



Authentieke toetsing, de beroepspraktijk in het vizier

Praktisch artikel

Dit artikel is het vijftiengste in een serie praktische artikelen over onderwijsinnovatie. Deze serie heeft de bedoeling om mensen die werkzaam zijn in het hoger onderwijs handreikingen en aandachtspunten te bieden voor eigen initiatieven in onderwijsinnovatie. De onderwerpen van deze reeks kunnen uiteenlopen, maar zullen altijd gaan over 'het maken van onderwijs' en dus over toepassingen van onderwijskundige en onderwijs-technologische inzichten in het dagelijks werk van de docent, het onderwijsteam of de studierichtingsleider.

Auteurs

Judith Gulikers, Theo Bastiaens, Paul Kirschner

Drs. Judith T.M. Gulikers is als onderzoeker in opleiding (oio) werkzaam bij het Onderwijstechnologisch expertisecentrum (OTEC) van de Open Universiteit Nederland.

Dr. Theo J. Bastiaens en prof. dr. Paul A. Kirschner zijn verbonden aan het OTEC van de Open Universiteit Nederland.

Inhoud

- _ Inleiding
- _ Verandering in de onderwijscultuur
- _ Toetsen in competentiegericht onderwijs
- _ Wat is authentiek toetsen?
- _ Authenticiteit is subjectief
- _ Vijf dimensies van authenticiteit
- _ De plaatst van authentieke toetsvormen in het huidige onderwijs
- _ Conclusie

Box 1: De piramide van Miller

Box 2: Algemeen kader voor authentiek toetsen

Box 3: De vijf toetsdimensies voor authenticiteit

Box 4: Het vijf-dimensionaal model (5DM) voor authentiek toetsen

Inleiding

In de jaren '70 van de vorige eeuw waren veel werkgevers niet tevreden over het niveau van de beginnend beroepsbeoefenaren. Die waren weliswaar theoretisch goed onderlegd, maar wisten in de praktijk vaak niet hoe ze die kennis moesten toepassen en gebruiken. De werkgevers gaven hiermee een signaal af richting het onderwijs, met name het beroepsonderwijs, dat naar hun idee te algemeen en theoretisch was. Het onderwijs richtte zich in die tijd met name op het verwerven van kennis en (basale) vaardigheden waarvan de relatie met de beroepspraktijk vaak niet

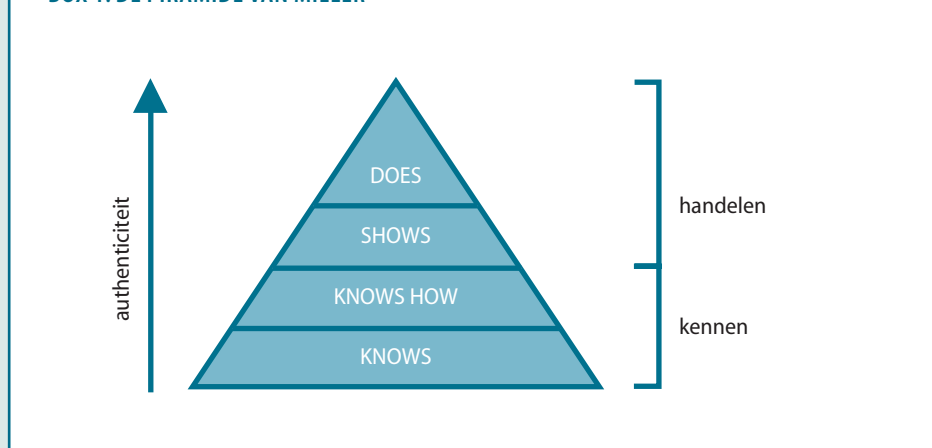
duidelijk was. Als reactie op deze kritiek zijn er in het (hoger) onderwijs de laatste decennia verschillende innovatieve thema's aan bod gekomen: duaal onderwijs, leren op de werkplek, levenslang leren, competentiegericht onderwijs en het erkennen van elders verworven competenties. Waar het in deze onderwijsvormen op neer komt, is dat er steeds meer nadruk is komen te liggen op het integreren van de beroepspraktijk in opleidingen om zo tegemoet te komen aan de klachten van het werkveld, én om studenten beter voor te bereiden op de praktijk. Ook van studenten komen de laatste jaren steeds meer vragen over de relevantie van de dingen die ze moeten bestuderen voor hun toekomstig functioneren. Studenten zijn meer gemotiveerd te studeren wanneer ze het nut van het studiemateriaal zien. Dit nut wordt vaak afgemeten aan de link met de beroepspraktijk. In dit kader gaat deze bijdrage over authentiek toetsen.

