**Recensie symposium Leren van toetsen**

Door Eveline de Beer, docent verpleegkunde aan het hbo en student master Onderwijswetenschappen

##### Inleiding

##### Op 1 juni 2018 heeft de 4e editie plaats van het symposium Leren van toetsen plaatsgevonden in vergadercentrum Domstad in Utrecht. Het symposium is georganiseerd door het Platform Leren van toetsen. De eerste keynote is Programmatisch toetsen bij het afstuderen van Dr. Liesbeth Baartman, Dr. Dominque Sluijsmans en Prof. Ir. Jan de Geus. De tweede keynote Innovatief toetsen is verzorgd door Prof. Dr. Jan Riezebos. Tijdens de twee workshoprondes heb ik de postermarkt onder leiding van Dr. Tamara van Schilt-Mol bijgewoond en de presentatie van Dr. Judith Gulikers over de toolkit formatief toetsen.

Dit symposium spreekt me aan om verschillende redenen. Ten eerste ben ik lid van de toetscommissie van een hogeschool. Vanuit die rol ben ik verantwoordelijk voor de borging van de kwaliteit van toetsing, waarbij de focus ligt bij summatieve toetsen. Echter, er is een trend te bespeuren richting formatief toetsen. Binnen onze opleiding voor verpleegkundigen staat formatieve toetsing in de kinderschoenen. Mijn drie vraagstellingen luiden: Wat kan de bijdrage van programmatisch toetsen zijn aan het verwerven van het eindniveau Bachelor of Nursing 2020? Welke rol kan de docent spelen bij de inzet van formatieve toetsing binnen onze opleiding tot verpleegkundige? En hoe kan de docent worden toegerust om formatieve toetsing in te zetten?

**Ochtendprogramma**

De dagvoorzitters Anne Potters (Hogeschool Rotterdam) en Martijn Leenknecht (Hogeschool Zeeland) verwelkomen ruim 250 deelnemers afkomstig van tenminste negen

verschillende hogescholen. Ze geven uitleg over de Hogescholenbingo. Bij binnenkomst heeft iedere deelnemer een bingokaart ontvangen. Als je iemand hebt gesproken en contactgegevens hebt uitgewisseld, mag je een kruisje zetten op je bingokaart. Wie het eerst de kaart vol heeft, ontvangt een prijs.

Keynote 1: Programmatisch toetsen bij het afstuderen

Keynote 1 bestaat uit drie delen. Dr. Liesbeth Baartman (hoofddocent bij het Lectoraat Beroepsonderwijs Hogeschool Utrecht), opent het eerste deel met de vraag: wat is programmatisch toetsen? Bij programmatisch toetsten kijk je naar het geheel van toetsen van je opleiding of een gedeelte van een opleiding. De combinatie van toetsen is meer dan de som der delen. Baartman legt uit: ‘Je ziet pas het geheel als je een stapje terugdoet en van een afstandje kijkt naar het geheel. Dan zie je pas of het totale plaatje valide en betrouwbaar is’. Van der Vleuten en Schuwirth (2005) zijn de grondleggers van een model voor assessment; een toetsprogramma, met als doel kwalitatieve en kwantitatieve informatie verkrijgen uit verschillende toetsbronnen. Bij programmatisch toetsen wordt het toetsprogramma als geheel en op doelbewuste wijze vormgeven (Van der Vleuten, Schut, & Heeneman, 2018). Molkenboer (2015, p. 37) verstaat onder een toetsprogramma “een bewuste en beargumenteerde combinatie van toetsvormen en toetsdoelen (formatief en summatief) passend bij de kwalificaties en de opbouw van de opleiding”. Van Berkel, Bax en Joosten-ten Brinke (2017) voegen hier aan toe dat het toetsprogramma moet passen bij de doelen, inhoud, opbouw en structuur van het curriculum.

Baartman vraagt zich af wat het toetsprogramma doet met het leren van de student. Ze heeft twee belangrijke boodschappen. Ten eerste: de inrichting van het toetsprogramma geeft signalen af over wat voor leren wordt gewaardeerd. Ten tweede: het gaat niet om transparantie, maar om betekenisgeving richting het beroep. Ze ligt beide boodschappen toe. We willen met de inrichting van het toetsprogramma signalen afgeven over wat er later in het beroep van een student wordt gevraagd. Het is daarom belangrijk dat een toetsprogramma een afspiegeling is van de beroepspraktijk. De boodschap die we afgeven met een toetsprogramma met veel summatieve toetsten is dat je eigenlijk geen fouten mag maken. We kunnen met een toetsprogramma dus een verkeerde boodschap afgeven, als we aanhangen dat leren alles te maken heeft met fouten mógen maken. Baartman zou graag zien dat je met toetsing een lange termijn leerproces in gang zet, waarin je fouten mág maken. De student heeft hiervoor wel iets nodig dat hij mee kan nemen, bijvoorbeeld feedback. Een toets moet informatie geven over de ontwikkeling op een competentie. De strekking van de tweede boodschap is dat een toetsprogramma betekenisgevend moet zijn richting het beroep waar voor wordt opgeleid. Het is belangrijk dat een student leert inschatten wat ‘goed werk’ betekent. Volledige transparant is niet nodig, studenten moeten zelf hun weg zoeken om betekenis te kunnen geven. Het formatief beoordelen is een continu cyclisch proces, dat doorloopt om de student te volgen in het betekenis geven over zijn ontwikkeling richting toekomstig beroep. Baartman eindigt met de oproep om je af te vragen of het toetsprogramma past bij het beroep waarvoor je opleidt en bij de boodschap die je wilt afgeven aan studenten.

