

Toetsen doe je samen met studenten

Dit artikel is het vijftiende 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 onderwijstechnologische inzichten in het dagelijks werk van de docent, het onderwijsteam of de opleidingsmanager.

Auteurs

Elly Vermunt
Dominique Sluijsmans

Vermunt is werkzaam bij de Faculteit Commercieel en Financieel Management van Zuyd Hogeschool en lid van het lectoraat Professioneel Beoordelen van Zuyd Hogeschool, Sluijsmans is lector Professioneel Beoordelen bij Zuyd Hogeschool.

Reacties op dit artikel naar:
elly.vermunt@zuyd.nl

Inhoud

- _ Samenvatting
- _ Drie belangrijke ontwikkelingen op het gebied van toetsing
- _ Het belang van het betrekken van studenten bij de toetscyclus
- _ De praktijkvraag en het doel van de uitgevoerde studie
- _ De onderzoeksoepzet
- _ Formatieve activiteiten per fase in de toetscyclus
- _ Discussie
- _ Vervolgonderzoek

Box 1: De toetscyclus

Box 2: Feedup, feedback en feedforward in de context van de toetscyclus

Box 3: Uitgewerkt voorbeeld voor de toetsvorm 'projectrapport'

Box 4: Uitgewerkt voorbeeld voor een toets met gesloten vragen

Box 5: Overzicht van besproken formatieve activiteiten per fase in de toetscyclus

Samenvatting

Om tot een goede toets te komen, is het van belang een aantal ontwerpfases te doorlopen. Deze fases vormen de zogenaamde toetscyclus. Deze toetscyclus wordt - in meer of mindere mate - in de praktijk vaak door een individuele docent doorlopen. Daarbij worden vooral toetsen ontwikkeld vanuit uitsluitend een summatieve functie: de toets leidt tot de beslissing of een student voor de toets zakt of slaagt. In dit artikel wordt de toetscyclus

vanuit een formatief perspectief bekeken: op welke wijze kunnen studenten in de verschillende fases van de toetscyclus actief worden betrokken teneinde de leerwaarde en motivatie van zowel de student als de docent te verhogen? In dit artikel worden praktische activiteiten beschreven die docenten tijdens de verschillende fases kunnen inzetten. We hopen met deze bijdrage docenten in het hoger (beroeps)onderwijs te inspireren om toetsen niet meer te benaderen als iets dat los staat van het onderwijs, maar als een integraal onderdeel van het onderwijsleerproces van hun studenten maar ook van henzelf als hbo-professional.

Drie belangrijke ontwikkelingen op het gebied van toetsing

Lange tijd zijn onderwijs en toetsing door zowel de wetenschap als de praktijk als twee gescheiden werelden benaderd. Het gevaar van het losweken van toetsing van het onderwijs leidt tot het bekende toetsgedrag ('wash-back effect'; Alderson & Wall, 1993): de boeken verdwijnen in de kast en de toets in de prullenbak als de voldoende binnen is (Frederiksen, 1984). Omdat toetsen een enorme impact op het leren en de motivatie van studenten hebben, is er veel onderzoek gedaan naar de wijze waarop onderwijs en toetsing beter op elkaar kunnen aansluiten. Dit heeft er toe geleid dat toetsing in toenemende mate wordt benaderd vanuit een ontwerp-perspectief (Janssen-Noordman & Van Merriënboer, 2002; Van Der Vleuten

& Schuwirth, 2005): toetsing vormt een integraal onderdeel in het ontwerp van het curriculum. Dit ontwerpdenken heeft tot drie belangrijke ontwikkelingen geleid.

Ontwikkeling 1: Ontwerpen van toetsen is een cyclisch proces

De eerste ontwikkeling is dat er veel meer aandacht is gekomen voor het cyclisch ontwerpen van toetsen en de eisen die dit stelt aan docenten. In het rapport 'Verantwoord toetsen en beoordelen in het hoger beroepsonderwijs' van de Expertgroep BKE/SKE (2013), geschreven in opdracht van de Vereniging Hogescholen, zijn bijvoorbeeld de leeruitkomsten en indicatoren benoemd waaraan de examinator in het hbo dient te voldoen om te mogen beoordelen. Doorgaans worden de stappen die bij het ontwerpen van toetsen dienen te worden genomen geïllustreerd met de toetscyclus. De toetscyclus beschrijft de fases die een docent doorloopt vanaf het ontwerp tot en met de interpretatie en nabespreking van een

toets. Er zijn meerdere vormen van toetscycli voorhanden die allen grote overeenkomsten vertonen, voor dit artikel hanteeren we de cyclus zoals beschreven door Joosten-Ten Brinke and Sluijsmans (2012) (zie box 1).

In de eerste stap van de toetscyclus, het toetsontwerp, worden de leerdoelen en de leeruitkomsten beschreven en wordt een passende toetsvorm gekozen. Vervolgens worden toetsvragen geformuleerd en wordt de toets geconstrueerd. De beoordelingsvoorschriften worden eveneens in deze fase opgesteld: Wat zijn de beoordelingscriteria en standaarden? Op basis van welke beslisregels wordt een beoordeling gegeven? Op welke wijze wordt het toetsresultaat vastgesteld (cijfer, voldoende/onvoldoende, kwalitatieve feedback, enzovoort). Vervolgens wordt de toets afgenomen. De toets wordt beoordeeld aan de hand van de opgestelde beoordelingsvoorschriften, waarna een beoordeling (bijvoorbeeld een cijfer) wordt toegekend. Op grond hiervan kan

een beslissing (bijvoorbeeld gezakt/ge-slaagd) worden genomen. Tijdens de analysefase worden de resultaten geanalyseerd en ontstaat inzicht in zaken zoals de moeilijkheidsgraad, betrouwbaarheid en validiteit van de toets als geheel en de onderliggende toetsvragen. Op grond hiervan kunnen beoordelingen worden aangepast en verbetervoorstellen voor de toets en het onderwijs worden geformuleerd. Het is een cyclisch proces, waarbij op grond van evaluaties verbeteringen worden doorgevoerd in de volgende ronde van de cyclus.

