

De touwtjes in handen: toetskwaliteit in het hbo

Dit artikel is het eenenzeventigste 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

Tamara van Schilt-Mol
Dominique Slujsmans
Martijn Peters
Fedor de Beer
Linda Jakobs

Van Schilt-Mol is associate lector Toetsen en Beoordelen bij het Kenniscentrum Kwaliteit van Leren van de Hogeschool van Arnhem en Nijmegen (HAN). Peters en De Beer zijn aan dit Kenniscentrum verbonden als onderzoekers. Slujsmans is lector Professioneel Beoordelen bij Zuyd Hogeschool en Jakobs is onderwijsadviseur bij de Service Unit Onderwijs & Onderzoek van de HAN.

Reacties op dit artikel naar:
Tamara.vanSchiltMol@han.nl of
toetsinggetoetst@han.nl

Inhoud

- _ Samenvatting
- _ Inleiding
- _ De ontwerpprincipes van De Toetsing Getoetst (DTG)
- _ Kwaliteit van toetsing in DTG
- _ DTG in de praktijk
- _ Transparantie, bruikbaarheid en uitvoerbaarheid van DTG voor opleidingen
- _ Wat levert werken met DTG op (stap 1 en 2)?
- _ Naar een herontwerp van DTG
- _ Conclusie

Box 1: De kwaliteitspiramide van eigentijds toetsen en beoordelen met beknopte definities

Box 2: Fasen van ontwikkeling in de weg naar duurzaam werken aan toetsing

Box 3: Kwaliteitscyclus analyse, verbetering en borging toetsing

Box 4: Uitwerking van de piramidelaag toetsprogramma

Box 5: Positie- en ambitiebepaling van een fictieve opleiding

Box 6: Toetsweb

Samenvatting

De Toetsing Getoetst (DTG) is een methodiek die opleidingen in staat stelt de regie te nemen bij het analyseren, verbeteren en borgen van toetskwaliteit. Met deze methodiek kunnen opleidingen op systematische wijze vaststellen hoe de huidige toetskwaliteit is (positiebepaling) en wat

de wenselijke toetskwaliteit is (ambitiebepaling). Omdat de methodiek een aantal concrete opdrachten bevat, wordt opleidingen houvast geboden bij het formuleren van een eigen visie op toetskwaliteit en zijn er handvatten om zelf de kwaliteitsprocessen rondom toetsing te verbeteren. In het studiejaar 2015-2016 is in een aantal pilots ervaring opgedaan met de methodiek en is de bruikbaarheid, transparantie en uitvoerbaarheid ervan onderzocht. Doel van dit artikel is deze opbrengsten te delen en toe te lichten hoe deze hebben geleid tot een herontwerp van de DTG-methodiek.

Inleiding

Kwaliteit van onderwijs betekent ook kwaliteit van toetsing: toetsen worden immers gebruikt om verantwoorde beslissingen over de beroepsbekwaamheid van studenten te nemen (Slujsmans, Van Schilt-Mol, Peeters & Joosten-Ten Brinke, 2015). Om zorgvuldige beslissingen te kunnen nemen en misclassificatie te voorkomen is een rijke verzameling aan bewijslast nodig die door bekwame beoordelaars wordt beoordeeld. De kwaliteit van toetsing in het hoger beroepsonderwijs is echter - naar aanleiding van een aantal incidenten rondom alternatieve afstudeertrajecten - in 2011 in twijfel getrokken. In 2012 is het rapport *Vreemde Ogen Dwingen* opgeleverd om hogescholen te stimuleren aandacht te besteden aan toetskwaliteit (Commissie externe validering examenkwaliteit, 2012). De verscherpte aandacht

voor toetsing heeft er ook toe geleid dat in accreditaties steeds meer nadruk wordt gelegd op toetsing. In 2014 zijn toetsing en gerealiseerde eindkwalificaties als aparte entiteiten opgenomen in de accreditatiesystematiek. Hiermee staan binnen opleidingen steeds vaker vragen centraal als: wat betekent kwaliteit van toetsing, wanneer is het systeem van toetsing adequaat, hoe goed is ons systeem van toetsen, waar liggen verbeterpunten en hoe houden we eigenaarschap over ons systeem van kwaliteitszorg?

Om deze vragen te kunnen beantwoorden en opleidingen een handvat te bieden zelf (weer) de regie te nemen over de kwaliteit van toetsing binnen de opleiding, is in 2011 op de Hogeschool van Arnhem en Nijmegen (HAN) door het lectoraat Toetsen en Beoordelen (voorheen het lectoraat Duurzaam Beoordelen) en de Service Unit Onderwijs & Onderzoek (SU O&O) het initiatief genomen om een methodiek te ontwikkelen waarmee betrokkenen in het hoger onderwijs de kwaliteit van toetsing kunnen analyseren, ver-

beteren en borgen. Daarmee sluiten we aan bij het idee van professionals governance: opleidingen hebben zelf systemen waarmee ze hun kwaliteit kunnen aan- en bijsturen (Huisman & De Vijlder, 2012). De basis van deze methodiek is vier jaar geleden gepubliceerd in dit tijdschrift (zie Sluijsmans, Peeters, Jakobs & Weijzen, 2012) en heeft geresulteerd in een eerste prototype van de methodiek De Toetsing Getoetst. Dit prototype is beschreven in het handboek *Kwaliteit van toetsing onder de loep* (Sluijsmans, Joosten-Ten Brinke, & Van Schilt-Mol, 2015).

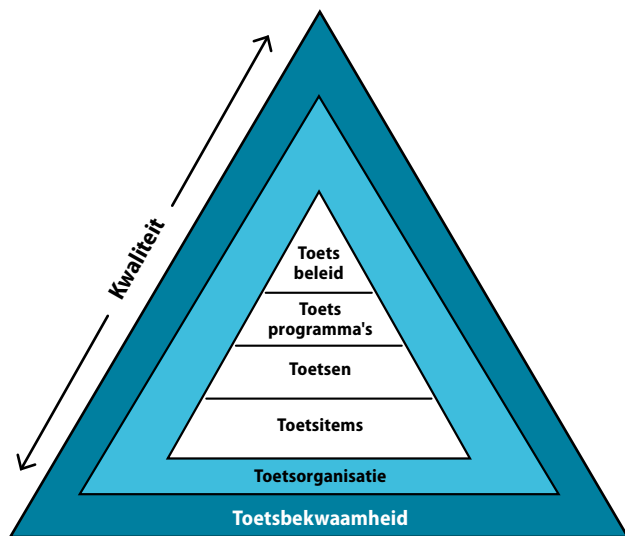
In een samenwerkingsverband tussen het lectoraat Toetsen en Beoordelen (HAN), het lectoraat Professioneel Beoordelen (Zuyd Hogeschool) en de Service Unit Onderwijs & Onderzoek (HAN) is dit prototype in studiejaar 2014-2015 bij negentien opleidingen getest met als doel de methodiek verder te verfijnen. Met de aangepaste methodiek is vervolgens in studiejaar 2015-2016 een onderzoekspilot uitgevoerd bij zeven opleidingen. Centrale vraag in dit onderzoekstraject was: helpt

de DTG-methodiek opleidingen om positie en ambitie ten aanzien van hun kwaliteit van toetsing te bepalen en deze te vertalen naar concrete doelen? In dit praktisch artikel wordt op deze vraag een antwoord gegeven en wordt de definitieve methodiek gepresenteerd.