In het tweede deel van keynote 1 gaat Dr. Dominique Sluijsmans (lector Professioneel Beoordelen bij Zuyd Hogeschool) in op de beslisfunctie van toetsen. Ze opent met een korte terugblik op het proces dat voorafging aan het Protocol Verbeteren en Verantwoorden van Afstuderen in het hbo 2.0 (Andriessen, Sluijsmans, Snel, & Jacobs, 2017). In 2011 ontstond rumoer over de diplomawaarde in het hbo, waarna de commissie Andriessen in 2012 kwam met het rapport Vreemde ogen dwingen. Daar zaten twee lijnen in (a) het ontwikkelen van toetsbekwaamheid door BKE/SKE en (b) het afstudeerdossier. Of de aanbevelingen uit dit rapport werken in de praktijk, hebben Sluijsmans en Andriessen tussen 2015 en 2017 onderzocht in het project Protocol afstuderen. Dit resulteerde in 2017 in een herziende versie van het Protocol Verbeteren en Verantwoorden van Afstuderen in het hbo 2.0. Hierin wordt een bijgesteld conceptueel model van het afstuderen in het hbo gepresenteerd, dat kan worden benut bij het verbeteren van een afstudeerprogramma (Andriessen et al., 2017). Het eerste protocol bevatte gesloten vragen, nu zijn het acht open vragen die een opleiding stimuleren bij het bepalen van de kwaliteit van het afstudeerdossier. Bij afstuderen gaat het om een set van bewijzen waar je beroepsbekwaamheid mee kan aantonen. Hoe kan je meer activiteiten organiseren om beroepsbekwaamheid aan te tonen? Ten eerste kan je kalibreren op basis van beoordelingscriteria. We kunnen veel leren van vreemde ogen, dat niet persé vier ogen hoeven te zijn. Vreemde ogen zijn, volgens Sluijsmans, zoveel ogen en zoveel type ogen dat nodig zijn om een oordeel te geven over beroepsbekwaamheid. Haar suggestie is om te streven naar meerdere beoordelaars en meerdere toetsen. Maar dan wel beoordelaars die beschikken over zowel inhoudelijke deskundigheid over het te beoordelen product, als over ontwerpdeskundigheid. Een samenhangend en consistent afstudeerprogramma ontwerpen vraagt bijzondere expertise. Als laatste noemt Sluijsmans het teamleren. In een collectief denk- en werkproces samenwerken met alle actoren, zoals examencommissie, assessor, docent, die een verantwoordelijkheid hebben binnen het afstudeerdossier, daar hoort de student ook bij. Sluijsmans eindigt met de aanbeveling om binnen de opleiding afspraken te maken over de manier waarop het protocol kan worden benut ter verantwoording van het eigen afstudeerdossier.

Prof. Ir. Jan de Geus (senior adviseur onderwijs Stenden Hogeschool Emmen) sluit de eerste keynote af met een *reality check* getiteld ‘…want tussen droom en daad staan wetten in de weg en praktische bezwaren’. Met wetten doelt De Geus op de Wet op het hoger onderwijs en wetenschappelijk onderzoek (WHW). Hij geeft aan het begin van zijn lezing aan dat hij vooral vragen heeft en geen oplossingen. Hij noemt een aantal artikelen uit de WHW en geeft aan waar het wringt met wat wij eigenlijk beogen met ons onderwijs en onze toetsing. Zo noemt hij dat in WHW artikel 7.3 lid 2 wordt vermeld dat een opleiding een samenhangend geheel is ‘van onderwijseenheden, gericht op de verwezenlijking van welomschreven doelstellingen op het gebied van kennis, inzicht en vaardigheden waarover degene die de opleiding voltooit, dient te beschikken’. Maar hoe verhoudt deze wetsbepaling zich tot de boodschap die we willen uitdragen, namelijk dat fouten maken mag? De wet is voortschrijvend, maar in zijn ogen sterk verouderd en daardoor niet meer passend bij de huidige onderwijspraktijk.