Authentieke toetsen zijn bedoeld om te evalueren hoe studenten kunnen functioneren als beginnend beroepsbeoefenaar door de toetsituatie (in grote mate) te laten overeenkomen met de toekomstige beroepspraktijk. In dit artikel wordt uitgebreid ingegaan op de vraag wat authentiek toetsen eigenlijk is en waarom authentiek toetsen belangrijk is in competentiegericht onderwijs. Vervolgens wordt een model beschreven dat vijf toetskenmerken beschrijft die de authenticiteit van een toets bepalen en dat praktische richtlijnen geeft voor het ontwerpen en evalueren van toetsen met verschillende soorten en niveaus van authenticiteit.

Verandering in de onderwijscultuur

Het werkveld klaagt steeds vaker dat het onderwijs studenten niet goed voorbereidt op de praktijk. Deze klachten zijn het gevolg van de veranderingen in de maatschappij in de afgelopen jaren waarin andere eisen aan werknemers worden gesteld. Om goed te kunnen functioneren in de praktijk moeten werknemers tegenwoordig kritisch kunnen denken en kennis kunnen toepassen in nieuwe situaties, om kunnen gaan met snel veranderende technologieën en communicatiemogelijkheden, goed kunnen samenwerken en voortdurend blijven leren. Om aan te sluiten bij deze nieuwe behoeften, moet het onderwijs zijn doelen aanpassen. Waar vroeger de nadruk lag op het opdoen van veel feitenkennis en routinematige vaardigheden, gaat het nu om het ontwikkelen van beroepscompetenties.

BOX 1: DE PIRAMIDE VAN MILLER



Birenbaum (2003) beargumenteert dat voor het bereiken van de onderwijsdoelen, ongeacht welk doel dit is, het noodzakelijk is dat instructie (I), leren (L) en assessment (A) op elkaar zijn afgestemd. De ILA-cultuur van het traditionele onderwijs werd gekenmerkt door: (1) instructie gericht op kennisoverdracht; (2) uit het hoofd leren; en (3) gestandaardiseerde toetsing. Competentiegericht onderwijs, waar veel hogescholen en universiteiten op overschakelen, wordt gekenmerkt door: (1) instructie gericht op het leren van studenten, (2) leren gericht op actieve kennisconstructie; en (3) gecontextualiseerde performance assessment, gericht op het evalueren van beroepscompetent gedrag. Bij het invoeren van competentiegericht onderwijs begint men vaak aan de instructiekant, terwijl de kant van de toetsing vaak tot het laatste blijft liggen. Birenbaums ILA-cultuur geeft echter aan dat een curriculum dat gebaseerd is op competentiegerichte instructie en leren, maar enkel getoetst wordt met gestandaardiseerde meerkeuzetoetsen, niet tot de gewenste

leerresultaten zal leiden. Onderzoek (Segers, Dierick, Dochy, 2001) ondersteunt dit en laat zien dat wanneer studenten ervaren dat het onderwijs (de instructie) en de toetsing twee verschillende boodschappen afgeven over wat er geleerd moet worden, dit een negatieve invloed heeft op de leerprestaties van studenten.

Toetsen in competentiegericht onderwijs

Competentiegericht onderwijs richt zich niet op het opdoen van losse kennis en vaardigheden maar op het integreren van kennis, vaardigheden en houdingen en het kunnen toepassen hiervan in beroepsrelevante situaties. Dit vraagt om een nieuwe manier van beoordelen. Segers, Dochy en Cascaller (2003) beschrijven verschillende soorten nieuwe toetsen, die gericht zijn op het toetsen van beroepscompetenties. Zij beargumenteren dat één van de belangrijkste kenmerken van deze nieuwe toetsen is dat studenten geconfronteerd worden met realistische situaties waarin zij hun kennis moeten toepassen op realistische problemen.

De beroepspraktijk moet zichtbaar worden in de toets, in tegenstelling tot wat vroeger vaak het geval was. Met andere woorden: de toetsing moet meer authentiek worden.

Er zijn twee belangrijke redenen voor authentiek toetsen: construct validiteit en consequentiële validiteit. Construct validiteit heeft te maken met de vraag of een toets meet wat deze bedoelt te meten. Miller (1990) beargumenteert dat wanneer een toets is bedoeld om beroepscompetenties te meten, het authentiek maken van de toets cruciaal is. Miller beschrijft beroepsbekwaamheid in vier lagen. Dit is schematisch weergegeven in de piramide van Miller (box 1) waarin onderliggende lagen het fundament vormen voor de laag erboven.