De ontwerpprincipes van DTG

Een belangrijk doel van DTG is opleidingen houvast bieden in het ontwikkelen van een eigen visie op toetskwaliteit en het komen tot een plan om deze visie en de daaraan gekoppelde doelen te realiseren. De methodiek is daarmee nadrukkelijk gericht op *zelfevaluatie*: opleidingen bepalen - op basis van hun visie op opleiden en toetsen - zelf de mate waarin zij duurzaam werken aan het realiseren van de gewenste kwaliteit van toetsing en formuleren op basis van deze analyse verbeterpunten. Kwaliteit van toetsing is in de methodiek geoperationaliseerd met de kwaliteitspiramide van toetsen en beoordelen (Joosten-Ten Brinke, 2011; Sluijsmans, Peeters, Jakobs & Weijzen,

BOX 1: DE KWALITEITSPIRAMIDE VAN EIGENTIJD TOETSEN EN BEOORDELEN MET BEKNOPTIE DEFINITIES (Sluijsmans, van Schilt-Mol, Peeters & Joosten-ten Brinke, 2015, p. 20-21)



Toetsbeleid

Met toetsbeleid wordt het geheel aan vastgestelde afspraken, zowel inhoudelijk als procedureel, over toetsen en beoordelen bedoeld.

Toetsprogramma

Een toetsprogramma is een bewuste en beargumenteerde combinatie van toetsvormen, passend bij de doelen en opbouw van een opleiding.

Toetsen

Een toets is een meetinstrument dat in het onderwijs wordt ingezet om na te gaan of de doelen van onderwijs door studenten worden bereikt.

Toetstaken

Een toetstaak is een item of opdracht binnen een toets waarmee studenten worden uitgedaagd hun kennis en vaardigheden te tonen.

Toetsorganisatie

Toetsorganisatie verwijst naar de wijze waarop docenten, examencommissie, toetscommissie, management en ondersteunende medewerkers doelgericht met elkaar samenwerken om de gewenste toetskwaliteit te realiseren.

Toetsbekwaamheid

Toetsbekwaamheid is de deskundigheid die bij opleidingen aanwezig moet zijn om te zorgen voor kwaliteit van toetsing op alle lagen van de kwaliteitspiramide.

BOX 2: FASEN VAN ONTWIKKELING IN DE WEG NAAR DUURZAAM WERKEN AAN TOETSING
(gebaseerd op Brouwer & Boer, 2013)

Activiteiten-georiënteerd



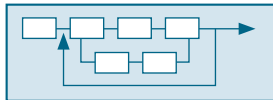
Individuele actoren in de opleiding voeren op ad-hoc basis losstaande activiteiten uit.

Proces-georiënteerd



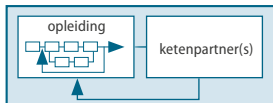
Relevante actoren in de opleiding voeren vanuit een korte termijnbeleid toetsactiviteiten uit die onderling samenhangen.

Systeem-georiënteerd



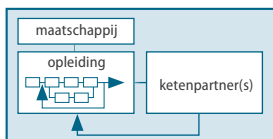
Alle relevante actoren in de opleiding voeren vanuit een middellangetermijnbeleid toetsactiviteiten uit die onderling samenhangen en geïntegreerd zijn in lopende processen in de opleiding.

Keten-georiënteerd



Alle relevante actoren in de opleiding voeren vanuit een middellangetermijnbeleid, samen met ketenpartners, toetsactiviteiten uit die onderling samenhangen en geïntegreerd zijn in lopende processen in de opleiding en de keten.

Maatschappij-georiënteerd



Alle relevante actoren in de opleiding voeren vanuit een langetermijnbeleid, samen met ketenpartners en de maatschappij, toetsactiviteiten uit die onderling samenhangen en geïntegreerd zijn in lopende processen in de opleiding en de keten/maatschappij.

2012). Hierin wordt kwaliteit van toetsing uitgewerkt in zes entiteiten: 1) kwaliteit van toetstaken, 2) kwaliteit van toetsen, 3) kwaliteit van toetsprogramma's, 4) kwaliteit van het toetsbeleid, 5) kwaliteit van de toetsorganisatie en 6) kwaliteit van toetsbekwaamheid (zie box 1). De kwaliteit van toetsing wordt bepaald door de mate van samenspel tussen de verschillende entiteiten: toetskwaliteit is zo sterk als de zwakste schakel uit de piramide. De visie die aan DTG ten grondslag ligt is dat er altijd aan toetskwaliteit wordt gewerkt. Om de mate waarin opleidingen *duurzaam* werken aan de kwaliteit van toetsing inzichtelijk te maken is inspiratie ontleend aan het Auditing Instrument for Sustainability in Higher Education (AISHE) (Brouwer & Boer, 2013; DHO, 2010). AISHE is gericht op het duurzaam werken aan ontwikkeling en het stimuleren van lange termijn denken: hoe komen we van de huidige situatie naar de gewenste situatie (Van Schilt-Mol, Peeters, Sluijsmans & Jakobs, 2015). Voor het formuleren van de

ontwerpprincipes van DTG is dan ook aangesloten bij vier kenmerken van AISHE: 1) zelfevaluatie, 2) fasen van ontwikkeling, 3) gestandaardiseerde set van elementen en 4) het doorlopen van de PDCA-cyclus.

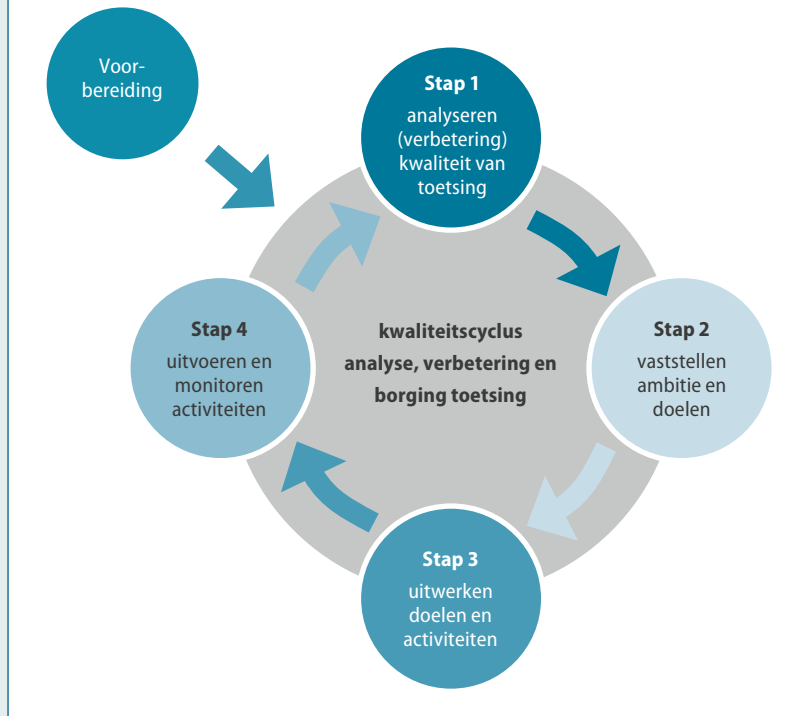
Zelfevaluatie: een zelfevaluatie biedt opleidingen bij uitstek de mogelijkheid om zelf de regie te nemen op de kwaliteit van toetsing. Een zelfevaluatie heeft een grote impact doordat de deelnemers zelf de evaluatiecriteria gebruiken en moeten begrijpen om te kunnen evalueren. Bovendien resulteert de evaluatie niet alleen in verbeterpunten, maar treedt er ook een bewustwordingsproces op (Dijkstra & Baartman, 2011).