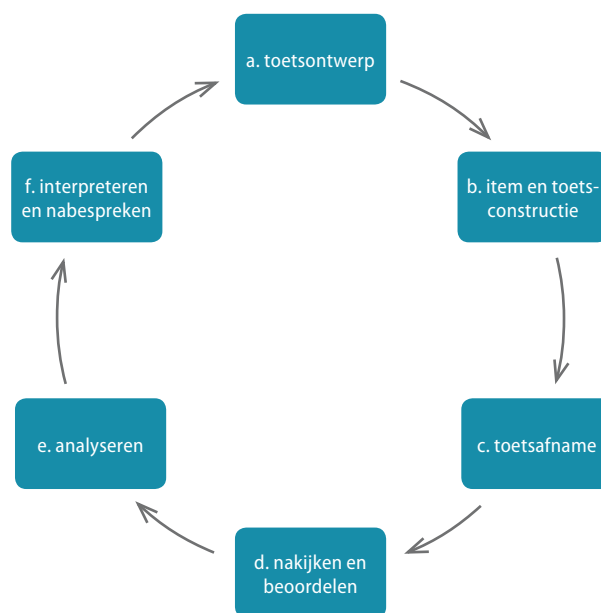
Ontwikkeling 2: Formatieve functie van toetsen is belangrijk

Een tweede ontwikkeling is dat toetsing niet meer alleen de functie heeft van certificering en/of selectie (de summatieve functie van toetsing), maar dat toetsen ook steeds meer zijn gericht op het stimuleren van verder leren (de formatieve functie van toetsing). Eerdere studies hebben aangetoond dat de klassieke summatieve functie van toetsing gericht op afsluiting en certificering niet toereikend is om kwaliteit van leren te borgen. Ook stimuleren deze toetsen onvoldoende de verantwoordelijkheid van studenten. Vanuit wetenschappelijk en maatschappelijk perspectief is veel bewijs voorhanden dat formatief toetsen het leren positief beïnvloedt. Formatief toetsen verwijst naar alle activiteiten die docenten en studenten uitvoeren om het leren en de motivatie te versterken (Sluijsmans, Joosten-Ten Brinke, & Van der Vleuten, 2013). Er is echter pas sprake van formatief toetsen wanneer de informatie door studenten en docenten daadwerkelijk is gebruikt voor verder leren (Ramaprasad, 1983; Sadler, 1989; Taras, 2005), met als gevolg bijvoorbeeld verbeterde leerprestaties, een verhoogde motivatie en betere toetsontwerpen.

Ontwikkeling 3: Toegenomen aandacht voor participatief ontwerp van onderwijs

Een derde ontwikkeling is de toegenomen aandacht voor het betrekken van studenten bij het ontwerpen van het cur-

BOX 1: DE TOETSCYCLUS (JOOSTEN-TEN BRINKE & SLUIJSMANS, 2012)



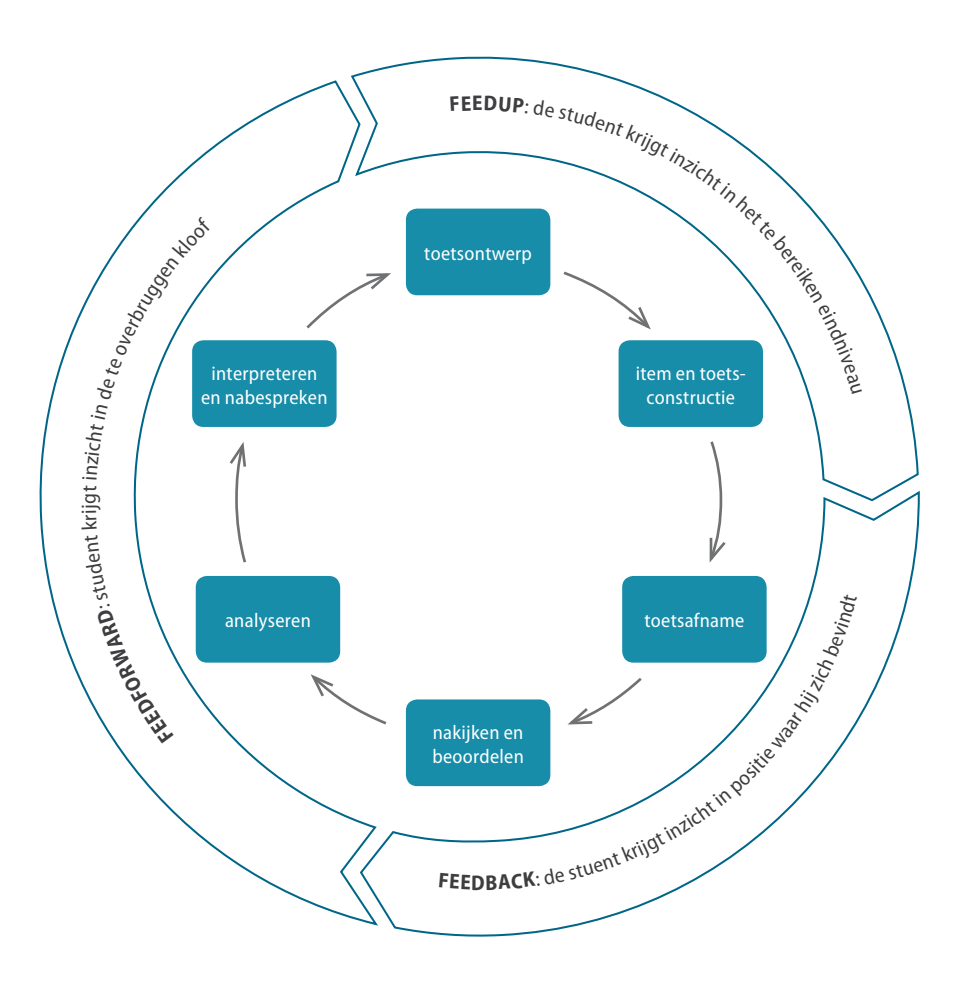
riculum, het zogenaamde participatief ontwerp (participatory design) van leeromgevingen. Hierbij worden alle stakeholders betrokken bij het ontwerpen van de leeromgeving (Könings, Seidel, & Van Merriënboer, 2014). Bron and Veugelers (2014) noemen een aantal argumenten waarom het zinvol is leerlingen bij het ontwerpen van curricula te betrekken. Eén van de argumenten is het zogenaamde normatieve argument: studenten hebben het recht om mee te denken over zaken die hen aangaan. Daarnaast noemen zij het onderwijskundig argument, namelijk dat studenten leren van het participeren in het ontwerpproces, en het relevantie argument: door het betrekken van studenten wordt het curriculum verbeterd. Het betrekken van studenten bij de toetsing is een ontwikkeling die naadloos aansluit bij het participatief ontwerpen.