De onderste laag wordt gevormd door de kennis die een student nodig heeft om zijn/haar toekomstige beroep uit te voeren. Vaak wordt gedacht dat kennis in competentiegericht onderwijs niet meer zo belangrijk is. Millers piramide laat echter zien dat een brede en gestructureerde kennisbasis het fundament is van een competentie. Kennis is dus wel degelijk belangrijk, maar komt alleen in een ander daglicht te staan: ze wordt op een andere manier aangesproken en aangewend. Het kennisniveau (laag 1) kan bijvoorbeeld getoetst worden met een schriftelijke multiple-choice toets.

In de volgende laag gaat het erom dat een student weet hoe hij/zij die kennis moet gebruiken bij het oplossen van probleemtaken. Een rechtenstudent moet bijvoorbeeld weten hoe hij zijn kennis moet gebruiken bij het opstellen van een betoog. Dit soort kennis kan bijvoorbeeld beoordeeld worden met schriftelijke casussen.

In de derde laag moet de student laten zien dat hij/zij kan handelen in een gesimuleerde situatie. Het gaat hierbij om kennis én handelen. Een geneeskundestudent moet bijvoorbeeld een diagnose kunnen stellen op basis van een onderzoek bij een simulatiepatiënt. Miller spreekt hier ook wel van 'performance assessment in vitro'.

De bovenste laag heeft de student bereikt wanneer hij/zij in de opleiding verworven kennis, vaardigheden en houdingen kan toepassen in een echte complexe praktijksituatie op de werkvloer. Beoordeling op does-niveau moet dan ook plaatsvinden in de beroepspraktijk, oftewel 'performance assessment in vivo'.

Afhankelijk van het doel van de toetsing (wat wil men toetsen) wordt het belang van authenticiteit van de toets bepaald. Echter, in veel onderwijssituaties blijft men steken bij het beoordelen van laag 1, 2 en eventueel 3. Om daadwerkelijk competentiegericht te toetsen en studenten inderdaad beter voor te bereiden op de beroepspraktijk, zal de nadruk meer gelegd moeten worden op laag 3 en 4. De lagen laten de beroepspraktijk sterk naar voren komen waardoor studenten gestimuleerd worden de competenties en denkprocessen te gebruiken die een professional in deze beroepssituatie zou gebruiken. Met andere woorden: voor het beoordelen van competenties is authentiek toetsen cruciaal.

Consequentiële validiteit van een toets beschrijft de gewenste en ongewenste effecten van een toets op het leren van studenten. De manier van toetsen blijkt één van de belangrijkste sturende factoren in het leergedrag van studenten: voor een multiple-choice toets studeert een student anders dan voor een mondeling tentamen.

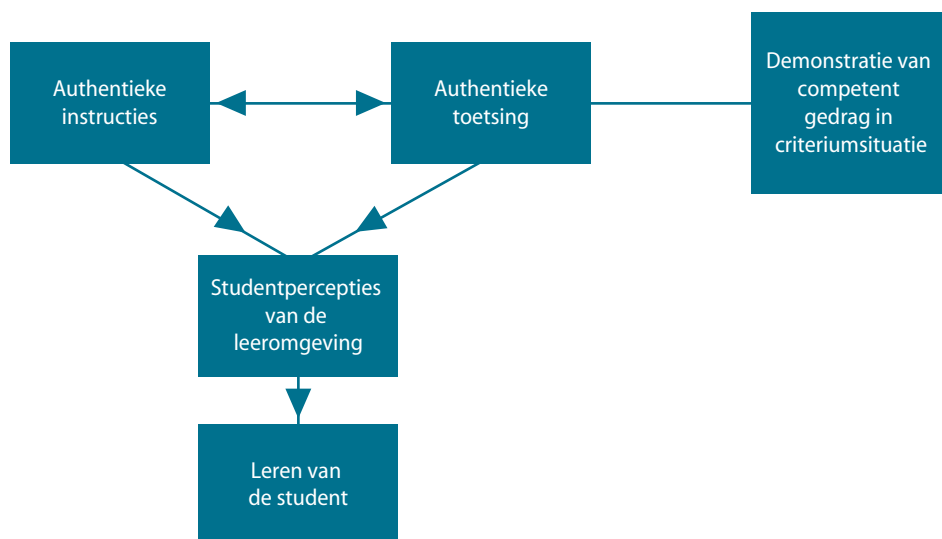
Van authentieke toetsen wordt verwacht dat ze het leren van studenten op twee manieren positief beïnvloeden. Ten eerste wordt verwacht dat authentieke toetsen die gericht zijn op het toetsen van beroepscompetenties, studenten ook daadwerkelijk aansporen om deze beroepscompetenties te ontwikkelen. Ten tweede zal een authentieke toets de motivatie om te leren verhogen, omdat studenten meteen de relevantie voor hun latere beroep kunnen zien en daarmee het nut van de toets snel ontdekken.