Fasen van ontwikkeling: binnen AISHE worden vijf ontwikkelingsfasen onderscheiden: 1) activiteiten-georiënteerd, 2) proces-georiënteerd, 3) systeem-georiënteerd, 4) keten-georiënteerd en 5) maatschappij-georiënteerd. Kenmerkend binnen deze fasen is een ontwikkeling in de wijze waarop actoren ingezet worden

(van individuele acties naar acties waarbij alle relevante actoren betrokken zijn), een ontwikkeling in het langetermijndenken (van ad-hoc naar een langetermijnstrategie) en een ontwikkeling in de integratie van processen in het onderwijsproces (van losstaande activiteiten naar activiteiten die bewust in samenhang worden uitgevoerd). Voor DTG is het kader van AISHE vertaald naar een kader voor toetsing (zie box 2).

Gestandaardiseerde set van elementen: voor iedere piramidelaag is een set van drie elementen benoemd op basis waarvan een analyse kan worden uitgevoerd. Het eerste element, kwaliteitscriteria, gaat over de vraag welke criteria de opleiding gebruikt om de kwaliteit van de lagen vast te stellen. Vragen die hierbij aan de orde komen zijn bijvoorbeeld: welke criteria worden gebruikt om de kwaliteit van toetsen en toetstaken vast te stellen? Wat zijn de richtlijnen bij het opstellen van het toetsbeleid? Het tweede element, ontwerp, is gericht op de vraag hoe de zes pi-

BOX 3: KWALITEITSCYCLUS ANALYSE, VERBETERING EN BORGING TOESTING
(Van Schilt-Mol, Peeters, Sluijsmans & Jacobs, 2015, p. 35)



ramidelagen aan de hand van deze kwaliteitscriteria zijn ontworpen. Vragen die bij de analyse kunnen worden gesteld zijn: hoe worden toetsen en toetstaken ontworpen en door welke actoren? Hoe worden docenten geprofessionaliseerd in toetsbekwaamheid? Het derde element, kwaliteitsborging, richt zich op de wijze waarop de kwaliteitsborging van de piramidelagen wordt ingericht. Voorbeelden van vragen zijn: wie stelt vast of interventies uit het toetsbeleid zijn behaald? Hoe worden verbeteracties geformuleerd en geïmplementeerd?

Doorlopen van de PDCA-cyclus: bij het ontwerp van de methodiek is aangesloten bij de kwaliteitscyclus van Deming (1986), waarbij vier activiteiten worden onderscheiden die gericht zijn op kwaliteitsverbetering: Plan, Do, Check en Act. DTG is opgebouwd uit vier methodische stappen die te koppelen zijn aan deze PDCA-cyclus (zie box 3). In de Plan-fase staat de ontwikkeling en de verbetering

van de kwaliteit van toetsing centraal. In deze fase vindt dan ook eerst de positiebepaling plaats (stap 1), vervolgens de ambitiebepaling (stap 2). Op basis van de ambitie worden reële doelen en activiteiten benoemd (stap 3) die in de Do-fase worden uitgevoerd (stap 4). In de Check-fase worden de resultaten van deze activiteiten gemeten en vergeleken met de gewenste fase (stap 1 en 2). In de Act-fase tot slot, wordt bijgestuurd op de verschillen tussen de (in de Plan-fase uitgesproken) gewenste situatie en de (in de Do-fase bereikte) daadwerkelijke situatie (stap 3 en 4). Concreet betekent dit dus dat bij de uitvoering van DTG de PDCA-cyclus tweemaal wordt doorlopen.

Kwaliteit van toetsing in DTG

De kwaliteitspiramide van toetsen en beoordelen, het AISHE-model en de vier ontwerpprincipes hebben de basis gevormd voor een rubric om positie en ambitie ten aanzien van toetskwaliteit te be-

palen: voor ieder element binnen de piramidelagen is voor elke ontwikkelingsfase een beknopte tekst ontwikkeld die de mate van kwaliteit omschrijft. In box 4 is een uitwerking van de piramidelaag toetsprogramma opgenomen. Een relevante en veel gestelde vraag is wanneer er sprake is van voldoende toetskwaliteit. Binnen de methodiek maken we onderscheid tussen enerzijds het *duurzaam werken* aan de kwaliteit van toetsing en anderzijds het geven van een *kwalitatief goede invulling* aan toetsing. Er wordt duurzaam gewerkt aan het realiseren van kwaliteit op alle lagen van de piramide als alle relevante actoren in de opleiding vanuit een middellangetermijnperspectief volgens de PDCA-cyclus activiteiten uitvoeren die zijn geïntegreerd in lopende processen binnen de opleiding (Van Schilt-Mol & Sluijsmans, 2015). Het realiseren van een kwalitatief goede invulling aan het ontwerp, de uitvoering en de borging van de verschillende piramidelagen is uitsluitend mogelijk door het vaststellen van kwaliteitscriteria voor deze lagen (Van Schilt-Mol & Sluijsmans, 2015). Aangezien opleidingen zelf verantwoordelijk zijn voor het vaststellen van deze criteria en hier ook eigen keuzes in kunnen, of zelfs moeten maken, kan de inhoudelijke invulling van wat verstaan wordt onder een 'kwalitatief goede invulling' per opleiding verschillen.

Voor het realiseren van kwaliteit wordt opleidingen aanbevolen om te streven naar ten minste de systeem-georiënteerde fase: in deze fase is de kans dat de kwaliteit van bijvoorbeeld toetsen of het toetsprogramma berust op toeval of op de kwaliteiten van die professionele docent die hiermee individueel aan de slag is gegaan immers het kleinste: in de systeem-georiënteerde fase zijn er gezamenlijk kwaliteitscriteria vastgesteld, wordt er ontworpen en geëvalueerd aan de hand van deze criteria en vinden alle toetsingsprocessen in samenhang plaatst. Echter, gezien de context van het hoger beroeps-onderwijs, waar opleidingen en beroepenveld zich gezamenlijk verantwoorde-

