde jaar. Begeleiding is aanwezig en wordt idealiter door de student zelf geregeld, bijvoorbeeld via een lectoraat of een externe opdrachtgever. Het portfolio dat wordt opgesteld bestaat uit een afspiegeling van het leertraject en eindproducten van de prestatie, bijvoorbeeld een onderzoeksverslag of evaluatieverslag en een reflectie op de prestatie. De verwachte prestatie overstijgt het reguliere bachelorniveau, waarbij de mate van complexiteit, zelfsturing en relatie met de beroepspraktijk belangrijke graadmeters zijn. Het excellentietraject kent vijf invalshoeken waaronder een prestatie en de eraan ten grondslag liggende competentie kan worden geplaatst: 1) vakmanschap, 2) internationale ori-

entatie, 3) leiderschapskwaliteiten, 4) innovatie en disseminatie en 5) gedrevenheid. Elke invalshoek kent eigen beoordelingscriteria (Flierman, 2011).

In deze kwalitatieve studie staan ervaringen en denkwijzen van beoordelaars centraal. In het descriptieve deel werden de argumentaties van beoordelaars, beschreven in 36 beoordelingsformulieren uit verzamelde excellentieportfolio's bestudeerd. In het exploratieve deel werden interviews gehouden met negen beoordelaars die dezelfde vignetten voorgelegd kregen ter beoordeling. Door gebruik van de constant vergelijkende analysemethode zijn uit de onderzoeksgegevens categorieën gededuceerd waarbij gezocht is naar samenhang tussen en nieuwe bewijslasten die beoordelaars beschreven en benoemden.

Participanten

Voor deze studie werd doelgericht gezocht naar participanten (Babbie, 2010). Hiervoor werden medewerkers benaderd die ervaring hadden met het beoordelen van excellentieportfolio's van de hogeronderwijsinstelling. Zes vrouwen en drie mannen met verschillende achtergronden (mondhygiënist: 1; fysiotherapeut: 2; oefentherapeut Cesar: 2; verpleegkundige: 1; docent lichamenlijk opvoeding: 1; neerlandicus: 1; cultureel antropoloog: 1) participeerden. Zes hebben een masteropleiding, twee zijn PhD kandidaat, een volgt een masteropleiding en een heeft een bacheloropleiding. Hun leeftijd varieerde van 29 – 63 jaar (gemiddeld 52 jaar). De ervaring als excellentie-examinator varieerde van 0,5 tot 6 jaar (gemiddeld 3 jaar).

Hoe meer ervaringen van participanten in een bepaald domein gelijk zijn, hoe eerder verwacht kan worden dat verzadiging zal optreden van nieuwe categorieën die uit de data ontstaan. Uit onderzoek van Guest, Bunce en Johnson (2006) bleek dat in een steekproef van zes interviews al ongeveer 80% van alle categorieën werden gevonden. Met negen interviews werd verwacht dat voldoende gegevens konden worden verzameld om bruikbare interpretaties te kunnen doen.

Dataverzameling

Beoordelingsformulieren van 36 excellentieportfolio's uit de periode 2010 tot en met 2013 zijn verzameld en geanalyseerd en met negen beoordelaars zijn individuele interviews gehouden. Voor deelvraag 1, de relatie tussen bewijslasten en beoordeling zijn zowel beoordelingsformulieren als gegevens uit de interviews geanalyseerd; voor de deelvragen 2 en 3 zijn de interviewdata geanalyseerd.

In semigestructureerde interviews met behulp van vignetten en een vragenlijst is nagegaan welke keuzes beoordelaars maken. De volgorde van het aanbieden van

de vignetten was niet vaststaand, de vragen waren vignet-gerelateerd. Na twee pilotinterviews bleek het niet nodig om vignetten en vragenlijst aan te passen. Daarom zijn de gegevens van de pilotinterviews in de totale sample opgenomen. De interviews duurden een uur, zijn opgenomen en getranscribeerd. Alle interviews zijn in een periode van vier weken afgenomen, inclusief de pilotinterviews. Hiermee werd gepoogd om geen invloed van de eerste geïnterviewde op de volgende geïnterviewde te laten ontstaan wanneer zij elkaar mogelijk ontmoeten en het interview zouden bespreken. Het transcriberen van een interview gebeurde in de week van afname; data-analyse vond plaats nadat eerst alle gegevens waren verzameld.