Wat is authentiek toetsen?

In het kader van het integreren van de beroepspraktijk in het onderwijs hebben opleidingen de laatste jaren nieuwe toetsen ontwikkeld onder de naam van authentiek toetsen. Cummings en Maxwell (1999) laten echter zien dat veel van deze toetsen in feite bestaande toetsen zijn die worden 'opgeleukt' met realistische elementen, zonder te weten welke elementen de toets eigenlijk meer realistisch maken. Het gevaar daarvan is dat deze 'authentieke' toetsen het leren van studenten juist negatief beïnvloeden omdat de authenticiteit te gekunsteld is, wat de toets eerder onbegrijpelijker dan realistischer maakt. Er bestaan veel verschillende ideeën over wat authenticiteit eigenlijk is en er zijn nog geen duidelijke richtlijnen voor het ontwerpen van authentieke toetsen.

In deze bijdrage zullen we daarom eerst het begrip authentiek toetsen duidelijk definiëren en vervolgens een model beschrijven dat meer inzicht geeft in wat de realistische elementen van authentieke toetsen zijn. Dit model geeft bovendien richtlijnen voor het ontwikkelen van authentieke toetsen.

BOX 2: ALGEMEEN KADER VOOR AUTHENTIEK TOETSEN



Messick (1994) stelt een vraag die cruciaal is voor een goede beschrijving van het begrip authenticiteit, namelijk: authentiek ten opzichte van wat? Hiermee geeft hij aan dat de authenticiteit van iets alleen beschreven kan worden in relatie tot iets anders. Zo kan een toets authentiek zijn ten opzichte van het onderwijs, maar niet authentiek ten opzichte van de werkelijkheid omdat het onderwijs niet is afgestemd op die werkelijkheid. Dit artikel definieert de authenticiteit van een toets ten opzicht van de beroepswerkelijkheid. Authentieke toetsen nemen de beroepspraktijk als uitgangspunt. Studenten moeten geconfronteerd worden met situaties uit de beroepspraktijk. Hier moet echter een kanttekening bij geplaatst worden. Van een eerstejaars student kan niet verwacht worden dat hij/zij

taken van een afgestudeerde of professional kan uitvoeren. Daarom zal een toets gebaseerd moeten zijn op situaties uit de beroepspraktijk, maar dan op het niveau van de student. Hierbij kan bijvoorbeeld gedacht worden aan situaties die een student op een bepaald niveau tijdens een stage moet uitvoeren. Een dergelijke situatie, die als uitgangspunt dient voor het ontwikkelen van de authentieke toets, heet de criteriumsituatie.

Een volgende stap in het definiëren van authentieke toetsing, is te beschrijven wat er van een student verwacht wordt in de toets. Savery en Duffy (1995) beschrijven dat een authentieke toets van studenten dezelfde cognitieve verwerking – denkprocessen – moet vragen, als de criteriumsituatie waar deze toets op gebaseerd is.

Darling-Hammond en Snyder (2000) zeggen dat alleen kijken naar het denkproces niet genoeg is. De beroepspraktijk vraagt namelijk dat studenten hun kennis en vaardigheden integreren en toepassen op nieuwe situaties. Daarom moet een authentieke toets gericht zijn op denken én handelen. Dit zou in de praktijk betekenen dat een authentieke toets een student moet vragen te demonstreren wat hij/zij geleerd heeft. Competent handelen vraagt bovendien niet alleen cognitieve activiteiten als probleem oplossen, maar ook naar meta-cognitieve competenties als reflectie, en sociale activiteiten als communicatie en samenwerking. Opleidingen op alle niveaus besteden hier steeds meer aandacht aan. Studenten moeten in een portfolio of persoonlijk ontwikkelingsplan (POP) laten zien dat ze kunnen reflecteren op hun eigen handelen. Samenwerking en communicatie wordt uitgelokt door onderwijsvormen als projectonderwijs en probleemgestuurd onderwijs (PGO), en in het (afstands)onderwijs komt steeds meer nadruk te liggen op samenwerking via de computer. Het beoordelen van dit soort vaardigheden blijkt echter nog moeilijk. Authentieke toetsen moeten dit brede spectrum van competenties in ogenschouw nemen.

Nog een belangrijk punt bij het beschrijven van wat authentiek toetsen is, is het uitgangspunt dat authenticiteit een continuüm is (Newmann & Wehlage, 1993). Het is een misconceptie dat iets ofwel helemaal authentiek, ofwel helemaal niet authentiek is. Dit betekent dat een toets in meer of mindere mate authentiek kan zijn en dat er dus vele variaties van authentiek toetsen mogelijk zijn. Dit idee, en hoe dit er in de praktijk uit

kan zien, komt later in deze bijdrage uitgebreid aan bod.