Daarna maakt De Geus de overstap naar de praktische bezwaren waar hij tegenaan loopt als hogeschool. Hij start met te vertellen dat wij uiteindelijk een student willen beoordelen op leeruitkomsten in de vorm van een beroepsproduct, een professionele houding en metacognitie zoals zelfregulatie. Een beroepsproduct beoordeel je op basis van vooraf opgestelde beoordelingscriteria. Dat zou een objectief oordeel moeten zijn, maar dat is niet zo. ‘Wij weten dat we in Nederland structureel lager beoordelen dan bijvoorbeeld in Duitsland’, stelt De Geus. Een ander praktisch bezwaar waar hij tegenaan loopt is dat het ontwikkelen van beoordelingscriteria voor elk beroepsproduct veel werk is. Stenden Hogeschool beoordeelt daarom niet zozeer het beroepsproduct zelf, maar vooral de manier waarop het eindproduct tot stand is gekomen. Voor hen is het proces naar vervaardiging van het beroepsproduct minstens zo belangrijk als het uiteindelijke beroepsproduct. Als laatste bezwaar noemt De Geus dat van studenten die in het buitenland stage lopen verwacht wordt dat een stagiaire volledig meewerkt. ‘Afstuderen doe je in je eigen tijd’ is het uitgangspunt van stagebegeleiders in een ander land. Bij het beoordelen van een professionele houding noemt De Geus als praktisch bezwaar het feit dat je als docent niet lijflijk aanwezig bent tijdens een stage. Hoe ga je die student beoordelen terwijl je er niet bij bent geweest? Over het beoordelen van metacognitie stelt De Geus dat een docent dit kan beoordelen, maar voor de betrouwbaarheid zijn vier ogen nodig. En het zorgen voor die vier ogen bij een beoordeling is vaak een praktisch probleem. Bovendien zijn voor holistische beoordelingen ervaren beoordelaars nodig. En ook hier blijkt dat waardering niet gelijk is tussen landen. Daarnaast is de navolgbaarheid en onderbouwing matig. Hierdoor is het lastig te achterhalen waarom een bepaald cijfer is gegeven. Holistisch beoordelen maakt onzeker in het kader van navolgbaarheid. ‘Dus dan ga je maar analytisch beoordelen zodat je beter kunt onderbouwen’, constateert De Geus.

De Geus sluit zijn betoog af met de opmerking dat je rekening moet houding met de populatie van studenten bij het ontwerpen van een toetsprogramma. Het ontwerpen van een goede, flexibele leeromgeving is belangrijk om je aan te kunnen aanpassen aan veranderingen in de praktijk en/of in het onderwijs. Stenden Hogeschool werkt om deze reden met een *sustainable learning environment*. *Sustainable development* verwijst naar het duurzaam ontwikkelen van competenties. Het inbedden van duurzaam ontwikkelen in curricula vraagt een multidisciplinaire benadering om innovatief, interactief, experimenteel en *real-world* leren te bewerkstelligen (Lozano, Merrill, Sammalisto, Ceulemans, & Lozano, 2017). Stenden Hogeschool geeft duurzaam ontwikkelen vorm in een aantal stappen (a) selectie van kritische beroepssituaties, (b) formuleren leeruitkomsten, (c) genereren alternatieve toetsprogramma’s op hoofdlijnen, (d), ontwerpen prototype onderwijs en toetsprogramma, (e) uitvoeren onderwijsprogramma en (f) onderzoeken effecten en aanpassen. De Geus eindigt met het statement dat curriculum- en toetsprogramma ontwikkeling teamwork is, ontwikkelen doe je samen!

Na de drie onderdelen van keynote 1 is er gelegenheid voor het publiek om vragen te stellen. Er wordt aan De Geus gevraagd of hij de ontwikkeling van een curriculum beschouwt als een iteratief proces. ‘Ja,’ bevestigt De Geus, ‘een prototype moet inderdaad telkens geëvalueerd en aangepast worden’.

Aan Sluijsmans wordt gevraagd of ze tips heeft hoe je kunt voorkomen dat holistisch beoordelen verzandt in eindeloos kalibreren? Sluijsmans zegt dat je werkt met meerdere ogen om betrouwbaar te beoordelen. Haar tip is om te kalibreren op niet teveel criteria, maar bijvoorbeeld op een uitgewerkt voorbeeld. Een andere optie is een *ranking* maken van gemaakte producten, en die rangschikking bespreken.