De ontwikkelde instrumenten

Een vignet is een korte beschrijving van een situatie waarbij zowel verklarende als contextuele factoren in een realistisch scenario worden gegeven. Een vignettenonderzoek is een krachtig hulpmiddel in onderzoek naar de mening van respondenten (Atzmüller & Steiner, 2010). Respondenten kunnen zo authentiek mogelijk een oordeel vormen zodat de onderzoeker een beeld krijgt van overeenstemming (Alexander & Becker, 1978). De geanonimiseerde vignetten bestonden uit teksten die waren geselecteerd uit gearchiveerde excellentieportfolio's van de invalshoeken: (reflectie op) vakmanschap, innovatie en disseminatie en internationale oriëntatie. In verband met de duur van het interview werd gekozen voor drie vignetten met niet te lange teksten die voldoende informatie gaven om tot op zekere hoogte op zichzelf begrijpelijk te zijn zodanig dat een afgewogen oordeel gegeven kon worden. Hieronder volgt een kort fragment uit een van de vignetten.

“...Tijdens mijn opleiding heb ik ook een onderzoek uitgevoerd naar de noodzaak van gerichte zorg in Afrika. Hiervoor heb ik eerst een plan van aanpak gehad en een klein vooronderzoek uitgevoerd naar de geboden zorg in Afrika en de cultuur van het betreffende land. Voor de opleiding zijn wij drie maanden naar Afrika geweest, waar we een beginnend project hebben voortgezet. ...Aangezien er eerst totaal geen sprake was van gerichte zorg in het betreffende land heb ik een bijdrage kunnen leveren aan de overdracht van kennis en vaardigheden. ...Tijdens ons verblijf hebben we een blog bijgehouden met onze ervaringen en zijn we geïnterviewd voor een Nederlands tijdschrift. Zo heeft ons project nog meer invloed gehad en meer bekendheid gekregen.” (*invalshoek innovatie en disseminatie*)

Naast vignetten werd een semigestructureerd interviewformat gebruikt waarin vooraf vastgestelde vragen stonden, zoals: “Welke factoren wijzen op het overstijgen van het bachelorniveau?” Verder werden vragen opgenomen ter verdieping van het begrip van de argumentaties tijdens het beoordelingsproces, zoals “hoe vormt u zich een indruk van de student?”

Data-analyse

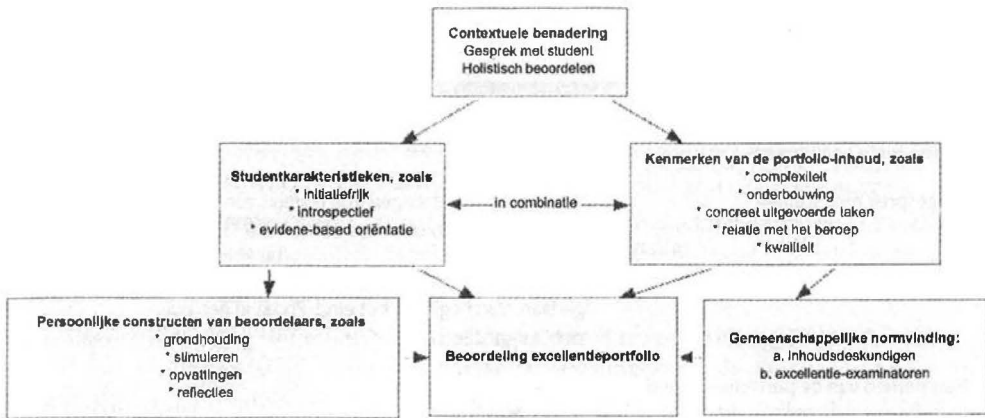
De teksten van de beoordelingsformulieren zijn per student doorgelezen voor een zo compleet mogelijk beeld. Vervolgens zijn de teksten op basis van de grounded theory benadering (cf. Maxwell, 2005) op inductieve wijze geanalyseerd: na de open codeerfase volgde een axiale codeerfase waarbij de codes zijn geordend en ingedeeld in subcategorieën en hoofdcategorieën. Op deze wijze zijn ook de interviewtranscripten geanalyseerd, gecodeerd en gegroepeerd. Daarna zijn de transcripten opnieuw gecodeerd op basis van cognitieve representaties van beoordelaars op de dimensies abstracte-concrete opmerkingen en positieve-negatieve evaluatie. Bij deze analyse was sprake van een meer deductieve databenadering (Babbie, 2010).