BOX 4: UITWERKING VAN DE PIRAMIDELAAG TOETSPROGRAMMA

Kwaliteitscriteria				
<i>Activiteiten-georiënteerd</i>	<i>Proces-georiënteerd</i>	<i>Systeem-georiënteerd</i>	<i>Keten-georiënteerd</i>	<i>Maatschappij-georiënteerd</i>
Bij het ontwerpen en evalueren wordt gebruik gemaakt van eigen of impliciet bekend veronderstelde kwaliteitscriteria. De keuze voor deze criteria wordt niet onderbouwd.	Bij het ontwerpen en evalueren wordt gebruik gemaakt van kwaliteitscriteria die op opleidingsniveau zijn vastgesteld. De keuze voor de criteria is gebaseerd op ervaringen van relevante actoren in de opleiding.	Bij het ontwerpen en evalueren wordt gebruik gemaakt van kwaliteitscriteria die op opleidingsniveau zijn vastgesteld. De keuze voor de criteria is gebaseerd op (wetenschappelijke) literatuur. De keuze is eveneens getoetst aan ervaringen van relevante actoren in de opleiding.	Bij het ontwerpen en evalueren wordt gebruik gemaakt van kwaliteitscriteria die opleidingsniveau zijn vastgesteld. De keuze voor de criteria is gebaseerd op (wetenschappelijke) literatuur. De keuze is eveneens getoetst aan ervaringen van relevante actoren in de opleiding, ketenpartners (zoals het werkveld) en/of externe deskundigen.	Bij het ontwerpen en evalueren wordt gebruik gemaakt van kwaliteitscriteria die opleidingsniveau zijn vastgesteld. De keuze voor de criteria is gebaseerd op (wetenschappelijke) literatuur. De keuze is eveneens getoetst aan ervaringen van relevante actoren in de opleiding, ketenpartners en/of externe deskundigen en collega-instellingen.
Ontwerp				
<i>Activiteiten-georiënteerd</i>	<i>Proces-georiënteerd</i>	<i>Systeem-georiënteerd</i>	<i>Keten-georiënteerd</i>	<i>Maatschappij-georiënteerd</i>
De opleiding beschikt over een overzicht van alle toetsen in de opleiding. De relatie tussen de toetsen wordt hierin niet beschreven/onderbouwd. Er vindt, met andere woorden, geen ontwikkeling plaats van een samenhangend toetsprogramma.	Het ontwerp en de evaluatie van het toetsprogramma wordt uitgevoerd door relevante actoren in de opleiding. Het programma wordt horizontaal ontworpen op basis van de opleidingskwalificaties. Bij het ontwerpen heeft afstemming plaatsgevonden over de uitvoerbaarheid van het programma.	Het ontwerp en de evaluatie van het toetsprogramma wordt uitgevoerd door alle relevante actoren in de opleiding. Het programma wordt horizontaal en verticaal ontworpen op basis van de opleidingskwalificaties en de uitwerking daarvan in niveaus. Bij het ontwerpen heeft afstemming plaatsgevonden over de balans tussen de uitvoerbaarheid van het programma, de meest geschikte toetsvormen, de functies van deze toetsen en de hoeveelheid toetsen.	Het ontwerp en de evaluatie van het toetsprogramma wordt uitgevoerd door alle relevante actoren in de opleiding. Ketenpartners en/of externe deskundigen en collega-instellingen worden hierbij betrokken. Het programma wordt horizontaal en verticaal ontworpen op basis van de opleidingskwalificaties en de uitwerking daarvan in niveaus. Bij het ontwerpen heeft afstemming plaatsgevonden over de balans tussen de uitvoerbaarheid van het programma, de meest geschikte toetsvormen, de functies van deze toetsen en de hoeveelheid toetsen.	Het ontwerp en de evaluatie van het toetsprogramma wordt uitgevoerd door alle relevante actoren in de opleiding. Ketenpartners en/of externe deskundigen en collega-instellingen worden hierbij betrokken. Het programma wordt horizontaal en verticaal ontworpen op basis van de opleidingskwalificaties en de uitwerking daarvan in niveaus. Bij het ontwerpen heeft afstemming plaatsgevonden over de balans tussen de uitvoerbaarheid van het programma, de meest geschikte toetsvormen, de functies van deze toetsen en de hoeveelheid toetsen.
Kwaliteitsborging				
<i>Activiteiten-georiënteerd</i>	<i>Proces-georiënteerd</i>	<i>Systeem-georiënteerd</i>	<i>Keten-georiënteerd</i>	<i>Maatschappij-georiënteerd</i>
Evaluatie vindt ad-hoc plaats. Op basis van de evaluatie worden incidentele aanpassingen gedaan.	Evaluatie vindt incidenteel plaats. Het resultaat van de evaluatie wordt besproken met relevante actoren in de opleiding en aanpassingen worden doorgevoerd.	Evaluatie vindt structureel en systematisch plaats als onderdeel van de kwaliteitscyclus van toetsing. Het resultaat van de evaluatie wordt besproken met alle relevante actoren in de opleiding en aanpassingen worden doorgevoerd.	Evaluatie vindt structureel en systematisch plaats als onderdeel van de kwaliteitscyclus van toetsing. Het resultaat van de evaluatie wordt besproken met alle relevante actoren in de opleiding ,met ketenpartners en/of externe deskundigen en collega-instellingen, en aanpassingen worden doorgevoerd.	Evaluatie vindt structureel en systematisch plaats als onderdeel van de kwaliteitscyclus van toetsing. Het resultaat van de evaluatie wordt besproken met alle relevante actoren in de opleiding , met ketenpartners (zoals het werkveld) en/of externe deskundigen, en aanpassingen worden doorgevoerd.

lijk zouden moeten voelen voor het vaststellen en bereiken van de eindkwalificaties van de opleiding, en gezien de maat-

schappelijke waarde van de diploma's die binnen het hoger beroepsonderwijs worden uitgegeven, zouden opleidingen

voor het realiseren van duurzaam werken aan de kwaliteit van toetsing moeten streven naar de maatschappij-georiën-

teerde fase: in deze fase worden zowel het beroepenveld als maatschappelijke partners - zoals collega-instellingen en externe experts - daar waar relevant betrokken bij de inrichting van het onderwijs en de toetsing hiervan.


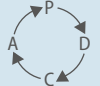



DTG in de praktijk

Een kernteam met leden uit de opleiding doorloopt de stappen van DTG onder leiding van een of twee begeleiders. In de pilots werd de externe begeleiding uitgevoerd door een lid van het onderzoeksteam zodat de condities van de uitvoering bij de verschillende opleidingen zo vergelijkbaar mogelijk waren. Tijdens de voorbereidende fase wordt het kernteam (dat bestaat uit tien tot vijftien personen) samengesteld. Aan dit kernteam nemen sleutelfiguren uit de opleiding op het gebied van toetsing deel, zoals leden van de examen-, toets- en curriculumcommissie, docenten, studenten en leden van het management. In vier dagdelen doorloopt het kernteam onder leiding van een externe begeleider de eerste twee stappen