De definitie van authentiek toetsen die hier gebruikt wordt, is dat een authentieke toets studenten vraagt dezelfde competenties, of combinatie van kennis, vaardigheden en houdingen, te demonstreren als in de criteriumsituatie in de beroepspraktijk. De mate van authenticiteit van een toets kan vervolgens beschreven worden door de mate van overeenkomst tussen de toetsituatie en de professionele situatie. Het vijf-dimensionaal model dat hierna beschreven wordt, geeft aan dat deze overeenkomst bepaald wordt door verschillende dimensies (toetskenmerken) die meer of minder op de praktijk lijken.

Authenticiteit is subjectief

Een bijkomstig probleem in de discussie over authentiek toetsen is dat authenticiteit subjectief is en afhankelijk is van de percepties van studenten. Onderzoek (bijvoorbeeld McDowell, 1995) heeft aangetoond dat niet de objectieve kenmerken van een toets bepalen hoe studenten leren, maar dat de manier waarop studenten de verwachtingen van een toets interpreteren bepaalt hoe en wat zij leren. Dat betekent dus dat het belangrijk is om te kijken of studenten een toets inderdaad zo interpreteren als de ontwerper van de toets dat bedoeld heeft. Als de docent poogt om competent handelen te beoordelen, maar studenten percipiëren de toets als een moment waarop de docent hen vraagt om te vertellen wat de theorie zegt over hoe je zou moeten handelen, dan zullen studenten voornamelijk de theorie uit hun hoofd gaan leren in plaats van zich richten op het oefenen van praktisch handelen. Bij authentiek toetsen is het belangrijk dat studenten vinden dat de toets

BOX 3: DE VIJF TOETSDIMENSIES VOOR AUTHENTICITEIT

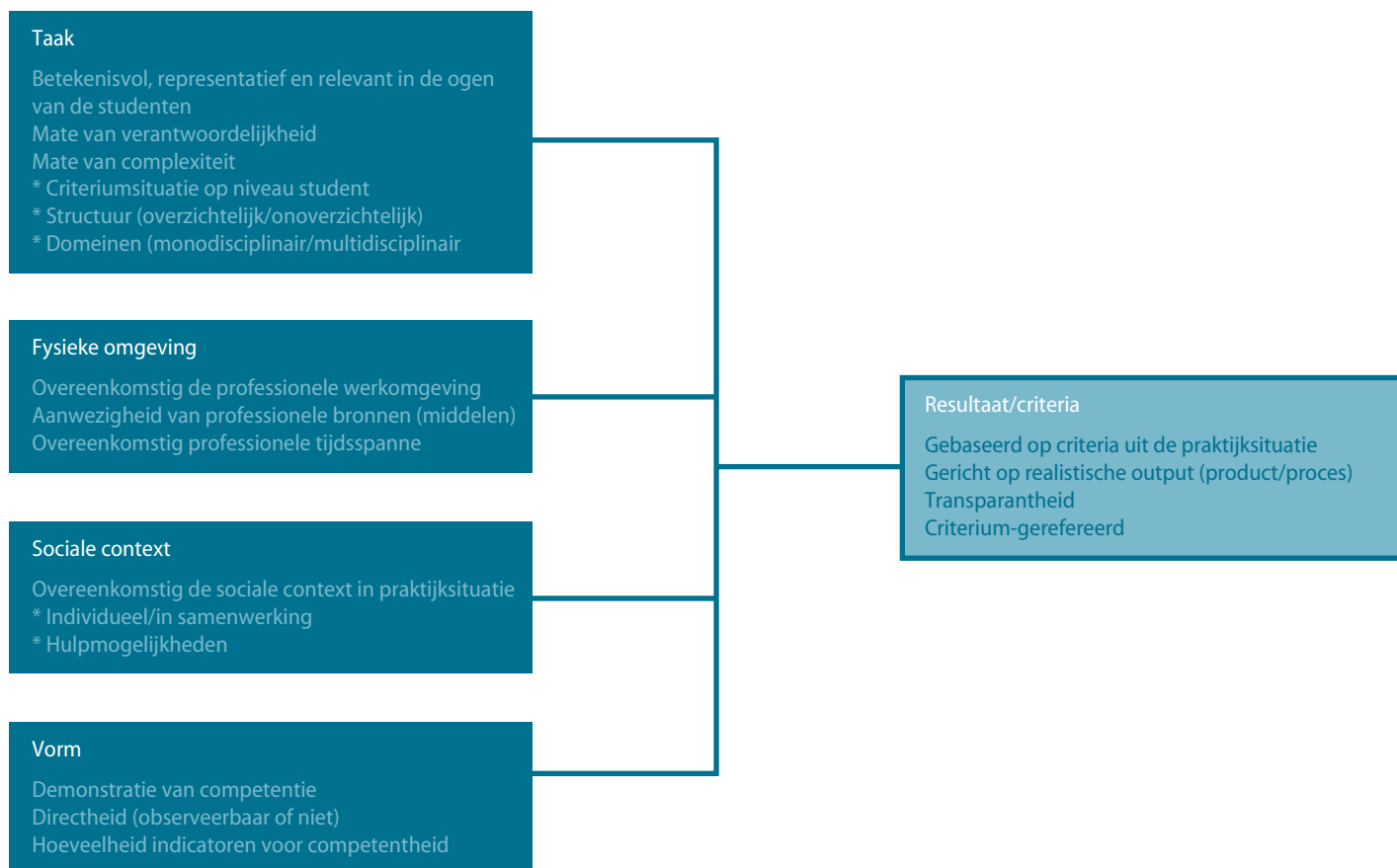
1. **Taak:** De toetsopdracht die de student moet uitvoeren tijdens de toets
2. **Omgeving:** de fysieke omgeving waarin de student de toetsopdracht moet uitvoeren
3. **Sociale context:** de interactiemogelijkheden in de toetsituatie
4. **Vorm:** de manier van toetsen (toetsmethode)
5. **Resultaat/criteria:** de output die uiteindelijk beoordeeld wordt (product/proces) en de punten waarop deze prestatie beoordeeld wordt