De axiale codes en bijbehorende fragmenten van beoordelingsformulieren en transcripten zijn besproken met een onafhankelijke collega. Op basis van dit overleg is aan een aantal fragmenten een andere code toegekend. Daarna zijn de categorieën van beide axiale codeerfasen vergeleken en zijn nogmaals naamgeving van de categorieën en de fragmenten die gecodeerd waren gecheckt op overeenkomsten tussen beide. Vervolgens zijn de categorieën en hun kenmerken geïntegreerd in één schema. In deze laatste stap, de selectieve codering, zijn de verschillende categorieën met elkaar verbonden. Hierin zijn de axiale codes van beoordelingsformulieren en transcripten in relatie tot elkaar gebracht. Ook dit schema is voorgelegd en besproken met de collega in een tweede overleg.

Om de kwaliteit van de analyse te vergroten is gebruik gemaakt van triangulatie (Kimchi, Polivaka & Stevenson, 1991; van Staa & Evers, 2010; Yin, 2009) door gebruik van verschillende bronnen (datatriangulatie), beoordelingsformulieren (documentanalyse) en interviews. Ook is samen met een collega gereflecteerd op de code-schema's (triangulatie van interpretaties) waarmee wordt gepoogd om subjectiviteit (researcher bias) zo veel mogelijk te voorkomen (Maxwell, 2005, Yin, 2010). Sandberg (1997) relateert dit door te stellen dat onderzoekers niet kunnen ontsnappen aan een eigen, subjectieve, interpretatie in een onderzoeksproces en spreekt over een interpretatief bewustzijn. Door dit te erkennen en er aandacht aan te schenken wordt het interpretatief bewustzijn ondersteund en een te sterke onderzoekersbias voorkomen.

Resultaten

Op basis van de uitkomsten van het analyseproces is een model ontwikkeld. In het model zijn de gevonden categorieën als thema's samengevoegd en zijn de onderlinge relaties aangegeven (zie figuur 1).



Figuur 1. Model van interacterende factoren bij een excellentieportfolio-beoordeling

Bij een het beoordelen van een excellentieportfolio blijkt dat beoordelaars contextueel beoordelen, zij geven aan holistisch te kijken (naar de totale prestatie) en de student te willen spreken, bijvoorbeeld voor aanvullende informatie. Zij beoordelen de portfolio-inhoud (wat hebben studenten gedaan en er over opgeschreven?) op basis van de beoordelingscriteria. De kenmerken van de portfolio-inhoud worden *in combinatie* met studentkarakteristieken (bijvoorbeeld initiatiefrijk zijn) beoordeeld. De portfolio-inhoud en de studentkarakteristieken vormen de bewijslasten. Verder wordt bij beoordelen gemeenschappelijke normvinding gebruikt en spelen persoonlijke constructen van beoordelaars een rol. De gevonden thema's zijn als factoren in het model opgenomen en geven de relatie aan van de factoren die interacteren bij excellentieportfolio-beoordeling. Hierna volgt een beschrijving van codes en categorieën die ten grondslag liggen aan de thema's in het model als resultaat van de analyse van de interviews en beoordelingsformulieren, zie tabel 1. In deze tabel wordt van elke factor een categorie met een citaat gegeven en is de frequentie opgenomen waarin de categorie ter sprake kwam of was opgeschreven.