Het belang van het betrekken van studenten bij de toetscyclus

Vanuit onderzoek zijn drie redenen te noemen waarom het betrekken van studenten bij de toetscyclus door het inzetten van formatieve activiteiten zinvol is. De eerste reden is dat hierdoor de leerwaarde voor de student verhoogd wordt doordat het handvatten biedt om de wijze van feedback te organiseren. Hattie and Timperley (2007) geven aan dat effectieve feedback drie vragen beantwoordt:

- 1 Waar werkt de student naar toe? Wat zijn de doelen en bijbehorende beoordelingscriteria en standaarden? Bij de beantwoording van deze vraag is sprake van feedup.
- 2 Waar staat student nu? Welke vooruitgang is geboekt ten aanzien van de beoordelingscriteria en standaarden? Bij de beantwoording van deze vraag is sprake van feedback.
- 3 Hoe komt de student in de gewenste situatie? Welke aanpak is nodig om te groeien en de gewenste leerdoelen en -uitkomsten te behalen? Wanneer ingegaan wordt op deze vraag is sprake van feedforward.

BOX 2: FEEDUP, FEEDBACK EN FEEDFORWARD IN DE CONTEXTE VAN DE TOETSCYCLUS



Door studenten bij de toetscyclus te betrekken, kan bij de verschillende fases, deze feedup, feedback en feedforward worden geboden. Hierbij zijn de drie vormen van feedback duidelijk te koppelen aan de diverse fases (zie box 2). In de praktijk kan soms een combinatie van de verschillende vormen van feedback tijdens één fase voorkomen. Zo kan bij een inzage zowel feedback als feedforward worden gegeven. Een tweede reden voor het actief betrekken van studenten bij de toetscyclus is dat hierdoor het zelfregulerend vermogen en de motivatie van studenten kan worden verhoogd. Eén van de belang-

rijkste vaardigheden van beginnende beroepsbeoefenaren is dat zij in staat zijn zichzelf te sturen en te blijven ontwikkelen. Zelfsturend vermogen wordt gezien als een van de meest essentiële vaardigheden voor het leren op de lange termijn (Sluijsmans, 2008; 2013). Het kunnen omgaan met feedback en deze vertalen naar aangepast handelen en denken is van belang om zelfregulatie te verhogen. Daarnaast biedt onderzoek aanwijzingen dat het betrekken van studenten bij het toetsproces leidt tot verhoging van de motivatie (White & Frederiksen, 1998).

Een derde reden is dat het actief betrekken van studenten bij toetsing ook leerwaarde kan hebben voor de docent. Door studenten in alle fases van de toetscyclus actief te betrekken krijgt de docent meer inzicht in het beheersniveau van de studenten en in de denkwijze en mogelijke misconcepties die bij de studenten heersen (bijvoorbeeld over de te behalen leerdoelen). Hierdoor kan de docent vervolgens een betere bijdrage leveren aan het overbruggen van de kloof van waar de student staat en waar deze naar toe wil of moet (William, 2011).

De praktijkvraag en het doel van de uitgevoerde studie

Ondanks het vele onderzoek dat wordt verricht naar het ontwerpen van toetsen met een formatieve functie bereiken de onderzoeksresultaten nog onvoldoende de praktijk. Docenten hebben behoefte aan handvatten die hen ondersteunen in het ontwerpen van toetsen en de wijze waarop zij hierbij de formatieve functie kunnen versterken door het verhogen van de betrokkenheid van studenten bij het toetsproces. Dit was voor het lectoraat Professioneel beoordelen van Zuyd Hogeschool aanleiding om een studie op te zetten naar het betrekken van studenten bij de toetscyclus en hierbij met praktische voorbeelden te komen waarmee docenten de formatieve waarde van hun toetsen kunnen versterken. In dit artikel worden de uitkomsten van het eerste deel van deze studie gepresenteerd. Het doel van de studie was meer zicht te krijgen op de manieren waarop een docent studenten kan betrekken bij de toetscyclus teneinde de formatieve waarde van de toets te vergroten. Een activiteit heeft formatieve waarde voor de student als de informatie die de student wordt geboden - door de docent of door anderen, bijvoorbeeld medestudenten - deze student helpt de kloof tussen waar hij staat en waar hij naar toe wil of moet, te over-

bruggen. Een activiteit heeft formatieve waarde voor de docent als de informatie die de docent tijdens die activiteit ontvangt deze docent helpt zijn onderwijs en/of toets aan te passen zodat hij de student beter kan helpen de kloof te dichten tussen waar hij staat en waar hij naar toe wil of moet.