een goede afspiegeling is van de beroepspraktijk. Het feit dat authenticiteit subjectief is, betekent dat wat studenten authentiek vinden niet automatisch hetzelfde is als wat docenten of toetsontwikkelaars authentiek vinden. Dit komt omdat ieders idee over wat authentiek is – wat volgens jou een goede afspiegeling van de werkelijkheid is – kan veranderen door bijvoorbeeld meer praktijkervaring of leeftijd (Honebein, Duffy, & Fishman, 1993). Docenten hebben vaak meer praktijkervaring en kijken daardoor wellicht anders tegen de beroepspraktijk aan dan studenten. Docenten zijn echter wel meestal degenen die de toets maken en dat doen zij op basis van wat zij authentiek vinden. We kunnen ons afvragen of dit de beste weg is. Aangezien de percepties van studenten van een toets in grote mate bepalen hoe en wat ze leren, is het relevant om rekening te houden met deze percepties. De praktijk laat zien dat opleidingen worstelen met dit probleem. Dit komt tot uiting in het steeds meer betrekken van studenten bij het vormgeven van de toetsing. Zo zie je steeds vaker dat studenten alleen een algemene casusbeschrijving krijgen of alleen de uitkomsten (criteria; meestal in termen van gedragingen die studenten

moeten laten zien) van de toets en dat ze vervolgens zelf de gelegenheid krijgen te bepalen hoe ze deze toets gaan uitvoeren. Op langere termijn streven veel opleidingen naar flexibel onderwijs met het doel om iedere individuele student maatwerk te bieden; onderwijs dat past bij de voorkennis en ervaring, behoeften en wensen van de studenten. Hoe en wanneer verschillende soorten van authentieke toetsen aan bod komen in deze trajecten zal dan ook bepaald worden door de behoeften en vaardigheden van de student op een bepaald moment.

Het bovenstaande kan samengevat worden in een schema dat een kader schept voor authentieke toetsingspraktijken (box 2). Dit schema is gebaseerd op drie elementen: (1) om het leren van studenten positief te beïnvloeden moet een authentieke toets ingebed worden in een authentieke leeromgeving (ILA-cultuur), (2) authentieke toetsing vraagt studenten hun competenties te demonstreren in een situatie die lijkt op de beroepspraktijk; en (3) authenticiteit is subjectief, wat aangeeft dat het zeer belangrijk is om rekening te houden met studentpercepties bij het ontwikkelen van effectieve authentieke toetsen.

BOX 4: HET VIJF-DIMENSIONAAL MODEL (5DM) VOOR AUTHENTIEK TOETSEN



Vijf dimensies van authenticiteit

De piramide van Miller (1990) suggereert dat echt authentiek toetsen alleen plaats kan vinden in de beroepspraktijk. Het staat natuurlijk buiten kijf dat beoordelen in de praktijk inderdaad het meeste overeenkomt met de beroepspraktijk, maar er is meer mogelijk. In veel gevallen is beoordelen in de praktijk niet wenselijk of niet mogelijk.