Contextuele benadering. Hiermee wordt bedoeld dat een beoordeling gebaseerd wordt op meer dan alleen het portfolio, beoordelaars willen 'het overzicht hebben'. Zes beoordelaars zeggen beoordelen gemakkelijker te vinden wanneer zij met de student gesproken hebben.

Kenmerken van de portfolio-inhoud. Beoordeeld wordt wat aan concreet uitgevoerde taken is beschreven met gebruikmaking van beoordelingscriteria. De prestatie moet het reguliere bachelorniveau overstijgen wat gemeten kan worden met de niveaumeter (Coppoolse & Vroegindewij, 2010): "Tijdens de testen moet creatief

Tabel 1. Overzicht van categorieën en onderliggende codes met voorbeelden en frequenties

Categorie	Voorbeeld citaat	Frequentie
Contextuele benadering		
a. gesprek met student	"Dit soort beoordelingen doe ik nooit zonder de student ook gesproken te hebben" (Be. 1)	6
b. holistisch kijken	"En een presentatie houden over wat ze hebben gedaan. Van begin tot het eind. Zodat je een overzicht krijgt" (Be. 3)	6
Kenmerken van de portfolio-inhoud		
a. onderbouwing	"...waar haalt ze dat vandaan? Dat kan zij wel vinden maar dat zie ik hier niet in" (Be. 1)	9
b. kwaliteit van het geschrevene	"...dat vind ik heel helder geformuleerd" (Be. 6)	9
c. complexiteit van het leertraject	"...maar ik mis een beetje de complexiteit in dit verhaal, in het hele gebeuren" (Be. 1)	6 – 8,3%
d. samenhang willen zien	"...en dat stukje en dat stukje, dit zijn een beetje losse fragmenten" (Be. 2)	3
e. relatie met beroep	"...wat gaat dat betekenen voor de beroepspraktijk?" (Be. 9)	8
f. concreet uitgevoerde taken / grotere taakbelasting dan regulier	"Dus ik zie nog niet zo goed welke extra inspanning ze nou eigenlijk geleverd heeft" (Be. 5)	9
g. eisen voor een artikel / onderzoeksverslag	"...wat is het niveau van het artikel en voor wie heb je dit geschreven. Is het publiceerbaar en is het sterwaardig" (Be. 7)	3
Gemeenschappelijke normvinding	"en toen heb ik met de commissie getoetst, denk maar even mee bij de beoordeling" (Be. 8)	9 – 100%
Studentkarakteristieken		
a. introspectief / feedback gebruiken / leergierig / groei willen zien	"De ontwikkeling geeft ze goed aan. Naar de toekomst ook, als je die cyclus door wilt trekken" (Be. 2)	8 – 27,8%
b. initiatiefrijk / zelfstandig / doorzettingsvermogen / gedrevenheid	"...dat gaat heel erg over intrinsieke motivatie, over idealen hebben en je daar mee verbinden en daar ook verantwoordelijkheid willen nemen" (Be. 6)	9 – 66,8%
c. samenwerken	"ik denk ook dat het goed is om te kijken volde ik nu aan die internationale competenties want ik heb niet samengewerkt met andere beroepsbeoefenaren in een andere cultuur" (Be. 7)	1 – 16,7%