Aan Baartman wordt gevraagd of ze ervoor pleit om *rubrics* los te laten? ‘Nee, niet persé, mits de rubric holistisch is geformuleerd’, antwoordt Baartman. Het is heel zinvol om met studenten samen te kalibreren over wat kan wel en wat kan niet. Je geeft met *rubrics* de grenzen aan. Die grenzen moeten de studenten ontdekken. Als een *rubric* een leerfunctie heeft, dan is deze analytisch goed. ‘Op eindniveau is een *single point rubric* wellicht voldoende’, eindigt Baartman haar antwoord.

Er wordt gevraagd aan de sprekers of ze een voorstander zijn van het maken van koppels van beoordelaars? Sluijsmans antwoordt dat het goed is om te investeren in een poule van beoordelaars. Ze hoeven niet op één lijn te liggen. Het gaat om de discussie die ze onderling voeren. Een wisselende samenstelling van een poule of koppel is daarom wenselijk.

Het publiek wil weten of De Geus een advies heeft voor het duurzamer maken van de WHW? Lachend antwoordt De Geus: ‘Helemaal opnieuw beginnen met schrijven en ruimte laten voor ontwikkelingen’. De eerste versie van 1991 moet herschreven worden, want de terminologie in de verschillende artikelen is niet consistent.

Deze drieluik tijdens keynote 1 heeft een breed beeld gegeven van programmatisch toetsen in het onderwijs. Baartman legt de nadruk op het lange termijn leerproces dat we willen bewerkstelligen en de boodschap die we willen uitdragen met ons toetsprogramma. Sluijsmans bouwt daarop voort door vanuit het protocol het afstuderen te verantwoorden. De Geus maakt een mooie koppeling naar de praktijk. Dit alles heeft me laten nadenken over de boodschap die wij willen afgeven met ons toetsprogramma. Ons toetsprogramma op eindniveau bevat uitsluitende summatieve toetsen. Dat komt niet overeen met de visie op toetsing uit het onderwijs- en toetsbeleidsplan (AGZ, 2017). Hierin wordt nadrukkelijk gezegd dat toetsing enerzijds gericht is op het bepalen of een student de gestelde competenties heeft verworven en anderzijds op het geven van informatie aan studenten en docenten over waar studenten staan in hun leerproces. Een goed onderwerp van gesprek in mijn docententeam.

Workshopronde 2: Postermarkt toetsprogramma’s

Op verschillende hogescholen in Nederland worden toetsprogramma’s bedacht en uitgeprobeerd. Onder leiding van Dr. Tamara van Schilt-Mol (associate lector Toeten en Beoordelen kenniscentrum Kwaliteit van Leren, Hogeschool Arnhem en Nijmegen) deelt een aantal opleidingen hun ervaringen met programmatisch toetsen door een posterpresentatie. In de korte inleiding vertelt Van Schilt-Mol dat een toetsprogramma niet per definitie betekent dat er programmatisch getoetst wordt. Van Schilt-Mol, Sluijsmans, Peters, De Beer, Jakobs (2016) zien een toetsprogramma als één van de zes entiteiten van toetskwaliteit. Voor elk van de entiteiten zijn kwaliteitscriteria te onderscheiden. Een kwalitatief goed toetsprogramma dat ingezet wordt om programmatisch te toetsen, kan verschillende verschijningsvormen hebben. Tijdens de postermarkt heb ik de ervaringen met programmatisch toetsen gehoord van drie verschillende opleidingen.

Binnen de Haagse Hogeschool heeft de opleiding Voeding en Diëtetiek gekozen om in leerjaar 4 de losse leerlijnen los te laten in het toetsprogramma. Leerjaar 4 wordt afgesloten met één summatief toetsmoment dat in één keer 60 studiepunten oplevert. De opleiding kon niet vertellen wat de ervaringen hiermee zijn, want deze manier van programmatisch toetsten wordt komend jaar uitgeprobeerd.

De opleiding Diergeneeskunde aan de Universiteit Utrecht werkt met een portfolio tijdens de laatste twee jaar co-schappen van de opleiding. Tijdens de co-schappen verzamelt de student verschillende bewijzen van functioneren bijvoorbeeld in de vorm van praktijkbeoordelingen. Aan het eind van deze twee jaar vindt een gesprek plaats met de begeleider waarin op basis van de inhoud van het portfolio wordt vastgesteld of de student als startbekwaam kan worden beschouwd. Deze laatste summatieve beoordeling wordt vooraf gegaan aan meerdere formatieve evaluatiemomenten.