uit de kwaliteitscyclus (zie box 3). In het eerste dagdeel stelt het kernteam de huidige fase van ontwikkeling vast op elk van de lagen uit de toetspiramide. Om tot deze positiebepaling te komen beschrijft het kernteam de huidige manier van werken met een casusbeschrijving. Ondersteunende materialen hierbij zijn de beknopte toelichting per piramidelaag (zie box 1), de karakteristieke eigenschappen van de vijf fasen van ontwikkeling (zie box 2) en een aantal hulpvragen per piramidelaag. Twee à drie deelnemers stellen per piramidelaag de casusbeschrijving vast, die vervolgens plenair in het kernteam besproken en aangescherpt wordt. Tot slot legt het kernteam op basis van de casusbeschrijving en bewijsvoering de huidige positie vast door consensus te bereiken over welke fase van ontwikkeling het meeste recht doet aan de huidige situatie. In het tweede dagdeel stelt het kernteam de ambities over de gewenste fase van ontwikkeling vast. Hierbij wordt benoemd welke concrete resultaten zicht-

baar zouden moeten zijn als de gewenste fase van ontwikkeling is bereikt en worden doelen geformuleerd. In box 5 is een positie- en ambitiebepaling van een fictieve opleiding weergegeven. Het kernteam beschrijft de argumentatie voor de positie- en ambitiebepaling in een werkboek. Dit werkboek heeft hiermee ook een functie als naslagwerk waarin informatie is opgenomen over de vastgestelde beginsituatie, de gewenste situatie en de doelen waar de opleiding aan wil gaan werken. In het derde dagdeel worden de doelen per piramidelaag bij elkaar gebracht, worden de verbanden tussen de piramidelagen en de doelen benoemd en worden de doelen geprioriteerd. Om de doelen te prioriteren die het beste passen bij de opleiding formuleert het kernteam een kijkkader. Dit kijkkader geeft hen antwoord op vragen als: Waar moeten de doelen bij aansluiten? Wat vinden jullie gezien jullie context nu het belangrijkste? Hoe past dit bij de condities waarbinnen jullie moeten werken? Met het kijkkader selecteert het kernteam drie à vier doelen

BOX 5: POSITIE- EN AMBITIEBEPALING VAN EEN FICTIEVE OPLEIDING

piramidelagen	elementen	ontwikkeling kwaliteit van toetsing				
		fasen van ontwikkeling				
		Activiteiten-georiënteerd	Proces-georiënteerd	Systeem-georiënteerd	Keten-georiënteerd	Maatschappij-georiënteerd
 toets(tak)en	kwaliteitscriteria					
	ontwerp					
	kwaliteitsborging					
 programma	kwaliteitscriteria					
	ontwerp					
	kwaliteitsborging					
 beleid	kwaliteitscriteria					
	ontwerp					
	kwaliteitsborging					
 organisatie	kwaliteitscriteria					
	ontwerp					
	kwaliteitsborging					
 bekwaamheid	kwaliteitscriteria					
	ontwerp					
	kwaliteitsborging					

huidige positie
 gewenste positie
 samenvallen huidige en gewenste positie

die zij uitwerkt in een beknopt actieplan. De positie- en ambitiebepaling en de gestelde doelen worden vervolgens in de vorm van een poster opgeleverd. In het vierde dagdeel vindt aan de hand van deze poster terugkoppeling plaats aan het volledige opleidingsteam en andere mogelijke relevante actoren. Eventueel vindt in dit dagdeel ook een aanzet plaats tot verdere concretisering van het actieplan. Na afronding van de vier dagdelen hebben de opleidingen de eerste twee stappen uit de kwaliteitscyclus van DTG uitgevoerd. De stappen drie en vier voeren zij zelfstandig uit.

Transparantie, bruikbaarheid en uitvoerbaarheid van DTG voor opleidingen

In het voorafgaande is uitgelegd wat de onderliggende ontwerpprincipes zijn van de DTG-methodiek en welke stappen de opleidingen doorlopen. In het studiejaar 2015-2016 is de waarde van de methodiek bij zeven pilotopleidingen onderzocht met als centrale vraag: Helpt de methodiek opleidingen om positie- en ambitie te aanzien van hun kwaliteit van toetsing te bepalen en deze te vertalen naar concrete doelen? Om deze vraag te beantwoorden zijn de volgende onderzoeksinstrumenten ingezet: observaties tijdens bijeenkomsten door een onderzoeker of onderwijsadviseur; een digitale vragenlijst over waarde van de methodiek en gepercipieerde effecten; telefonische interviews met de betrokken onderwijsadviseurs; analyse van de werkboeken en posters om zicht te krijgen op de wijze waarop opleidingen komen tot een positie- en ambitiebepaling, en op de kwaliteit van de uitgewerkte doelen en activiteiten. De waarde van de methodiek is geoperationaliseerd in drie variabelen: transparantie, bruikbaarheid en uitvoerbaarheid.

Transparantie verwijst naar de mate waarin de werkwijze van de methodiek inzichtelijk is voor de gebruikers. Uit het onderzoek blijkt dat de methodiek en de onder-

liggende werkvormen duidelijk en toepasbaar zijn. In de werkboeken is te zien dat alle opleidingen hun positie en ambitie hebben bepaald en daar doelen aan hebben gekoppeld. Door het sturende format van het werkboek zijn de positiebepalingen expliciet beargumenteerd. Hoewel het format hier wel om vraagt, onderbouwen niet alle opleidingen de gekozen huidige fase van ontwikkeling met bronnen, zoals bijvoorbeeld opleidingsdocumenten. De argumentatie achter de ambitiebepaling blijft impliciet, doordat in het werkboek niet wordt gevraagd deze op te schrijven. De pilotdeelnemers waarden de actieve werkvormen, omdat deze hen op een speelse en interactieve manier laten kennismaken met de soms toch wel onbekende en/of lastige terminologie. Een concreet voorbeeld hiervan is het memoryspel dat tijdens het eerste dagdeel wordt gebruikt om de karakteristieken van de verschillende ontwikkelingsfasen (van activiteiten-georiënteerd naar maatschappij-georiënteerd) toe te lichten. Hoewel de deelnemers aangeven de terminologie lastig te vinden, blijkt zowel uit de werkboeken als tijdens de terugkoppeling van de opbrengsten in het vierde dagdeel dat de leden van de kernteams de termen wél correct gebruiken. Dit lijkt erop te duiden dat de deelnemers door het werken met de methodiek de terminologie in de vingers hebben gekregen en hiermee kunnen werken.

Bruikbaarheid verwijst naar de mate waarin de methodiek leidt tot het stellen van concrete doelen gericht op duurzame ontwikkeling van toetskwaliteit. Uit de werkboeken blijkt dat de argumentatie achter de doelen waar de opleidingen mee aan de slag willen impliciet blijft doordat in het werkboek niet wordt gevraagd deze te expliciteren. Dit betekent niet dat hier geen (goede) argumenten onder liggen, zo blijkt uit observatie, maar deze zijn slechts benoemd in het gesprek. Ook blijkt het moeilijk om te achterhalen of er sprake is van valide am-

bities en argumenten: sluiten de doelen daadwerkelijk aan bij de visie van de opleiding, passen de activiteiten binnen de context en dragen ze voldoende bij tot het bereiken van de gewenste situatie? Uit de vragenlijstuitkomsten en interviews blijkt dat de deelnemers zelf vinden dat de methodiek de verwachtingen waarmaakt: zij geven aan dat zij worden gestimuleerd over de eigen grenzen heen te kijken, positie te bepalen en vandaaruit een ambitie te formuleren. Dit werkte naar hun eigen zeggen, mede dankzij de input vanuit de methodiek, kwaliteitsverhogend.