De onderzoeksvraag die centraal stond was: *Op welke manier kan een docent studenten betrekken bij de voorbereiding, uitvoering en evaluatie van een toets (de gehele toetscyclus) met het doel de formatieve waarde voor de student en docent te vergroten?*

De onderzoekopzet

Voor de literatuurstudie is het boek *Embedded formative Assessment* van Dylan Wiliam (2011) als uitgangspunt genomen. Binnen het kennisdomein van formatief assessment en dan specifiek de meer praktische handboeken wordt dit boek tot de standaardwerken gerekend. Wiliam (2011) is geanalyseerd op relevante literatuur over de zes fases van de toetscyclus. Er is aanvullende literatuur gezocht met betrekking tot formatieve activiteiten in fase vijf en zes van de toetscyclus (analyseren; interpreteren en nabespreken). In totaal zijn twaalf studies opgenomen om de onderzoeksvraag te beantwoorden: Butler (1988); Carless (2006); Fontana and Fernandes (1994); Foos, Mora, and Tkacz (1994); Higgins, Hartley, and Skelton (2002); Kohn (1994); Li and De Luca (2014); Macdonald (2005); Taras (2001); White and Frederiksen (1998); Winingier (2005) en Yu and Li (2014).

Formatieve activiteiten per fase in de toetscyclus

In deze paragraaf beschrijven we de formatieve activiteiten die per fase in de toetscyclus zinvol zijn om in te zetten. Hierbij worden, gezien de sterke onderlinge samenhang, steeds twee van de zes fases samen besproken, waarbij achtereenvolgens de nadruk ligt op het bieden

van feedup, feedback en feedforward. Ter illustratie zijn in box 3 en 4 de beschreven formatieve activiteiten praktisch uitgewerkt voor een tweetal verschillende typen toetsen.

Toetsontwerp en item- en toetsconstructie (feedup)

Bij deze twee eerste fases van de toetscyclus staan de leerdoelen en de te bereiken leeruitkomsten centraal. Ook de beoordelingscriteria en standaarden worden in deze fase opgesteld. Juist deze aspecten, die het startpunt moeten vormen van het leerproces voor de student (feedup: wat is het te bereiken eindniveau?), zijn voor studenten vaak onduidelijk. Studenten hebben vaak moeite met het decoderen van de beoordelingscriteria onder andere doordat het taalgebruik waarin leerdoelen en criteria geformuleerd zijn, niet altijd helder is voor studenten (Carless, 2006; Higgins, Hartley, & Skelton, 2002). Er is een tal van methoden voorhanden om studenten actief bij deze fases te betrekken. In de eerste plaats kunnen de leerdoelen en criteria actief met studenten besproken worden. Daarnaast kunnen studenten ook actief betrokken worden bij het opstellen van de beoordelingscriteria en het vertalen van deze criteria in beoordelingsrubrics. In rubrics wordt per criterium een aantal niveaus kwalitatief beschreven, bijvoorbeeld wanneer aan een bepaald criterium in onvoldoende, voldoende mate of zeer goed wordt voldaan¹. Uit onderzoek blijkt dat gezamenlijk opgestelde criteria groten-deels overeenkomen met door de docent opgestelde criteria, maar dat deze werkwijze leidt tot een groter begrip van de criteria bij de studenten. Daarnaast leidt het gezamenlijk opstellen tot een groter gevoel van eigenaarschap van de criteria bij de studenten (Taras, 2001; William, 2011).

Een andere methode die tot een verheldering van de beoordelingscriteria leidt, is het gebruik van deze criteria om eigen

¹ Voor meer informatie zie hoofdstuk 17 in 'Toetsen in het hoger onderwijs' (Van den Bos, Burghout, & Joosten-Ten Brinke, 2014)

werk te beoordelen. Dit effect is onder andere aangetoond in een groot onderzoek van White and Frederiksen (1998). Leerlingen die actief met de beoordelingscriteria werkten – onder andere door zichzelf te beoordelen aan de hand van deze criteria – presteerden beter op pro-

jecten en een onderzoekstoets dan leerlingen uit een controlegroep. Opmerkelijk was dat met name de minder goede leerlingen een grotere vooruitgang boekten dan de goede leerlingen. White and Frederiksen veronderstellen dat goede leerlingen beter uit zichzelf weten waar

goed werk aan moet voldoen en over de metacognitieve vaardigheden beschikken om hun werk te monitoren dan minder goede leerlingen. Daarnaast concludeerden ze dat de motivatie en het zelfvertrouwen van de leerlingen door deze methode verhoogd wordt.

BOX 3: UITGEWERKT VOORBEELD VOOR DE TOETSVORM 'PROJECTRAPPORT'

Context: Gedurende acht weken onderwijs werken studenten in een groep of individueel aan een project. In week 8 wordt het rapport ingeleverd. In week 2 van het volgend blok wordt de beoordeling inclusief feedback teruggekoppeld aan student. In dit voorbeeld wordt uitsluitend de beoordeling van het rapport uitgewerkt. De beoordeling van het proces en het persoonlijk functioneren kan bij een project uiteraard ook beoordeeld worden. Dit wordt in dit voorbeeld echter niet uitgewerkt.