De Hogeschool Utrecht heeft bij de opleiding Communicatie gekozen voor twee integrale summatieve assessments in leerjaar 2; één halverwege leerjaar 2 en één assessment aan het eind van leerjaar 2. Beide assessments leveren 30 studiepunten op en beslaan alle competenties die op dat niveau verworven moeten worden. Het integraal assessment bestaat uit drie onderdelen (a) een portfolio, (b) een presentatie en (c) een criteriumgericht interview. Alle onderdelen worden beoordeeld door twee beoordelaars. Gedurende leerjaar 2 werken studenten in zogenaamde leerteams aan projecten die afgesloten worden met formatieve beroepsproducten. In een leerteam wordt gewerkt aan een project, taak, probleem enzovoort om kennis te construeren en leerdoelen te bereiken (Fransen, Weinberger, & Kirschner, 2013). De docent geeft in de rol van coach begeleiding en formatieve feedback op de geleverde producten. De opleiding en de studenten zijn enthousiast over deze manier van programmatisch toetsen in leerjaar 2. Studenten ervaren een grote mate van vrijheid voor de invulling van hun individuele leerproces. Docenten geven aan dat de begeleiding en het uiteindelijke integrale assessment arbeidsintensief is, maar dat het de creativiteit, zelfsturing en het lerende vermogen van studenten vergroot.

Wat mij opvalt is dat bij alle opleidingen gewerkt wordt met grote toetseenheden aan het eind van een bepaalde leerperiode. De nadelen die kleven aan toetsen die een groot aantal studiepunten opleveren, staan haaks op wat beoogd wordt met een kwalitatief goede toets, namelijk een betrouwbaar en valide oordeel uitspreken over de prestatie van een student. Het is me niet duidelijk geworden op welke manier de toetskwaliteit van de afsluitende summatieve toetsen geborgd wordt.

**Middagprogramma**

Het middagprogramma wordt gestart met de tweede workshopronde. Daarna volgt de tweede keynote. De middag wordt afgesloten met de lancering van de toolkit formatief toetsen.

Workshopronde 2 Toolkit formatief toetsen

Dr. Judith T.M. Gulikers (Universitair docent, Wageningen University & Research) presenteert tijdens deze workshop een toolkit (Gulikers & Baartman, 2018) met professionaliseringsmateriaal voor docenten. Deze toolkit is ontwikkeld in samenwerking met vier hogescholen (Hogeschool Utrecht , Saxion Hogescholen, Avans Hogeschool, Zuyd Hogeschool) op basis van de overzichtsstudie naar formatief toetsen van Gulikers en Baartman (2017). Vandaag wordt de eerste versie van de toolkit gelanceerd, in september komt de definitieve versie. Gulikers toont twee filmpjes (a) de introductie van de formatieve toetscyclus en (b) het doel van de toolkit. Tijdens beide filmpjes zie je Gulikers en Baartman zelf in beeld om hun boodschap te vertellen. Dit beeldmateriaal is één van de instrumenten uit de toolkit.

Gulikers en Baartman (2017, p. 11) definiëren formatief toetsen als: “alle vormen van formele en informele toetsing in de klas die gebruikt worden om zicht te krijgen op het leren van studenten om op basis hiervan het leren van de studenten te bevorderen en/of onderwijsactiviteiten aan te passen”. In deze overzichtsstudie presenteren Gulikers & Baartman (2017) de formatieve toetscyclus als theoretisch- en analysekader. De formatieve toetscyclus bestaat uit vijf fasen (a) verhelderen van verwachtingen, (b) ontlokken van studentreacties, (c) analyseren en interpreteren van studentreacties, (d) communiceren over resultaten met studenten, en (e) vervolgactiviteiten ondernemen (Gulikers en Baartman, 2017). De toetscyclus voor formatief toetsen is gebaseerd op twee stromingen (a) *assessment for learning* en (b) *data-based decision making.* *Assessment for learning* is vooral gericht op het stimuleren van leerprocessen in de klas middels interactie tussen student en docent. Feedback speelt hierin een belangrijke rol als het gaat om het verbeteren van leren (Hattie & Timperley, 2007). *Data-based decision making* speelt daarentegen meer op school- of vakniveau en is vooral gericht op het verhogen van leerresultaten (Lai & Schildkamp, 2013). Beide stromingen benadrukken dat de docent een cruciale rol heeft in het optimaliseren van een leerproces (Gulikers & Baartman, 2017).

Het doel van de toolkit is kritisch en doelgericht formatief toetsen in de les verbeteren. De toolkit bevat daarom allerlei materialen en instrumenten die ingezet kunnen worden als professionaliseringsinterventies in docententeams. Tijdens de workshop hebben we zelf een training ontworpen voor een docententeam. Op deze manier heb ik goed kennis kunnen maken met de materialen uit de toolkit.

Voor de tweede keynote rijkt dagvoorzitter Leenknecht een boek uit aan de deelnemer die als eerste de bingo heeft gehaald. Ze heeft inderdaad acht mensen gesproken van acht verschillende hogescholen. De bingo is een ludieke manier om verschillende mensen te spreken.