Uitvoerbaarheid verwijst naar de mate waarin de methodiek gezien de context en condities van de opleiding te organiseren is. Uit het onderzoek komt naar voren dat er enkele factoren zijn die het kunnen werken met de methodiek beïnvloeden. Een eerste factor is de *samenstelling en expertise van het kernteam*. De methodiek vraagt van de deelnemers uit het kernteam basiskennis over toetsing en diepgaandere kennis over de lagen van de piramide. Daarnaast vereist de methodiek dat de deelnemers van het kernteam goed op de hoogte zijn van de huidige situatie van toetsing binnen de opleiding. Het kiezen van een dwarsdoorsnede van de opleiding in het kernteam is zinvol, maar kan beperkend werken bij het uitwerken van de positie van een piramide-laag in kleinere groepjes als hierbij niet alle leden voldoende op de hoogte zijn van de stand van zaken binnen deze piramide-laag. Dit heeft uiteraard effect op de kwaliteit en de mate waarin de beschrijving recht doet aan de daadwerkelijke positie.

Een tweede factor die de uitvoerbaarheid bepaalt is de *aanwezigheid en mate van participatie aan kernteam*. Het primaire onderwijsproces heeft altijd voorrang. Consequentie was dat tijdens de uitvoering van DTG niet altijd het volledige kernteam aanwezig was. De - in een aantal gevallen voorkomende - onrust en wisselingen in teamsamenstelling leidde

soms tot wendingen in de visie op de positie- en/of ambitiebepaling. Facilitering om steeds deel te kunnen nemen lijkt dan ook geboden. Daarnaast hebben ook de (verschillen in de) mate van inbreng tussen deelnemers, het ervaren gevoel van betrokkenheid (sommige deelnemers waren verplicht tot deelname zonder goede informatie over doel van de pilot) en (het ontbreken van) veiligheid om plenair de stand van zaken te bespreken en dus ook de huidige knelpunten te benoemen, een duidelijke invloed op de waardering en de ervaren opbrengsten van de methodiek.

Een derde belangrijke factor betreft de *kwaliteit van de begeleiders*. Hoewel er sprake is van een zelfevaluatie komt uit het onderzoek naar voren dat het raadzaam lijkt deze onder begeleiding van een externe begeleider uit te voeren. Deelnemers geven aan dat dit ervoor zorgt dat zij gestructureerd de methodiek doorlopen, maar vooral ook worden uitgedaagd diepgaander te kijken naar de stand van zaken dan zij wellicht gedaan zouden hebben zonder 'vreemde ogen'. De begeleiders moeten niet alleen kunnen sturen op groepsprocessen, maar moeten daarnaast ook diepgaande kennis hebben van alle toetsentiteiten (toetsen, toetsstaken, toetsprogramma, toetsbeleid, toetsorganisatie en toetsbekwaamheid). De kwaliteit van de methodiek rust voor een belangrijk deel op de begeleider: de (inhoudelijke) interactie tussen het kernteam en de begeleider beïnvloedt de kwaliteit van de positie- en ambitiebepaling, van de argumenten die hierbij gebruikt worden en van de doelen die worden gesteld.

Wat levert werken met DTG op (stap 1 en 2)?

Naast uitspraken over de bruikbaarheid, transparantie en uitvoerbaarheid, levert het onderzoek naar DTG ook veel inzicht in de gepercipieerde effecten. Deze effecten zijn weliswaar niet vastgesteld en bevestigd in aanvullend onderzoek, maar

bieden wel een mooi vooruitblik voor vervolgonderzoek naar de effectiviteit van de methodiek. In de vragenlijsten en de interviews worden door de deelnemers aan de zeven pilots veel extra opbrengsten genoemd. De observaties bevestigen dit.

Een eerste opbrengst is het *gemeenschappelijk vocabulaire*. De methodiek geeft de deelnemers een begrippenkader. De semantische verwarring in toetsbegrippen is groot. Zo wordt bijvoorbeeld het begrip toetsplan zowel gebruikt ter aanduiding van het toetsprogramma als het toetsbeleid. De methodiek geeft de deelnemers daarnaast de taal om de verschillende fasen in ontwikkeling te kunnen duiden: de vijf ontwikkelingsfasen, de drie elementen en de karakterisering hiervan in de rubric helpen hen om te kunnen argumenteren waarom er sprake is van een bepaalde fase van ontwikkeling.

Een tweede opbrengst betreft het werken aan een *gemeenschappelijke visie op toetsing*. Tijdens de uitvoering van DTG wordt vaak inzichtelijk dat er ofwel geen visie op toetsing is geformuleerd, ofwel dat er sprake is van een impliciet bekend veronderstelde visie die in de praktijk bij de verschillende deelnemers verschillende invullingen blijkt te hebben. Doordat de deelnemers tijdens de uitvoering met elkaar in gesprek gaan over de kwaliteitscriteria van toetsing vormen zij gaandeweg een (aanzet tot een) gemeenschappelijke visie op toetsing, of formuleren zij als doel deze te gaan formuleren.

Een andere opbrengst is de professionalisering in toetsing. Voor veel deelnemers van DTG is deelname aan de methodiek een vorm van professionalisering geweest. Veelal zijn deelnemers wel bekend met bijvoorbeeld kwaliteitscriteria voor toetsen, over de piramidelagen toetsprogramma, toetsorganisatie en toetsbeleid blijkt minder kennis voorhanden.

Bovendien worden de deelnemers zich bewust van de complexiteit van toetsing: het gaat niet alleen om het toetsproces (het maken en afnemen van toetsen), maar ook om het formuleren van goed

toetsbeleid, het inrichten van een toetsorganisatie en het realiseren van toetsbekwaamheid bij alle relevante actoren. Juist de verbondenheid van de piramidelagen voor het realiseren van duurzame kwaliteit van toetsing is voor velen een eyeopener. Bijkomende voordeel was dat er ook nu feitelijk tijd was om als team te spreken over de visie op opleiden en toetsen en de kwaliteit van toetsing.

Tot slot versterkt de methodiek de *bewustwording van (het belang van) toetsing binnen het opleiden*. Hoewel de leden van het kernteam vaak een belangrijke rol spelen bij toetsing binnen de opleiding, zien velen van hen toetsing vooral als een verantwoordingsinstrument. Doordat de deelnemers tijdens DTG veel met elkaar in gesprek zijn over de visie op toetsing, toetsprogramma's en het toetsbeleid vindt er een bewustwording plaats van het belang van toetsing voor de ontwikkeling van studenten en het bereiken/dekken van de eindkwalificaties. Naast het verkrijgen van zicht op de mate waarin duurzaam gewerkt wordt aan kwaliteit van toetsing op alle lagen van de piramide, wordt ook inzichtelijk welke processen er binnen de opleiding al georganiseerd zijn, welke kennis er aanwezig is en welke rollen en verantwoordelijkheden er op het gebied van toetsing zijn belegd. Tijdens de uitvoering van DTG blijkt vaak dat er meer beschikbaar is dan deelnemers weten (bijvoorbeeld toetsprotocollen) en blijkt regelmatig dat deelnemers van elkaar niet weten over welke deskundigheid zij beschikken en welke rollen, taken en verantwoordelijkheden collega's vervullen binnen de toetsorganisatie. Een uitgebreid onderzoeksverslag is in voorbereiding (Peters, et al.).