Fase in toetscyclus	Formatieve activiteiten	Opbrengst voor de student	Opbrengst voor de docent
Toetsontwerp/ item- en toets-constructie	Week 1: docent bespreekt met studenten de leerdoelen, opdracht en het beoordelingsformulier. Hierbij bespreekt de docent waarom deze leerdoelen juist leiden tot de gekozen toetsvorm: het projectrapport. Is het helder wat er verwacht wordt? Is taalgebruik bij beoordelingscriteria helder voor student?	Voor de student wordt het helder waar hij naar toe moet werken. Student begrijpt de leerdoelen en de relatie met de beoordelingscriteria.	Voor de docent wordt het duidelijk wat bij studenten nog onduidelijk is en waar hij dient te verhelderen.
Toetsafname en nakijken/ beoordelen	Week 3 of 4: docent laat tijdens bijeenkomst studenten 'oude' rapporten beoordelen aan de hand van een beoordelingsrubric. De studenten lezen en beoordelen in eerste instantie individueel het rapport; vervolgens worden in groepjes van drie à vier studenten deze oordelen onderling besproken. Tot slot bespreekt de docent dit klassikaal na waarbij hij zijn oordeel ook bespreekt.	Door actief met beoordelingsrubric te werken worden de criteria en de standaarden voor de student verhelderd.	Docent bemerkt tijdens deze sessie waar onduidelijkheid zit, waar hij vervolgens extra aandacht aan kan besteden.
	Week 7: voor het inleveren van het rapport laat docent studenten een selfassessment uitvoeren van hun eigen conceptrapport met behulp van de vastgestelde rubric. Hierop geeft de docent feedback.	Student krijgt inzicht in de discrepantie tussen eigen beoordeling en het oordeel van docent. Dit euid tot beter begrip van de standaard waaraan het werk dient te voldoen. Op basis van de feedback kan student rapportage verbeteren.	Docent krijgt inzicht in manier waarop student criteria interpreteert.
Analyseren/ interpreteren en nabespreken	Week 2: docent geeft aan de hand van de beoordelingsrubric een beoordeling en duidelijke feedback en laat hierbij de link naar volgende rapporten en andere thema's ook zien.	Student krijgt inzicht in de beoordeling en de te overbruggen kloof.	Docent krijgt inzicht in lacune bij student.

BOX 4: UITGEWERKT VOORBEELD VOOR EEN TOETS MET GESLOTEN VRAGEN (bijvoorbeeld juist/onjuist of meerkeuzevragen)

Context. Gedurende acht weken werken studenten met behulp van diverse werkvormen als PGO, hoorcolleges, werkcolleges en zelfstudie aan bepaald thema. In week 9 wordt de afsluitende meerkeuze toets afgenomen. In week 2 van het volgend blok wordt de inzage en nabespreking verzorgd.

Fase in toetscyclus	Formatieve activiteiten	Opbrengst voor de student	Opbrengst voor de docent
Toetsontwerp/ item- en toets-constructie	week 1: docent bespreekt leerdoelen met studenten: is het helder wat er verwacht wordt? De docent legt de relatie met de toetsvorm uit. Komt dit overeen met wat zij willen leren, zijn er aanvullende leerdoelen te formuleren?	Voor de student wordt het helder waar hij naar toe moet werken.	Voor de docent wordt het duidelijk wat voor studenten nog onduidelijk is en wat hij dient te verhelderen.
	week 5 en 6: docent laat de studenten individueel toetsvragen ontwikkelen en bespreekt een deel vervolgens klassikaal waarbij hij feedback geeft over het niveau. De docent neemt deze toetsvragen in, waarbij hij aangeeft dat hij een deel van de vragen in toets verwerkt	Door het maken van de vragen werkt student actief met lesstof en krijgt hij een beter beeld van essentie van leerstof. Omdat goede vragen gebruikt worden voor de toets werkt het motiverend.	Docent krijgt inzicht in het beeld dat studenten hebben van te bereiken eindniveau.
Toetsafname / Nakijken en beoordelen	week 9: docent gebruikt voor de samenstelling van de toets een deel van de vragen die door de studenten zijn ontwikkeld. Bij de toetsafname wordt aan het antwoordformulier een feedback-formulier toegevoegd, waarop student kan aangeven welke items hij onduidelijk vindt en waarom. De docent gebruikt deze informatie bij de analyse van de toets.	De in de klas ontwikkelde vragen kunnen in toets voorkomen; de student heeft profijt van activiteit in de fase hiervoor. De student kan bij onduidelijkheid zijn feedback kwijt. Dit verhoogt zijn inbreng.	Docent krijgt inzicht in interpretatie van toetsitems. Op grond hiervan kan docent besluiten eventueel onduidelijk gestelde items te verwijderen en/of aan te passen voor volgende toets.
Analyseren/ interpreteren en nabespreken	week 2 volgend blok: docent bespreekt klassikaal de juiste antwoorden en de toetsanalyse. Hierbij gaat docent actief met de studenten in gesprek over de antwoorden waarbij studenten gevraagd wordt waarom bepaalde vragen goed of niet goed zijn. Daarnaast vraagt docent de studenten om feedback over toetsitems.	Student krijgt een beter beeld van hetgeen hij nog niet kent of begrijpt. Daarnaast krijgt hij een beter beeld van de manier waarop hij de vragen beantwoordt. Eveneens wordt de transparantie van de beoordeling verhoogd.	Docent krijgt feedback over toets-items waardoor hij deze kan verbeteren.