Categorie	Voorbeeld citaat	Frequentie
d. sociaal vaardig	"die kan je alles in handen leggen, sociaal vaardig, gaat het uitwerken..." (Be. 4)	1 – 27,8%
e. evidence-based oriëntatie / systematisch en doelgericht werken	"dat er weinig evidence voor is en dat ze daar wel een rol in wil gaan spelen" (Be. 6)	3 – 22,2%
Persoonlijke constructen		
a. positieve grondhouding: vertrouwen	"Begin met alles positief te zien. De student komt niet met zijn flauwekul. In principe" (Be. 1)	6
b. willen stimuleren	"...wil je je studenten stimuleren om te excelleren" (Be. 7)	8
c. opvattingen / aannames	"...dan is het lastig inschatten of dat boven bachelorniveau is. Ik denk dat het toch wel boven bachelorniveau is dat ze zich hierin verdiept heeft" (Be. 3)	4
d. overwegingen die tot reflectie leiden	"Dan zal je misschien meer tussen de regels doorlezen van wat er gedaan is. Waardoor je misschien toch met een ander oog kijkt. Niet zo strak. Ik weet niet of dat nou oké is of juist niet oké" (Be. 9)	6
e. aangeven van spanningsveld, het lastig vinden	"Zolang ik dat niet zie zou ik zeggen niet voldaan. Hoewel... in potentie is het wel overstijgend, dit is een lastige" (Be. 5)	8
f. Cognitieve representaties		
• concreet	"Ik zie het er niet in" (Be. 1)	9
• abstract	"Interessante studente" (Be. 2)	3
• evaluatie, positief of negatief	"Goed dat je dat zo ziet" (Be. 2)	9

Noot 1: De frequentie geeft het aantal interviews / beoordelingsformulieren waarin de categorie werd genoemd. Een percentage geeft de frequentie aan in de 36 beoordelingsformulieren en een los getal geeft het aantal van de negen interviews waarin dit besproken werd door de beoordelaar. Noot 2: Be. 1 = beoordelaar interview 1; etc.

worden omgegaan met problemen van de patiënt individueel. Niveaumeter 12 en 15" (beoordelingsformulier student 4). Ook wordt gelet op onderbouwing en of de relatie wordt gelegd met het beroep waarvoor wordt opgeleid (8 beoordelaars).

Studentkarakteristieken. Bij driekwart van de beoordelingsformulieren en bij alle interviews wordt gelet op eigenschappen en gedragingen van studenten, samengevat onder de categorie "studentkarakteristieken". Hierbij zijn de categorieën onderscheiden: introspectief; initiatiefrijk: "Zij liet hierbij een grote mate van pro-activiteit zien"

(beoordelingsformulier student 10); evidence-based oriëntatie; sociaal vaardig zijn, kunnen samenwerken: “Meerdere leeromgevingen gecreëerd met noodzaak tot eigen initiatief, besluitvorming, samenwerking” (beoordelingsformulier student 2).

Gemeenschappelijke normvinding. Bij elke excellentiebeoordeling is het voorwaarde om de excellentie-aanvraag door minimaal één deskundige te laten ondersteunen, bijvoorbeeld docenten van de bacheloropleiding of praktijkbegeleiders en heeft direct te maken met de geleverde prestatie. Gemeenschappelijke normvinding is in elk beoordelingsformulier met ‘voldaan’ ingevuld en genoemd door alle beoordelaars.

Persoonlijke constructen van beoordelaars. Hieronder wordt verstaan: opvattingen die zowel positief als negatief kunnen zijn ten aanzien van de beoordeling van een excellentieportfolio en reflecties erop. Beoordelaars stimuleren studenten het excellentietraject te volgen en benoemen hun eigen positieve grondhouding: “begin met alles positief te zien” (beoordelaar 1). Vier beoordelaars spreken hun opvatting of aanname uit op basis van het vignet-interview. De vignettenbeoordeling werd door acht beoordelaars als “lastig” ervaren. Beoordelaars benoemen hun overwegingen over het beoordelen van een door hen begeleide student: “Dan zal je misschien meer tussen de regels door lezen van wat er gedaan is. Waardoor je misschien toch met een ander oog kijkt. Niet zo strak. Ik weet niet of dat nou oké is of juist niet oké” (beoordelaar 9).

Twee interviewvragen betroffen de invloed van het begeleidingsproces op de beoordeling. Met deze data zijn de cognitieve representaties gecodeerd. De concrete representaties zijn gecodeerd onder de categorie “kenmerken van de portfolio-inhoud”. Alle beoordelaars gebruiken concrete representaties, bijvoorbeeld: “ik zie het er niet in” (beoordelaar 1). Drie gaven abstracte representaties: “als ik haar dan tegen kwam op de gang hadden we wel even een gesprekje ... ze kon een beetje in het zeurderige zitten. Maar dan zit ik helemaal mezelf een soort van bril op te zetten...” (beoordelaar 3). Beoordelaars ervaren het begeleidingsproces als aanvulling op hun inzicht in de prestatie van studenten, wat werd gecodeerd onder de categorie “contextuele beoordeling”.