Naar een herontwerp van DTG

Het uitvoeren van zeven pilots en het onderzoek naar de bruikbaarheid, transparantie en uitvoerbaarheid van de DTG-methodiek heeft ons handvatten gegeven de methodiek te herontwerpen. Om de *bruikbaarheid* en *transparantie* van de methodiek te verhogen zijn de volgende

aanpassingen doorgevoerd in het ondersteunende materiaal:

- Er is een handleiding opgesteld voor de begeleiders en de opleidingen, evenals een profiel voor de begeleiders van de methodiek.
- De opdrachten in het werkboek zijn concreetiseerd, zodanig dat de deelnemers niet alleen op een kwalitatief goede en zinvolle wijze de argumenten kunnen vastleggen, maar ook de opbrengsten uit de plenaire discussies weer kunnen geven en beter in staat zijn de argumenten voor de gewenste fase van ontwikkeling te vertalen naar concrete doelen. Ook zijn opdrachten opgenomen om de keuzes bij de prioritering van de doelen beter vast te leggen. Deze aanpassingen dragen uiteraard bij aan het effectief kunnen doorlopen van de PDCA-cyclus.
- Er is meer aandacht voor de terugkoppeling van de opbrengsten van DTG aan het gehele opleidingsteam, zodat de opleidingen beter toegegerust zijn om zelfstandig stap drie en vier uit de kwaliteitscyclus van DTG (het uitwerken, uitvoeren en monitoren van de (verbeter) activiteiten) uit te kunnen voeren.

Uit de pilots bleek dat de uitvoerbaarheid in grote mate beïnvloed wordt door de samenstelling van het team en door de cultuur binnen de opleiding. Om echt zelf de regie te nemen over kwaliteit van toetsing moeten opleidingen zich vragen stellen als: wat is onze visie op toetsing? Hoe moeten wij ons toetsprogramma inrichten zodanig dat dit aansluit bij onze visie op onderwijs en toetsing en alle eindkwalificaties dekt? Wat verstaan wij onder een goede toets en een goed toetsbeleid?

Het stellen én beantwoorden van deze vragen vereist een cultuur waarin het gesprek over kwaliteit, en het al dan niet bereiken hiervan, mogelijk is. Als er hiervan geen sprake is, is het realiseren van zowel duurzaam werken aan kwaliteit als een kwalitatief goede invulling van toetsing niet haalbaar. Het is daarmee dus van essentieel belang dat goed nagedacht

wordt over de samenstelling van het kernteam en dat vooraf gesproken wordt over de wijze waarop binnen de opleiding omgegaan zal worden met de opbrengsten van DTG.

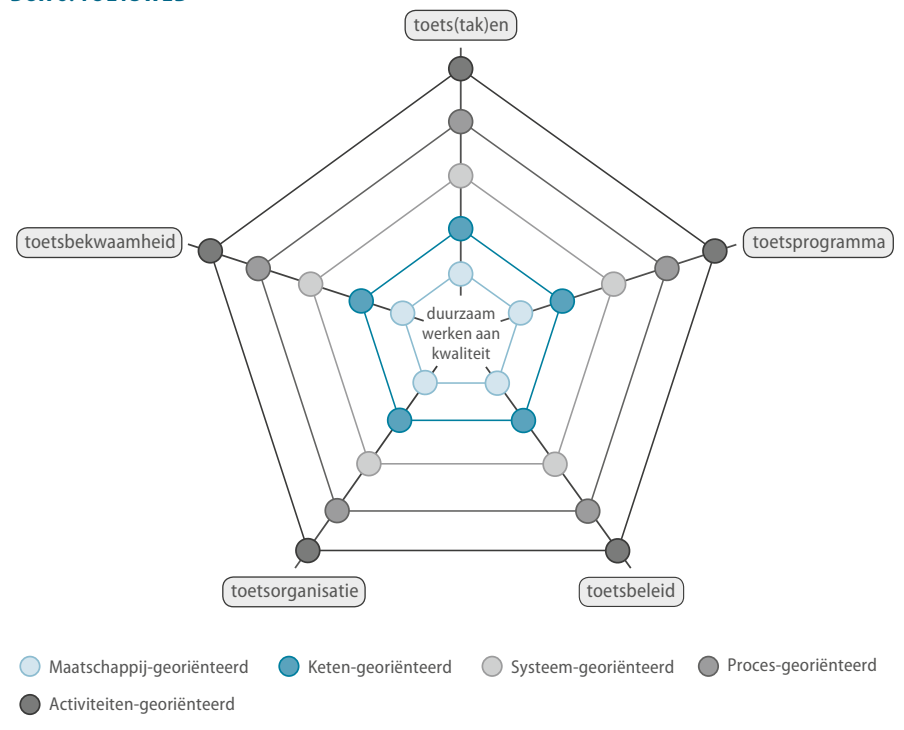
De pilots hebben - naast het doorvoeren van redelijk praktische aanpassingen - ook geleid tot een fundamentele inhoudelijke aanpassing van de methodiek. Bij de operationalisatie van kwaliteit van toetsing is er bij het oorspronkelijke ontwerp voor gekozen gebruik te maken van de kwaliteitspiramide van eigentijds toetsen en beoordelen (box 1): kwaliteit van toetsing wordt hierbij omschreven als het samenspel van de verschillende lagen. Echter, uit de pilots bleek dat de weergave in een piramide ten onrechte leidt tot aannames over voorwaardelijkheid en hiërarchie tussen de verschillende toetsentiteiten. Zo veronderstelt een piramide dat er sprake is van een basis en een top, in de huidige weergave zijn er daarnaast ook twee lagen aan de piramide toegevoegd (organisatie en bekwaamheid) die de andere toetsingsentiteiten lijken te omvatten. Deze aannames staan haaks op het principe dat bij het duurzaam werken aan de kwaliteit van toetsing alle toetsentiteiten van even groot belang zijn en dat juist de combinatie van de afzonderlijke kwaliteit van de lagen én mate van samenspel tussen deze entiteiten bepalend is voor de mate van duurzaam werken aan kwaliteit.

Bovendien is in de pilots geconstateerd dat de piramide nu uitsluitend fungeert als visuele weergave om de verschillende entiteiten waaruit toetsing bestaat te duiden; de onderlinge relatie tussen de lagen, de verschillende fasen van kwaliteit die hierin bereikt kunnen worden en het effect dat kwaliteit van afzonderlijke toetsentiteiten op zowel het duurzaam werken aan kwaliteit als het realiseren van inhoudelijke kwaliteit heeft, wordt hierin niet zichtbaar.