Wanneer een geneeskundestudent nog nooit een injectiespuit heeft gezet, is het niet (ethisch) verantwoord om meteen in een huisartsenpraktijk te beoordelen of de student een griep prik kan zetten bij een echte patiënt. Maar om te beoordelen of een eerstejaars student rechten een goede argumentatie kan opbouwen voor een betoog, voldoet een contextrijke casustoets wellicht prima.

Het vijf-dimensionaal model (5DM) (Gulikers, Bastiaens, & Kirschner, 2004) dat hierna beschreven wordt, beargumenteert dat een toets op verschillende manieren meer of minder authentiek kan worden gemaakt door verschillende toetskenmerken wel of niet op de beroepswerkelijkheid te laten lijken. Hoe meer kenmerken op de beroepspraktijk lijken, hoe authentieker.

Hoe authentiek een toets moet zijn wordt met name bepaald door het doel van de toets, waarbij duidelijk is dat wanneer het om competentietoetsing gaat, meer authenticiteit noodzakelijk is. Hoe kan dit alles nu vertaald worden naar de praktijk? Ten eerste, een beroepssituatie moet als uitgangspunt worden genomen voor het ontwerpen van een toets.

Alleen door echt in de beroepspraktijk te kijken kun je beschrijven hoe relevante (kritische) beroepssituaties eruit zien, welke competenties de beroepsbeoefenaar in deze situatie gebruikt en op welke manier.

Vervolgens kunnen verschillende toetsen worden ontwikkeld die meer of minder lijken op deze situatie. Het vijf-dimensionaal model geeft richtlijnen voor het ontwikkelen van dit soort toetsen. Dit model beschrijft vijf toetsdimensies met bijbehorende kenmerken. Door deze onderdelen mee te nemen in de toetsontwikkeling kunnen verschillende authentieke toetsen ontwikkeld worden. Box 3 geeft een beschrijving van de vijf dimensies van authentieke toetsen en box 4 is een weergave van het volledige vijf-dimensionaal model met bijbehorende kenmerken.

Het vijf-dimensionaal model beschrijft geen definitieve richtlijnen voor authentieke toetsen. Met andere woorden, het model zegt niet dat een authentieke toets altijd complex is, of dat studenten in een authentieke toets altijd moeten samenwerken. Of een toets authentiek is, wordt namelijk bepaald door de mate van overeenkomst met de professionele situatie. Bijvoorbeeld: een toets in samenwerking uitvoeren is alleen authentiek als deze taak in de praktijk ook in een team wordt uitgevoerd. De hoofdgedachte achter het vijf-dimensionaal model is dat alle vijf

elementen in meer of mindere mate op de professionele situatie kunnen lijken en daarmee de toets meer of minder authentiek maken.

De plaats van authentieke toetsvormen in het huidige onderwijs

Wat is het nut van toetsen met verschillende maten van authenticiteit, en is maximale authenticiteit altijd het beste? Hierover is discussie mogelijk. Ten eerste is authenticiteit slechts één van de kwaliteitskenmerken van nieuwe vormen van toetsing (Baartman, Bastiaens, Kirschner, 2004) en kan meer authenticiteit ten koste gaan van andere kwaliteitsaspecten zoals bijvoorbeeld vergelijkbaarheid. In de onderwijspraktijk moet altijd een afweging gemaakt worden tussen verschillende kwaliteitscriteria, het doel van de toets en praktische mogelijkheden. Zo kan het zijn dat het uitvoeren van een bepaalde taak in de beroepssituatie een aantal uren duurt en dat deze uren bovendien verspreid zijn over een week. Praktische belemmeringen op school kunnen dit onmogelijk maken. Het vijf-dimensionaal model geeft nu mogelijkheden om de toets op een bepaald punt iets minder authentiek te maken, maar op andere punten wel authenticiteit te behouden. Bovendien kunnen traditionele kennistoetsen nuttig zijn als onderdeel van een competentiegericht assessmentprogramma, of als het doel van de toets vraagt om individuele, gestandaardiseerde toetsing. Echter, wanneer het gaat om het beoordelen van beroepscompetent gedrag, zal het verhogen van de authenticiteit steeds relevanter worden. Ten tweede is het mogelijk dat het belang van authentieke toetsing verandert naarmate een student vordert in een opleiding. Authenticiteit is afhankelijk van hoe een student tegen de beroepspraktijk aankijkt.