Naast het beoordelen van eigen werk is ook het beoordelen van werk van medestudenten zinvol. Studenten zijn vaak beter in staat om fouten of beperkingen in ander werk te signaleren dan bij henzelf. Indien ze bepaalde fouten hebben herkend en bespreken, is de kans groter dat ze deze fouten in het eigen werk niet herhalen (William, 2011). Door studenten

werk van anderen te laten beoordelen aan de hand van de beoordelingscriteria en deze te bespreken, wordt een beter begrip van de criteria en standaarden bereikt. Tot slot kunnen studenten actief betrokken worden bij de fase van toetsconstructie door henzelf actief toetsvragen te laten ontwikkelen. Dat dit een leerzame

activiteit is, toonde het onderzoek van Foes, Mora and Tkacz (1994) aan. Studenten die zelf toetsvragen ontwikkelden, scoorden significant beter op vergelijkbare vragen (i.e. vragen over hetzelfde onderwerp) dan studenten die (dezelfde) oefenvragen uitgereikt kregen. Uit het onderzoek kwam ook naar voor dat studenten moeite hebben met het bepa-

len van de hoofdlijnen van de stof (Wat zijn de onderwerpen waar de toetsvragen over moeten gaan?). William (2011) bespreekt laatstgenoemd punt als een voordeel van het actief betrekken van studenten bij de ontwikkeling van toetsvragen. De door de student ontwikkelde vragen geven de docent een goed beeld van wat de studenten blijkbaar als hoofdlijnen ervaren. Aan de hand hiervan kan een docent bijsturen.

Zoals blijkt zijn in deze fase tal van activiteiten mogelijk. In box 5 staat een samenvatting van de activiteiten weergegeven die in deze en de volgende fases kunnen worden ingezet.

Toetsafname en nakijken/beoordelen (feedback)

Bij de toetsafname en de beoordeling wordt duidelijk waar de student nu staat (feedback). Het actief betrekken van een student bij deze fases kan het inzicht van de student in de huidige positie en het begrip voor de beoordeling vergroten. Belangrijke activiteiten die hierbij ingezet kunnen worden zijn self- en peerassessment. Zoals aangegeven in de fase hiervoor blijkt uit onderzoek dat als studenten hun eigen werk en/of andermans werk beoordelen dit tot een verheldering

van de beoordelingscriteria en hun eigen prestaties ten opzichte van deze criteria leidt, waardoor hun eigen scores verhoogd worden (Fontana & Fernandes, 1994; White & Frederiksen, 1998). Indien hierbij gewerkt wordt met heterogene groepen wordt het leereffect vergroot. Met name laagsoorders profiteren van een heterogene groepssamenstelling, terwijl hoogsoorders hiervan geen onderdeel ondervinden. Het is zinvol self- en peerassessment in te zetten in de fase vóór de docentbeoordeling, dat wil zeggen tijdens de lessen op opdrachten of op conceptversies. Tal van onderzoek toont aan dat leerlingen meer open staan voor feedback als ze het nog kunnen gebruiken voor de verbetering van hun werk (Butler, 1988; Carless, 2006; Kohn, 1994).

Daarnaast kan self- en peerassessment tegelijk met of ná de docentbeoordeling worden georganiseerd. Hierbij wordt de studenten gevraagd het eigen werk te beoordelen en van een oordeel te voorzien aan de hand van beoordelingscriteria. Zo beschrijft Taras (2001) een methode waarbij een student, na de beoordeling door de tutor, het eigen werk moet beoordelen aan de hand van de criteria en feedback van de tutor. Hierbij wordt het oordeel

van de tutor pas getoond nadat student het eigen werk beoordeeld heeft.

Vervolgens wordt het studentoordeel vergeleken met het docentoordeel.

Eventuele onduidelijkheid over het oordeel wordt daarna besproken (eventueel met een tweede beoordelaar), waarbij nog ruimte is om te onderhandelen over het oordeel. Bij het onderzoek is bewust gekozen de studenten de feedback wél en het oordeel nog niet te tonen. Veel onderzoek toont aan dat indien studenten een beoordeeld werk terugkrijgen de meeste aandacht naar het eindoordeel (meestal een cijfer) gaat en niet naar de feedback (Taras, 2001; William, 2011).

Een vergelijkbare methode wordt beschreven door Macdonald (2005), waarbij de student nadat het werk gemaakt is, het werk zelf gaat beoordelen en becijferen aan de hand van beoordelingscriteria, waarna de docent deze beoordelingen beoordeelt. Bij beide methodes wordt het uiteindelijke cijfer wel bepaald door docent, maar de student wordt wel zeer actief bij de beoordeling betrokken. Dit leidde tot meer begrip voor het behaalde cijfer en tot een vergroting van het inzicht in de te verbeteren aspecten. Bij beide onderzoeken was er grote overeenstemming tussen de docent- en studentbeoordeling.

BOX 5: OVERZICHT VAN BESPROKEN FORMATIEVE ACTIVITEITEN PER FASE IN DE TOETSCYCLUS

Fase	Mogelijke activiteiten
Toetsontwerp / Item- en toetsconstructie (feedup)	<ul style="list-style-type: none"> • Leerdoelen actief met studenten bespreken. • Leerdoelen en beoordelingscriteria (bijvoorbeeld rubrics) samen met studenten ontwikkelen. • Tijdens de lessen ander werk beoordelen aan de hand van de beoordelingscriteria. • Studenten actief zelf toetsvragen laten ontwikkelen.
Toetsafname / Nakijken en beoordelen (feedback)	<ul style="list-style-type: none"> • Studenten eigen werk laten beoordelen aan de hand van beoordelingscriteria (selfassessment). • Studenten elkaars werk laten beoordelen aan de hand van beoordelingscriteria (peerassessment).
Analyseren / Interpreteren en nabespreken (feedforward)	<ul style="list-style-type: none"> • Groepsgewijze nabespreking van een toets door studenten onderling. • Nabespreking waarbij docent student actief betreft en feedback vraagt over vragen. • Studenten de eigen toets laten beoordelen (naast de docent beoordeling) aan de hand van beoordelingscriteria.