Tot slot hebben we geconstateerd dat het onderscheid tussen de entiteiten toetsen en toetsstaken - nu twee aparte lagen in de piramide - in de dagelijkse onderwijs-

praktijk niet gemaakt wordt. Er is daarom een nieuwe weergave ontwikkeld die de toetsentiteiten, de verschillende ontwikkelingsfasen en de mate waarin dit het duurzaam werken aan kwaliteit beïnvloedt, samenbrengt: het toetsweb (zie box 6). Voor deze visualisatie is inspiratie ontleend aan het curriculaire spinnenweb (Van den Akker, 2003). In het toetsweb zijn de vijf entiteiten van toetskwaliteit - toets(tak)en, toetsprogramma, toetsbeleid, toetsorganisatie en toetsbekwaamheid - nog steeds herkenbaar. Deze vijf zijn verbonden met elkaar én ook met de kern van het web. Vanuit elk van de vijf entiteiten kan worden gewerkt naar duurzame toetskwaliteit via de vijf ontwikkelingsfasen: de buitenste ring is de activiteiten-georiënteerde fase, de binnenste ring de maatschappij-georiënteerde fase. Hierdoor kan een opleiding als het ware omhoog klimmen: hoe verder de opleiding klimt, hoe dichter ze bij de kern - het duurzaam werken aan kwaliteit - komen. De oorspronkelijk ontwikkelde rubric is nu integraal onderdeel van het toetsweb: voor ieder knooppunt is een beknopte omschrijving geschreven waarin geschetst wordt op welke wijze de kwaliteitsprocessen ingericht zijn voor die betreffende toetsentiteit. Deze kern is de verbindende schakel in het web: zowel de toetsentiteiten als de ontwikkelingsfasen beïnvloeden de stabiliteit van deze kern. Het web symboliseert, net als het curriculaire spinnenweb, de kwetsbaarheid van het geheel: te veel of te weinig aandacht voor één entiteit of te weinig aandacht voor de onderlinge verbondenheid van de entiteiten kan de kern van het web verstoren. Maar ook een disbalans in de fasen van ontwikkeling per toetsentiteit kan de kern uit het midden halen. Met deze praktische en fundamentele aanpassingen hebben we een belangrijke stap gezet in het verder verfijnen van de methodiek. Het in dit praktisch artikel beschreven onderzoek heeft echter nog geen antwoord gegeven op de vraag hoe effectief DTG is. Leiden de ambities en doelen daadwerkelijk tot kwalitatief bete-

BOX 6: TOETSWEB



re toetsing en waaruit blijkt dat er sprake is van duurzame toetskwaliteit? Om hier antwoord op te krijgen zal aanvullend praktijkgericht onderzoek worden uitgevoerd. Uiteraard worden opleidingen die deel willen nemen aan dit onderzoek van harte uitgenodigd zich hiervoor aan te melden, want zonder de praktijk is het onmogelijk valide en betrouwbare uitspraken te doen.

Conclusie

In december 2012 presenteerden wij de eerste contouren van de DTG-methodiek. Exact vier jaar later kunnen wij de eerste ervaringen delen dankzij de inzet van vele hbo-opleidingen. Zij zijn van grote waarde geweest in het komen tot een herziene versie die weer voldoende basis geeft voor onderzoek naar de waarde ervan. Graag delen wij wat er nu is. Alle definitieve materialen zijn derhalve opgenomen op de website han.nl/toetsinggetoetst. Hier is ook een digitale versie van het

toetsweb te vinden. We hopen opleidingen met deze materialen te stimuleren zelf de regie te nemen als het gaat om het duurzaam werken aan het en borgen van toetskwaliteit. Kwaliteit van toetsing is van essentieel belang voor de kwaliteit van onderwijs in het hbo, maar het is ook essentieel dat opleidingen zich realiseren dat zij zelf verantwoordelijk zijn voor het formuleren van een onderbouwd verhaal: een verhaal dat recht doet aan de eigen visie op opleiden en toetsen en de specifieke opleidingscontext.

Literatuur

- Akker, J. van den (2003). Curriculum perspectives: an introduction. In J. van den Akker, W. Kuiper & U. Hameyer (eds.), *Curriculum Landscapes and Trends*, (pp. 1-10). Dordrecht: Kluwer Academic Publishers.
- Brouwer, F., & Boer, P. (2013). Auditing Instrument for Sustainability in Higher Education - AISHE 2012 [Beoordelingskader duurzame ontwikkeling]. Den Haag: Hobéon.
- Commissie externe validering examenkwaliteit (2012).

Vreemde ogen dwingen. Den Haag: Vereniging Hogescholen.

- Deming, W. E. (1986). *Out of the Crisis*. Cambridge: MIT-CAES.
- DHO (2010). *AISHE Auditing instrument for Sustainability in Higher Education*. Den Haag: Duurzaam Hoger Onderwijs.
- Dijkstra, A., & Baartman, L. (2011). Zelfevaluatie van de kwaliteit van assessment. *Tijdschrift OnderwijsInnovatie*, 1, 17-25.
- Huisman, P. W. A., & de Vijlder, F. J. (2012). De sleutelrol van professionals governance. *Tijdschrift voor Hoger Onderwijs & Management*, 3, 26-31.
- Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen. Lectorale rede*. Tilburg: Fontys Lerarenopleiding Tilburg.
- Peters, M., De Beer, F., Van Schilt-Mol, T., Sluijsmans, D., Jakobs, L. & Van der Linden, J. (in voorbereiding). *De Toetsing Getoetst onderzocht: herontwerp van de methodiek om de kwaliteit van toetsing te analyseren, verbeteren en borgen*.
- Schilt-Mol, T. van, & Sluijsmans, D. M. A. (2015). Kwaliteit van toetsing geoperationaliseerd. In D. Sluijsmans, D. Joosten-ten Brinke, & T. Van Schilt-Mol (Red.), *Kwaliteit van toetsing onder de loep. Handvatten om de kwaliteit van toetsing in het hoger onderwijs te analyseren, verbeteren en borgen*. Antwerpen-Apeldoorn: Garant, pp. 151-196.
- Schilt-Mol, T. van, Peeters, A., Sluijsmans, D., & Jakobs, L. (2015). Het ontwerp van de methodiek De Toetsing Getoetst. In D. Sluijsmans, D. Joosten-ten Brinke & T. van Schilt-Mol, *Kwaliteit van toetsing onder de loep. Handvatten om de kwaliteit van toetsing in het hoger onderwijs te analyseren, verbeteren en borgen*. Antwerpen-Apeldoorn: Garant, pp. 25-36.
- Sluijsmans, D., Peeters, A., Jakobs, L., & Weijzen, S. (2012). De kwaliteit van toetsing onder de loep. *Tijdschrift OnderwijsInnovatie*, 4, 17-25.
- Sluijsmans, D., Van Schilt-Mol, T., Peeters, A., & Joosten-ten Brinke, D. (2015). Kwaliteit van toetsing in het hoger onderwijs onder de loep. In D. Sluijsmans, D. Joosten-ten Brinke, & T. Van Schilt-Mol (Red.), *Kwaliteit van toetsing onder de loep. Handvatten om de kwaliteit van toetsing in het hoger onderwijs te analyseren, verbeteren en borgen*. Antwerpen-Apeldoorn: Garant, pp. 15-24.