Een ouderejaars student, die al vaker met de echte beroepspraktijk in aanraking is gekomen, kan anders tegen de beroepspraktijk aankijken dan een eerstejaars student. Wellicht hechten ouderejaars studenten meer waarde aan een grotere overeenkomst met de beroepspraktijk, omdat die voor hen steeds dichterbij komt. Aan de andere kant is het mogelijk dat een bepaalde mate of vorm van authenticiteit juist voor jongere studenten zeer belangrijk is om hen de relevantie te laten zien van de dingen die ze moeten leren. Dit artikel valt binnen een project waarin onderzoek wordt gedaan naar hoe belangrijk studenten van verschillende studiejaren de vijf toetsdimensies van het vijf-dimensionaal model vinden, en hoe de authenticiteit van deze kenmerken hun leergedrag stuurt. Een voorbeeld uit het onderzoek laat zien dat een authentieke fysieke context niet zoveel invloed heeft op het de leeruitkomst en motivatie van (onervaren) studenten zolang de taak en het op te leveren resultaat maar authentiek zijn (Gulikers, Bastiaens, Martens, 2005). Een authentiekere fysieke context kan echter veel belangrijker zijn voor ouderejaars studenten die steeds meer in de echte beroepspraktijk moeten functioneren. Dit zou kunnen pleiten voor een toename van beoordeling in de beroepspraktijk naarmate een student in het studietraject komt.

Conclusie

Door authenticiteit op te splitsen in verschillende onderdelen en te benaderen als een dimensie, biedt het vijf-dimensionaal model veel praktische aanknopingspunten om de authenticiteit van toetsen te variëren tijdens een opleiding. Bovendien komt het model daarmee ook tegemoet aan praktische problemen rondom, of bezwaren tegen, authentiek toetsen.

Zoals hierboven beschreven, zou écht authentiek toetsen betekenen dat men studenten moet beoordelen tijdens hun taakuitvoering in de beroepspraktijk (bijvoorbeeld tijdens stages). De ontwikkeling van nieuwe manieren van toetsen, die bovendien (geheel) worden uitgevoerd in de praktijk, vraagt echter veel tijd, geld en moeite. Bovendien brengt het allerlei moeilijkheden met zich mee. Zo wordt de toets minder controleerbaar en gaan de toetsituaties meer verschillen per student. Daarnaast ligt de verantwoordelijkheid voor de beoordeling nog steeds bij de opleiding en niet bij de praktijk. Een ander probleem ontstaat er bij universitaire opleidingen of opleidingen met een breed en minder duidelijk gedefinieerd beroep. Hier is authentiek toetsen (wel of niet in de beroepspraktijk) sowieso moeilijker door de onduidelijkheid van het toekomstige beroep. Echter, relevante (kritische) beroepssituaties zijn voor iedere opleiding te definiëren, op basis waarvan bijvoorbeeld wel authentiekere assessmenttaken kunnen worden ontwikkeld. Het vijf-dimensionaal model laat namelijk zien dat de keuze niet ligt tussen totaal niet authentiek toetsen of volledig authentiek toetsen in de echte beroepspraktijk, maar dat er verschillende tussenvarianten zijn.

Het model geeft hiermee concrete handvatten voor het ontwikkelen van gedeeltelijk authentieke toetsen als gehele authentieke toetsing niet gewenst of praktisch niet haalbaar is. Bovendien maakt het model het mogelijk om de toetsing beter aan te passen aan wensen en behoeften van studenten. Dit past mooi in het kader van competentiegerichte en flexibele of geïndividualiseerde opleidingstrajecten, waar de laatste jaren steeds meer aandacht voor is.

Referenties

- Baartman, L.K.J., Bastiaens, T.J., & Kirschner, P.A. (2004). Requirements for Competency Assessment Programmes. Paper presented at the Onderwijs Research Dagen 2004, Utrecht.
- Birenbaum, M. (2003). New insights into learning and teaching and the implications for assessment. In M. Segers, F.J.R.C. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht: Kluwer Academic Publishers.
- Cummings, J. J., & Maxwell, G. S. (1999). Contextualising authentic assessment. *Assessment in Education: Principles, Policy & Practice*, 6 (2), 177-194.
- Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment in teaching in context. *Teaching and Teacher Education*, 16, 523-545.
- Gulikers, J., Bastiaens, Th., & Martens, R. (in press). The surplus value of an authentic learning environment. *Computers in Human Behavior*.
- Gulikers, J., Bastiaens, Th., & Kirschner, P. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52 (3), 67-85.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23 (2), 13-23.
- Newmann, F. M., & Wehlage, G. G. (1993). Five standards for authentic instruction. *Educational Leadership*, 50 (7), 8-12.
- Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. In B. G. Wilson (Ed.), *Constructivist learning environments: Case studies in instructional design*. Englewood Cliffs, NJ: Educational Technology Publications.
- Segers, M., Dierick, S., Dochy, F. (2001). Quality standards for new modes of assessment. An exploratory study of the consequential validity of the OverAll test. *European Journal of Psychology of Education*, 16(4), 569-586.
- Segers, M., Dochy, F., & Cascallar, E. (2003). *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht: Kluwer Academic Press.