Conclusies en discussie

Als resultaat van het onderzoek hoe beoordelaars bewijslasten van studenten beoordelen bij het verlenen van excellentiepunten kon een model worden ontwikkeld (zie figuur 1) van interacterende factoren bij het beoordelen van een excellentieportfolio. De factoren zijn de thema's van de categorieën van de beoordeelde bewijslasten, cognitieve representaties en de invloed van het begeleidingsproces, namelijk: context,

portfolio-inhoud, studentkarakteristieken, gemeenschappelijke normvinding en persoonlijke constructen. Behalve over studentkarakteristieken zijn over alle andere factoren studies beschreven in relatie tot portfoliobeoordelingen (Driessen & van Tartwijk, 2006; Tichelaar et al., 2005; van der Schaaf et al., 2005). Opvallend is het beoordelen van de portfolio-inhoud *in combinatie* met studentkarakteristieken. Een mogelijke verklaring hiervoor vormt wellicht het gebruikte referentiekader voor de ontwikkeling van het bestudeerde excellentieonderwijs: Renzulli's drie-ringen model (1978), waarin studenteigenschappen intelligentie, motivatie en creativiteit centraal staan. De beoordelingscriteria van de invalshoek "gedrevenheid" kennen de studenteigenschappen: intrinsieke motivatie, interesse, doorzettingsvermogen, werklust en pro-activiteit, een visie hebben (Excellentiecommissie Faculteit Gezondheidszorg, 2013). Beoordelaars lijken deze eigenschappen als beoordelingscriteria te gebruiken om een volledig beeld te krijgen van een student. Uit onderzoek onder zorgprofessionals naar wat excellente zorgprofessionals kenmerkt kwam een conceptueel profiel met acht domeinen van karakteristieken: initiatiefrijk, innovatief, introspectief, evidence-based oriëntatie, kennis, coöperatief, communicatief en kijken in een breed perspectief (Paans, Wijkamp, Wiltens & Wolfensberger, 2013). Het meewegen van studentkarakteristieken lijkt daardoor een goede voorbereiding op wat van excellente afgestudeerden wordt verwacht, eenmaal werkzaam in de praktijk. Omdat de beoordelingscriteria beschikbaar zijn voor studenten zijn zij geïnformeerd over de studentkarakteristieken.

Er lijkt sprake te zijn van een hermeneutische wijze van beoordelen zoals aanbevolen door Moss (1994) en Tigelaar et al. (2005): beoordelaars zoeken bewijsmateriaal wat wordt geïnterpreteerd in relatie tot eerder gevonden bewijs (zoals mondeling verkregen informatie) en in relatie tot beoordelingscriteria. Deze benadering komt overeen met wat als een zorgvuldige, evenwichtige expertbeoordeling wordt gezien (Cook et al., 2006; Van der Schaaf et al., 2005; Schuwirth & Van der Vleuten, 2006). Omdat het excellentieportfolio samengesteld materiaal bevat vanuit verschillende bronnen maakt dat gesproken kan worden van constructvaliditeit (Cook & Beckman, 2006). Door de verschillende bronnen is betrouwbaar en objectief beoordelen lastig omdat het materiaal in samenhang moet worden geïnterpreteerd.

Persoonlijke constructen spelen een rol en beoordelaars zijn zich daarvan bewust ('biased subjectivity'); dit is kenmerkend voor professioneel werken (Sandberg, 1997). In dit onderzoek blijkt uit concrete representaties dat alle beoordelaars hun oordeel voor een belangrijk deel baseren op de portfolio-inhoud.

Beperkingen van het onderzoek en toekomstig onderzoek.