Keynote 2: Leren van toetsen

In keynote 2 geeft econoom Prof. dr. Jan Riezebos (hoogleraar onderwijsinnovatie Faculteit Economie en Bedrijfskunde Rijksuniversiteit Groningen) een lezing, getiteld ‘Toetsing continue verbeteren: inspiratieverhaal’. In zijn oratie stelt Riezebos (2017) dat onderwijsinnovatie begint met luisteren; met vragen stellen en verbinding zoeken. Luisteren in termen van klantwaarde identificeren en contact zoeken met de klant. Dat betekent dat we in gesprek moeten met de student, want hij staat centraal in toetsing, zo stelt Riezebos. Een toets moet meerwaarde hebben voor het afnemende werkveld, voor het leerproces van de student en voor het leerproces van de organisatie. Riezebos heeft dat leerproces van dichtbij begeleid middels *assurance of learning, Assurance of learning* verwijst naar processen om aan te tonen dat studenten leerverwachtingen bereiken door de programma's waaraan zij deelnemen. *Assurance of learning* helpt onderwijsinstellingen om verantwoordelijkheid voor de beoogde doelen aan te tonen, hun onderwijsprogramma te verbeteren en studiesucces te evalueren (AACSB, 2017). Riezebos omschrijft *assurance of learning* als het proces dat je als organisatie doormaakt op basis van toetsen van studenten. Je gebruikt toetsen om zelf van te leren als organisatie.

Riezebos onderscheidt vier problemen waar we binnen onderwijsland mee geconfronteerd worden. Ten eerste is er sprake van ketenverliezen. Iedereen doet goed zijn best, maar wel op een eilandje. Ketenverliezen ontstaan door gebrek aan afstemming en door onnodige herhaling. Onderwijslogistiek houdt in dat je je richt op processen in plaats van op prestaties. Toetsen is een vorm van prestatie, zo stelt Riezebos. Veel toetsing kan passen in procesdenken, maar dat vraagt een andere manier van kijken naar toetsen. Het tweede probleem dat Riezebos herkent in het onderwijs is de variatie van verschillende instroom van studenten. Als onderwijsinstelling durven we geen ‘nee’ te zeggen als een student zich aanmeldt. Wat volgens Riezebos nodig is voor kwaliteit, is variatie in het doorlopen van het proces. Het derde probleem dat Riezebos aanstipt is dat onderwijs mensenwerk. Mensen zijn essentieel in het uitvoeren van processen. Riezebos pleit ervoor om mensen en hun onderlinge samenwerking te plaatsen boven processen en hulpmiddelen. Als laatste probleem noemt Riezeboes de ondersteuning van processen. Het doel zou moeten zijn het juiste aanbod op het juiste tijdstip bij de juiste student. Maar is de manier waarop we onze logistieke systemen organiseren effectief en efficiënt?

Na dit betoog verrast Riezebos de deelnemers met een toets over innovatief toetsen op papier. Er worden blaadjes uitgedeeld met vijf vierkeuze vragen en een krasblaadje om de antwoorden op in te vullen. Bij een juist antwoord verschijnt een sterretje in het vakje.

Riezebos eindigt zijn presentatie met een aantal oplossingen vanuit onderwijslogistiek. Kijk vanuit de klant naar jezelf en het proces. Creëer meerwaarde voor de klant, in dit geval de student. Dat kunnen verschillende soorten waarden zijn (a) functionele waarde, (b) emotionele waarde, (c) sociale waarde en (d) economische waarde. Riezebos sluit af met het *statement* dat je moet ‘begrijpen voordat je gaat verbeteren’. Dat kan je doen door daadwerkelijk naar de werkplek van de student toe te gaan en te kijken, open vragen te stellen en respect te tonen voor wat er gebeurt binnen een bedrijf.

De processen binnen het bedrijfsleven en de problemen die daarin optreden, zijn voor mij herkenbaar in de onderwijssetting. Zoals Riezebos aangeeft, toetsing past in het denken in processen. Ik moet denken aan de presentatie van vanochtend waarin Baartman spreekt over formatief toetsen als middel voor een lange termijn leerproces. De formatieve toetsvorm die Riezebos kiest als voorbeeld met de korte toets op papier, was goed gekozen. Ook met zo’n simpele formatieve toetsvorm, geef je de lerenden feedback mee. Het was jammer dat de toetsvragen van slechte kwaliteit waren, zeker bij de doelgroep van dit symposium die toetskwaliteit hoog in het vaandel heeft staan.

Lancering toolkit formatief toetsen

Met deze toolkit komen Gulikers en Baartman (2018) tegemoet aan de wens vanuit de onderwijspraktijk voor bruikbare materialen. De vier hogescholen die instrumenten hebben uitgeprobeerd geven aan dat ze er inspiratie en enthousiasme uit hebben gehaald. Baartman zegt dat het verbeterpunten heeft opgeleverd voor de toolkit. De toolkit is te vinden op de site van het Platform Leren van toetsen <https://lerenvantoetsen.nl/toolkit-formatief-toetsen/>. In de toolkit zijn kennisclips, powerpointpresentaties, artikelen en veel praktische tools opgenomen. Gulikers en Baartman presenteren in een filmpje hoe je de toolkit kunt gebruiken.