Alhoewel studenten de methodes als zinvol ervaren, moesten zij wel wennen aan de methodes. De docenten waren ook van mening dat het inzetten van de methode met name zinvol is vanaf het eerste of tweede studiejaar omdat studenten dan nog open staan voor nieuwe methodes en een maximaal cumulatief effect kan worden bereikt. Het is een goede investering op de lange termijn, maar niet op de korte termijn.

Zoals aangegeven heeft selfassessment een positief effect op het leereffect. Een ander belangrijk effect is dat het bijdraagt aan het ontwikkelen van de zelfregulatievaardigheden van studenten (Sluijsmans et al., 2013; Taras, 2001; William, 2011).

Analyseren/interpreteren en nabespreken (feedforward)

Bij de laatste twee fases uit de toetscyclus vindt de analyse en nabespreking plaats. Uit onderzoek van de Landelijke Studenten Vakbond (LSVb) onder hoger onderwijs studenten uit diverse sectoren bleek dat bij het overgrote deel van de inzagemomenten alleen gelegenheid geboden werd tot inzage, maar dat verder weinig tot geen feedback en feedforward geboden werd. Deze zeer beperkte feedback was een van de grootste kritiekpunten van de studenten op de toetsing (Binsbergen, 2014). Er zijn vele methodes om studenten actiever bij deze fases te betrekken en hierbij meer feedback en feedforward te bieden.

Eén van de methodes is de door Winger (2005) ontwikkelde FSA-methode (Formative Summative Assessment), waarbij de formatieve en summatieve functie van beoordelen samenkomen. Bij deze methode worden bij de inzage van de toets alle vragen besproken waarbij de studenten actief betrokken worden. Hierbij wordt de studenten ook om feedback gevraagd over onduidelijke of misleidende vragen en wordt de itemanalyse besproken waarbij de docent motiveert waarom bepaalde vragen vervallen zijn. De methode bleek effectief: een onver-

wachte herkansing na één week leidde tot hogere scores bij de groep die via deze methode inzage had gehad dan bij controlegroep. Daarnaast geeft Winger aan dat studenten deze wijze van inzage als zeer positief ervaren en beschrijft hij de waarde van de FSA-methode voor de docent: het levert veel informatie over de toets-items, waardoor de docent deze kan verbeteren.

Een andere methode voor het betrekken van studenten bij de inzage is de Group Based Formative Assessment-methode, beschreven door Yu and Li (2014). De inzage/nabespreking wordt in groepsverband uitgevoerd waarbij de docent ervoor zorgt dat de groepen zo zijn samengesteld dat in elke groep verschillende niveaus aanwezig zijn. De studenten bespreken hun werk na en trachten gezamenlijk de goede antwoorden te vinden. Na een halfuur krijgen de groepen de goede antwoorden. Bij de controlegroep kregen leerlingen aan het begin van de inzage direct de goede antwoorden naast hun eigen werk. De methode leidde tot positieve resultaten: leerlingen waren zeer tevreden over de methode, daarnaast werd de herkansing significant beter gemaakt door de groep die deze methode toegepast had dan door de controlegroep.

Studenten benoemen als belangrijke criterium voor goede feedback in het algemeen (en dus ook voor feedforward) dat deze bruikbaar is voor toekomstige verbetering (Carless, 2006; Taras, 2001; William, 2011). Dit heeft diverse implicaties: ten eerste betekent het dat feedback en feedforward niet alleen zinvol zijn na een eindbeoordeling maar vooral ook in een eerder stadium, daar waar nog verbetering mogelijk is: halverwege het leerproces, op een tussentijdse toets, op een conceptverslag, enzovoort (Higgins et al., 2002). Daarnaast betekent dit, dat feedforward extra waardevol is als deze ook voor toekomstige onderwijsseenheden bruikbaar is. Het feit dat veel curricula thematisch zijn opgebouwd vormt hierbij een probleem (Higgins et al., 2002; Li & De

Luca, 2014). Vaak is, indien juist geformuleerd, feedforward wel bruikbaar voor andere onderwijsseenheden; door studenten wordt dit echter niet altijd herkend. Het lijkt hierbij zinvol het verband tussen de verschillende thema's te verduidelijken waardoor de bruikbaarheid van de feedforward voor toekomstige verbetering voor studenten wel zichtbaar wordt. Uiteraard is het hierbij van belang dat de docenten die bij verschillende thema's betrokken zijn, de beoordelingscriteria vergelijkbaar interpreteren en toepassen.

Discussie

In deze studie zijn diverse activiteiten besproken die kunnen worden ingezet om studenten actief te betrekken bij de toetscyclus. Hierbij blijkt dat deze cyclus heel geschikt is als kapstok voor het bieden van feedup, feedback en feedforward tijdens de verschillende fases van de toetscyclus. Door de actieve betrokkenheid van studenten neemt de formatieve waarde van de toets toe, zowel voor de student maar ook voor de docent. De toetscyclus staat momenteel volop in de aandacht, mede als gevolg van de rol die deze cyclus inneemt bij de voorgestelde BKE-trajecten. We hopen met dit artikel docenten te inspireren om bij het werken met deze cyclus de focus niet uitsluitend te leggen op het ontwerpen en afnemen van een verantwoorde toets maar eveneens de kansen te zien en benutten om tijdens deze fases meer leerwaarde voor studenten maar ook voor henzelf te creëren.

Het actief betrekken van studenten betekent niet dat we de toetsing overdragen aan de student en dat de docent hier minder verantwoordelijk voor wordt. Het betekent wel dat het leereffect en de transparantie van toetsing vergroot kan worden en dat de verantwoordelijkheid van de student hierbij toeneemt.