De resultaten uit deze studie zijn gerelateerd aan de beoordelingspraktijk van een hogeschoolfaculteit voor gezondheidszorg. Of de factoren 'de kenmerken van de portfolio-inhoud' en 'studentkarakteristieken' die in combinatie worden beoordeeld representatief zijn voor overige hogere beroepsopleidingen of discipline-afhankelijk zijn is niet afleidbaar. De portfoliobeoordelingen zijn niet vergeleken met andere beoordelingen van de studenten, dit kan een onderzoeksvraag zijn voor aanvullende studies. Door het kleinschalige karakter van deze studie kon gedetailleerd praktijkkennis van beoordelaars worden beschreven. De hypothesen die hieruit naar voren komen (i.e. de vijf factoren) zullen in een vervolgonderzoek getoetst moeten worden. Bijvoorbeeld door een vragenlijst te ontwikkelen waarmee de invloed van de vijf factoren op portfoliobeoordelingen onder een bredere populatie bestudeerd kunnen worden. Andere suggesties voor onderzoek zijn of een dergelijke beoordelingspraktijk op de ontwikkeling van excellentiestudenten van invloed is en of specifieke studentkarakteristieken ook buiten het excellentieonderwijs studenten kunnen ondersteunen om zo een goede beroepsbeoefenaar te worden.

Implicaties voor de praktijk

Wanneer een procedurebeschrijving van het portfolioproces wordt toegevoegd op basis waarvan de beoordeling tot stand is gekomen bevordert dit *betrouwbaarheid* als interpretatief bewustzijn; daardoor wordt niet alleen gefocussed op het eindresultaat (Sandberg, 1997; Schuwirth en van der Vleuten, 2006; Tigelaar et al., 2005). Check beoordelaarsinterpretaties van een portfoliobeoordeling bij de student (member-check) om betrouwbaarheid van interpretaties te toetsen, bijvoorbeeld in een tripartite meeting (Tigelaar et al., 2005).

Vul het profiel van de excellentiestudent aan met de karakteristieken die uit onderzoek onder zorgprofessionals zijn verkregen (Paans et al., 2013). Karakteristieken kunnen als referentiekader voor persoonlijke groei worden gebruikt (Paans et al., 2013; Van Eijl, Wolfensberger & Pilot, 2006) bijvoorbeeld middels 360-graden-feedback. Deze assessmentvorm werkt op basis van verzamelende vragenlijsten van professionals waarmee wordt samengewerkt en een zelfevaluatie. In medische opleidingen wordt hier naar tevredenheid mee gewerkt (Davies & Archer, 2005).

De in deze studie beschreven praktijkkennis betreft een klein aantal respondenten. Totdat vervolgonderzoek gedaan is onder een bredere populatie kunnen de interacterende factoren wellicht een hulpmiddel zijn bij normvinding van de beoordelingen. Tot slot wordt aanbevolen startende excellentie-examinatoren met behulp van authentieke vignetten te laten oefenen met beoordelen onder leiding van ervaren beoordelaars.