Gulikers en Baartman sluiten af met de oproep om je ervaringen met de toolkit te delen via de blogpagina formatieve strategieën. Op basis van onze ervaringen kunnen ze nog wat sleutelen aan de eerste versie, voordat in september de definitieve versie verschijnt.

Dagvoorzitter Potters bedankt alle sprekers en vertelt dat er een digitale evaluatie volgt. De organisatie hoopt dat we veel nieuwe mensen hebben leren kennen met de bingo.

**Evaluatie en conclusie**

Ik ben met drie vraagstellingen dit symposium begonnen: Wat kan de bijdrage van programmatisch toetsen zijn aan het verwerven van het eindniveau Bachelor of Nursing 2020? Welke rol kan de docent spelen bij de inzet van formatieve toetsing binnen onze opleiding tot verpleegkundige? En hoe kan de docent worden toegerust om formatieve toetsing in te zetten?

Is een toetsprogramma hetzelfde als programmatisch toetsen? Nee, een goed toetsprogramma wil niet per definitie zeggen dat er sprake is van programmatisch toetsen. Deze spraakverwarring leefde voor dit symposium bij mij, maar is nu uit de wereld. Programmatisch toetsen omvat veelal een toetsprogramma van summatieve en formatieve toetsvormen. De nadruk ligt op meerdere formatieve beoordelingsmomenten en slechts enkele summatieve beoordelingen. De bijdrage die programmatisch toetsen kan leveren zit in de toevoeging van het formatief beoordelen gericht op de ontwikkeling van de student en minder op het beoordelen van een prestatie. De voorbeelden die ik heb gezien tijdens de postermarkt, hebben me geïnspireerd om binnen de opleiding verpleegkunde te kijken naar de inzet van formatief toetsen.

De docent speelt een cruciale rol bij formatief toetsen. Hoe die rol gedefinieerd moet worden, kwam niet expliciet aan bod tijdens dit symposium. Gulikers en Baartman (2017, p. 6) beschrijven in hun overzichtsstudie wel een aantal concrete docentcompetenties die noodzakelijk zijn voor formatief toetsen, te weten ‘(a) begrip hebben van formatief toetsen, (b) vakinhoudelijke kennis en inzicht in *learning progressions* kunnen toepassen, (c) een vakdidactisch handelingsrepertoire kunnen inzetten, (d) student-docent relaties expliciet aan de orde kunnen stellen met een focus op student-sturing en (e) technologie kunnen gebruiken ter ondersteuning van het formatieve toetsproces’. De toolkit (Gulikers & Baartman, 2018) bevat een scala aan professionaliseringsinterventies om de docent toe te rusten om formatieve toetsing in te zetten. In ons team staat formatief beoordelen in de kinderschoenen. Om met mijn team in gesprek te gaan over formatief beoordelen zijn de kaartspellen ‘Gezamenlijke taal creëren’ en ‘Formatieve toetscyclus’ behulpzaam.

Meerdere sprekers benadrukken het belang van samenwerking als het gaat om curriculum- en toetsontwikkeling. Baartman stelt dat programmatisch toetsen teamwork is. De gezamenlijke verantwoordelijkheid vat Sluijsmans samen met de term teamleren. De Geus stelt dat ontwikkelen een proces is dat je samen doet. Riezebos onderstreept dit met de stelling dat we samen op weg gaan naar beter onderwijs.

Zoals de titel van het symposium al aangeeft: er is veel te leren van toetsen. Formatief toetsen wordt ook toetsen om van te leren genoemd. Niet vreemd dat de weg naar beter onderwijs gezien kan worden als een lange termijn leerproces, dat onlosmakelijk verbonden is met formatieve toetsen. *Assessment for learning* kan nooit zonder feedback. Ja, er valt nog veel te leren van en rnet elkaar over toetsen!

**Referenties**

Academie Gezondheidszorg (2017). Onderwijs- en toetsbeleidsplan Academie Gezondheidszorg 2017.