Vanuit de toekomstige functie-uitoefening bestaat eveneens een grote legitimatie om met studenten in gesprek te

gaan over het onderwijs en de toetsing. Zij leren hiermee verantwoordelijkheid te nemen voor hun eigen ontwikkeling. Deze houding/vaardigheid is van elementair belang voor hun ontwikkeling tijdens de rest van hun loopbaan. Het is belangrijk dat het onderwijs hen ook hierop voorbereidt.

Op basis van het voorgaande willen we de volgende aanbevelingen doen:

- Ga met studenten in gesprek over de toetsing.
- Neem studenten serieus.
- Zie de kansen die het jezelf als docent biedt.
- Betrek bij de ontwikkeling van BKE-trajecten direct de formatieve waarde die kan worden geboden tijdens het doorlopen van de toetscyclus.

Vervolgonderzoek

In deze studie zijn activiteiten beschreven die docenten kunnen inzetten om studenten actief te betrekken bij de toetsing. In hoeverre deze momenteel al worden toegepast bij Zuyd Hogeschool en elders is onduidelijk. Het gepresenteerde onderzoek vormt dan ook de basis voor twee vervolgstudies. In de eerste vervolgstudie wordt onderzocht in hoeverre bij bepaalde opleidingen binnen Zuyd Hogeschool daadwerkelijk studenten betrokken worden bij de toetscyclus en in hoeverre studenten en docenten hier een meerwaarde in zien. Vervolgens wordt een studie uitgevoerd naar de effecten van het betrekken van studenten bij de toetscyclus: hierbij zullen zowel de effecten op de formatieve waarde voor student als docent worden onderzocht, alsook de effecten op de toetscores en de motivatie en ervaren zelfregulatie door studenten.

Referenties

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115-129.
- Binsbergen, M., Verstappen, J., Selm, M. van (2014). Toetsing in het hoger onderwijs. Percepties van studenten over toetsing in het hoger onderwijs. : Bureau ICE en Onderzoeksbureau van de Landelijke

Studenten Vakbond (LSVb).

- Bron, J., & Veugelers, W. (2014). Why we need to involve our students in curriculum design: Five Arguments for Student Voice. *Curriculum & Teaching Dialogue*, 16.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58(1), 1-14.
- Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher education*, 31(2), 219-233.
- Expertgroep BKE/SKE. (2013). *Verantwoord toetsen en beslissen in het hoger beroepsonderwijs. Een voorstel voor een programma van eisen voor een basis- en seniorkwalificatie examinering (BKE/SKE)*. Vereniging Hogescholen.
- Fontana, D., & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology*, 64(3), 407-417.
- Foos, P.W., Mora, J. J., & Tkacz, S. (1994). Student study techniques and the generation effect. *Journal of Educational Psychology*, 86(4), 567.
- Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, 39(3), 193.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64.
- Janssen-Noordman, A. M., & van Merriënboer, J. J. (2002). *Innovatief onderwijs ontwerpen: via leertaken naar complexe vaardigheden*: Wolters-Noordhoff Groningen/Houten.
- Joosten-ten Brinke, D., & Sluijsmans, D. (2012). Tijd voor toetskwaliteit: het borgen van toetsdeskundigheid in examencommissies. *TH&MA*, 19(4), 16-21.
- Kohn, A. (1994). Grading: The issue is not how but why. *Educational Leadership*, 52, 38-38.
- Könings, K. D., Seidel, T., & van Merriënboer, J. J. (2014). Participatory design of learning environments: Integrating perspectives of students, teachers, and designers. *Instructional Science*, 42(1), 1-9.
- Li, J., & De Luca, R. (2014). Review of assessment feedback. *Studies in Higher education*, 39(2), 378-393.
- Macdonald, A. (2005). Student self-evaluation of coursework assignments: A route to a better perception of quality. *Learning and Teaching in Higher Education*, 1(1), 102-107.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1), 4-13.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Sluijsmans, D. M. A. (2008). *Duurzaam beoordelen in vraaggestuurd leren* (lectorale rede) Nijmegen: Hogeschool van Arnhem en Nijmegen.
- Sluijsmans, D. M. A. (2013). *Verankerd in leren: Vijf bouwstenen voor professioneel beoordelen in het hoger beroepsonderwijs* (lectorale rede). Heerlen: Zuyd Hogeschool.
- Sluijsmans, D. M. A., Joosten-ten Brinke, D., & van der Vleuten, C. P. M. (2013). *Toetsen met leerwaarde. Een reviewstudie naar de effectieve kenmerken van formatief toetsen*. Den Haag: NWO.
- Taras, M. (2001). The use of tutor feedback and student self-assessment in summative assessment tasks: Towards transparency for students and for tutors. *Assessment & Evaluation in Higher Education*, 26(6), 605-614.
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- Van den Bos, P., Burghout, C., & Joosten-ten Brinke, D. (2014). Toetsen met rubrics *Toetsen in het hoger onderwijs* (pp. 193-204): Springer.
- Van der Vleuten, C. P. M., & Schuwirth, L. W. (2005). Assessing professional competence: From methods to programmes. *Medical education*, 39(3), 309-317.
- White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16(1), 3-118.
- William, D. (2011). *Embedded formative assessment* Bloomington: Solution Tree Press.
- Wininger, S. R. (2005). Using your tests to teach: Formative summative assessment. *Teaching of Psychology*, 32(3), 164-166.
- Yu, H., & Li, H. (2014). Group-based formative assessment: A successful way to make summative assessment effective. *Theory and Practice in Language Studies*, 4(4), 839-844.