Referenties

- Alexander, C. S. & Becker, H. J. (1978). The Use of Vignettes in Survey Research. *Public Opinion Quarterly* 42(1): 93-104.
- Atzmüller, C. & Steiner, P. M. (2010). Experimental Vignette Studies in Survey Research. *Methodology* 6(3): 128-138.
- Babbie, E. (2010). *The Practice of Social Research*. Belmont, USA: Wadsworth.
- Cook, D. A. & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine*, 119, 166.e7-166.e16.
- Coppoolse, R. & Vroegindewij, D. (2010). *75 Modellen van het onderwijs*. Houten: Noordhoff Uitgevers.
- Davies, H. & Archer, J. (2005). Multi source feedback: development and practical aspects. *The Clinical Teacher*, 2(2), 77-81.
- Driessen, E., & Van Tartwijk, J. (2006). Toetsen met portfolio's; in: H. van Berkel en A. Bax, Editors; *Toetsen in het hoger onderwijs (2^e druk)*. Houten: Bohn, Stafleu van Loghem.
- Excellentiecommissie Faculteit Gezondheidszorg (2013). *Kader, beoordelingsprocedure en registratie van sterren. Uitwerking Faculteit Gezondheidszorg*. Hogeschool Utrecht. Interne publicatie.
- Flierman, T. (2011). *Nadere uitwerking Sterrenstelsel. Handreiking voor studenten, docenten en beoordelingscommissies*. Utrecht: Hogeschool Utrecht. Interne publicatie.
- Gregori-Giralt, E., & Menendez-Varela, J. L. (2015). Validity of the learning portfolio: analysis of a portfolio proposal for the university. *Instructional Science*, 43, 1-17.
- Guest, G., Bunce, A. & Johnson, L. (2006) How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, 18; 59.
- Kimchi, J., Polivka, B. & Stevenson, J. S. (1991). Triangulation: Operational Definitions. *Nursing Research*, 40(6), 364-366.
- Maxwell, J. A. (2005). *Qualitative Research Design. An Interactive Approach (second edition)*. Thousand Oaks: Sage.
- Moniz, T., Arntfield, S., Miller, K., Lingard, L., Watling, C., & Regehr, G. (2015). Considerations in the use of reflective writing for student assessment: issues of reliability and validity. *Medical Education*, 49, 901-908.
- Moss, P. A. (1994). Can There Be Validity Without Reliability? *Educational Research*, 23 (2), 5-12.
- Paans, W., Wijkamp, I., Wiltens, E. & Wolfensberger, M. V. (2013). What constitutes an excellent allied health care professional? A multidisciplinary focus group study. *Journal of Multidisciplinary Healthcare*, 6, 347-356.
- Qvortrupa, A., & Keiding, T. B. (2015). Portfolio assessment: production and reduction of complexity. *Assessment & Evaluation in Higher Education*, 40, 407-419.

- Reis, S. M., & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning and Individual Differences*, 20, 308–317.
- Renzulli, J. S. (1978). What Makes Giftedness? Reexamining a Definition. *Phi Delta Kappan* 60 (3), 180 - 184.
- Sandberg, J. (1997). Are Phenomenographic Results Reliable? *Higher Education Research & Development*, 16(2), 203-212.
- Scager, K., Akkerman, S. F., Pilot, A., & Wubbels, T. (2014). Challenging high-ability students. *Studies in Higher Education*, 39, 659-679.
- Schuwirth, L. W. T. & van der Vleuten, C. P. M. (2006). AI plea for new psychometric models in educational assessment. *Medical Education*, 40, 296-300.
- Tigelaar, D. E. H., Dolmans, D. H. J. M., Wolfhagen, I. H. A. P. & van der Vleuten, C. P. M. (2005). Quality issues in judging portfolios: Implications for organizing teaching portfolio assessment procedures. *Studies in Higher Education*, 30(5), 595–610.
- van Eijl, P. V., Pilot, A., Wolfensberger, M. V. C., & Schreve-Brinkman, E. J. (2010). Talentontwikkeling met honoursprogramma's. *Tijdschrift voor Hoger Onderwijs*, 28(4), 182-198 (182).
- van der Schaaf, M., Stokking, K. & Verloop, N. (2005). De invloed van cognitieve representaties van beoordelaars op hun beoordeling van docentportfolio's. *Pedagogische Studiën* (82), 7-26.
- van Staa, A. & Evers, J. (2010). 'Thick analysis': Strategie om de kwaliteit van kwalitatieve data-analyse te verhogen. *Kwalon* 43(1), 5-12.
- van Tartwijk, J., Driessen, E., van der Vleuten, C. & Stokking, K. (2007). Factors Influencing the Successful Introduction of Portfolio's. *Quality in Higher Education* 13(1) 69-79.
- van Tartwijk, J., van Rijswijk, M., Tuithof, H. & Driessen, E.W. (2008). Using an analogy in the introduction of a portfolio. *Teaching and Teacher Education*, 24, 927–938.
- Wolfensberger, M. V. C. (2015). *Talent development in European Higher Education*. Dordrecht: Springer.
- Yin, R. K. (2009). *Case Study Research. Design and methods (fourth edition)*. Washington DC: Sage.