Andriessen, A., Sluijsmans, D., Snel, M., & Jacobs, A. (2017). Protocol Verbeteren en Verantwoorden

van Afstuderen in het hbo 2.0. Geraadpleegd op 18 april 2018, van

<http://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/000/744/original/Vreemde_Ogen_Dwingen-Protocol_Verbeteren_en_Verantwoorden_van_Afstuderen_in_het_hbo.pdf?1491999591>

Association to Advance Collegiate Schools of Business International (AACSB) (2017). *Eligibility*

*Procedures and Accreditation Standards for Business Accreditation.* Geraadpleegd op 4 juni 2018,

van <https://www.aacsb.edu/-/media/aacsb/docs/accreditation/standards/business-accreditation->

2017-update.ashx?la=en

Baartman, L. & Kloppenburg, R. (2013). *KwaliteitsInstrument Toetsprogramma’s in*

*beroepsgericht onderwijs (KIT)*, Hogeschool Utrecht. Geraadpleegd op 10 mei 2018, van

[www.kwaliteit-toetsprogramma.nl](http://www.kwaliteit-toetsprogramma.nl)

Fransen, J., Weinberger, A., & Kirschner, P. (2013). Team effectiveness and team

development in CSCL. *Educational Psychologist, 48*(1), 9-24.

doi:10.1080/00461520.2012.747947

Gulikers, J. T. M. & Baartman, L. (2017). *Doelgericht professionaliseren. Formatief toetsen met*

*effect! Wat DOET de docent in de klas?* Eindrapport NRO-PPO overzichtsstudie dossiernummer

405-15-722. Geraadpleegd op 6 juni 2018, van <https://lerenvantoetsen.nl/wp-content/uploads/2018/05/Eindrapport-NRO-Review.pdf>

Gulikers, J. T. M. & Baartman, L. (2018). *Toolkit formatief toetsten.* Geraadpleegd op 6 juni 2018, van <https://lerenvantoetsen.nl/toolkit-formatief-toetsen/>.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research 77*(1), 81-112. doi: 10.3102/003465430298487

Lai, M. K. & Schildkamp, K. (2013). Data-based decision making: an overview. In: K. Schildkamp, M. K. Lai, & L. Earl, (Eds). *Data-based Decision Making in Education* (pp. 9-12). Dordrecht: Springer.

Lozano, R., Merrill, M. Y., Sammmalisto, K., Ceulemans, K., & Lozano, F. J. (2017).

Connecting competences and pedagogical approaches for sustainable development in higher

education: a literature review and framework proposal. *Sustainability, 9,* 1-15. doi: 10.3390/su9101889

Molkenboer, H. (2015). *Toetsen volgens de toetscyclus. Fase 1,2 en 3 – tot en met gesloten en open*

*vraagvormen.* Enschede: Bureau voor Toetsen & Beoordelen.

Riezebos, J. (2017). *Op weg naar beter onderwijs* (niet gepubliceerde inaugurele rede).

Rijksuniversiteit Groningen: Groningen. Geraadpleegd op 4 juni 2018, van

http://www.janriezebos.nl/userfiles/file/oratie%20Riezebos.pdf

Van Berkel, H., Bax, A., & Joosten-ten Brinke, D. (Eds.) (2017). *Toetsen in het hoger onderwijs.*

Houten: Bohn Stafleu van Loghum.

Van der Vleuten, Schut, & Heeneman (2018). Programmatisch toetsen als motor voor professioneel

leren in het hoger onderwijs. In: D. Sluijsmans, & M. Segers, (Eds.). *Toetsrevolutie (pp. 124-135).*

Culemborg: Phronese.

Van der Vleuten, C.P.M. & Schuwirth, L. W. T. (2005). Assessment of professional competence:

From methods to programmes. *Medical Education, 39,* 309-317.

Van Schilt-Mol, T., Sluijsmans, D. M. A., Peters, M., De Beer, F., & Jakobs, L. (2016). Bovenkant formulier

[De touwtjes in](https://youlearn.ou.nl/web/om0403162114/cursus/-/coursenavigator/8463170?_nl_ou_dlwo_courseview_WAR_nloudlwocourseplanportlet__facesViewIdRender=%2Fxhtml%2Fviewer%2FcourseNavigator.xhtml)

[handen: toetskwaliteit in het hbo.](https://youlearn.ou.nl/web/om0403162114/cursus/-/coursenavigator/8463170?_nl_ou_dlwo_courseview_WAR_nloudlwocourseplanportlet__facesViewIdRender=%2Fxhtml%2Fviewer%2FcourseNavigator.xhtml) *[OnderwijsInnovatie, 4](https://youlearn.ou.nl/web/om0403162114/cursus/-/coursenavigator/8463170?_nl_ou_dlwo_courseview_WAR_nloudlwocourseplanportlet__facesViewIdRender=%2Fxhtml%2Fviewer%2FcourseNavigator.xhtml)*[, 15-24.](https://youlearn.ou.nl/web/om0403162114/cursus/-/coursenavigator/8463170?_nl_ou_dlwo_courseview_WAR_nloudlwocourseplanportlet__facesViewIdRender=%2Fxhtml%2Fviewer%2FcourseNavigator.xhtml)