

Tijd voor toetskwaliteit:

het borgen van toetsdeskundigheid van examencommissies

Desirée Joosten-ten Brinke, Lector Eigentijds toetsen en beoordelen, d.tenbrinke@fontys.nl;

Universitair docent Open universiteit, desiree.joosten-tenbrinke@ou.nl

Dominique Sluijsmans, Lector Professioneel beoordelen, dominique.sluijsmans@zuyd.nl

Artikel voor het themanummer in TH&MA

Versie 120903, aantal woorden (excl. Referenties): 2997

“Opleidingen beschikken over een adequaat systeem van toetsing en tonen aan dat de beoogde eindkwalificaties worden gerealiseerd”. Deze standaard, beschreven in het NVAO accreditatiekader (NVAO, 2010), zal menig opleiding inmiddels zeer bekend in de oren klinken. De laatste jaren wordt de kwaliteit van toetsing in het Hoger onderwijs aan een kritisch oog onderworpen en worden opleidingen gestimuleerd om aan deze standaard te voldoen. De druk is hoog: de standaard moet minimaal met een voldoende beoordeeld zijn om de accreditatie goed te doorstaan. Meer dan ooit tevoren wil het Hoger onderwijs investeren in kwaliteit van toetsing. Daarbij komen echter vele vragen op: Wat betekent kwaliteit van toetsing? Wanneer is het systeem van toetsing adequaat? Welke toetsdeskundigheid wordt van examencommissies verwacht? In dit artikel wordt het belang van toetskwaliteit beschreven met bijzondere aandacht voor de rol van examencommissies bij het borgen van toetskwaliteit.

Belang van goed toetsen

In de afgelopen jaren is de aandacht voor kwaliteit van toetsing in het hoger onderwijs toegenomen. Momenteel staat landelijk de kwaliteit van HBO-opleidingen in het centrum van de publieke aandacht. De Inspectie van het Onderwijs (2009) stelde vast dat bij een aantal HBO-opleidingen de wet- en regelgeving rond het afstuderen onvoldoende wordt nageleefd. In de nasleep daarvan komt een ontwikkeling op gang waarin publieke verantwoording van de HBO-instellingen gevraagd wordt. Evaluaties van kwaliteit van de opleidingen gebeuren veelal vanuit reflecties op instroom- doorstroom- en uitstroomrendementen. Doelstelling van onderwijsinstellingen is een zo hoog mogelijk rendement. Kwaliteit van beoordelen is hierbij essentieel. Opleidingen met een goede combinatie van toetsing en feedback in de opleiding halen relatief hoge rendementen (Gibbs, 2010). Verkeerde beoordelingen kunnen leiden tot een lager rendement. Het Interstedelijk Studenten Overleg (ISO, 2009) merkte op dat er nog lang niet bij alle instellingen en opleidingen voldoende aandacht voor toetsing is. Studenten constateren dat toetsing en examinering een onderschoven kindje zijn bij gesprekken en debatten over onderwijskwaliteit en dat de kwaliteit van onderwijs nog zo hoog kan zijn, als het niveau van de

toetsing en examinering dat niet is, een groot deel van de investeringen in het onderwijs teniet wordt gedaan. Simpelweg omdat studenten afgerekend worden op de (studie-)punten die zij voor de toetsen behalen en niet op wat zij daadwerkelijk geleerd hebben. Studenten merken direct het effect van voldoende of onvoldoende kwaliteit van de toets (Mager & Nowak, 2012). Hierdoor wordt hun tevredenheid beïnvloed en deze tevredenheid wordt door onderwijsinstellingen gebruikt als indicator voor de kwaliteit van hun toetsing (Gerritsen – Van Leeuwenkamp, 2012). In september 2010 is de Wet Versterking Besturing van kracht geworden die geleid heeft tot diverse wijzigingen in de Wet op het Hoger Onderwijs en Wetenschappelijk Onderzoek (WHW). De meeste wijzigingen zijn per september 2010 in werking getreden. Conform deze wet is de examencommissie het orgaan dat op objectieve en deskundige wijze vaststelt of een student voldoet aan de voorwaarden die de onderwijs- en examenregeling (OER) stelt ten aanzien van kennis, inzicht en vaardigheden die nodig zijn voor het verkrijgen van een graad. Daarmee verschuift de rol van examencommissies van het uitvoeren van procedures rondom toetsing naar het borgen van de kwaliteit van de toetsing. Dat vraagt een onafhankelijke en deskundige examencommissie. In januari 2011 is het nieuwe accreditatiestelsel in werking getreden waarbij 'Toetsing en gerealiseerde eindkwalificaties' als aparte standaard is opgenomen. Deze standaard geldt als knock-out criterium voor de accreditaties. Dat wil zeggen dat opleidingen die een onvoldoende halen op dit criterium de accreditatie verliezen. Examencommissies moeten zich bij de accreditatie kunnen verantwoorden over het aspect 'toetsen en beoordelen' (Smits, 2011).

Ook wetenschappelijk gezien is kwaliteit van toetsing een thema dat toenemend object van onderzoek is. Hoewel de nadruk vooral heeft gelegen op kwaliteit van toetsen en gebruikte toetsinstrumenten, is er recentelijk meer aandacht voor de kwaliteit van programma's van toetsing, kwaliteit van de betrokkenen van toetsing of de organisatie en procedures rondom toetsing.

De maatschappij en de politiek hebben echter in de loop van de jaren steeds meer vraagtekens gezet bij het niveau van de afgestudeerden en de kwaliteit van toetsing in het algemeen (Ministerie van OC&W, 2012; Onderwijsraad, 2006). Het louter verzamelen van een reeks 'stempels' per docent en per vak ter afronding blijkt niet meer toereikend om het HBO-niveau aan te tonen. Vele voorbeelden in de media leiden ertoe dat aan de kwaliteit van de bewijsvoering wordt getwijfeld.

Omdat toetsen informatie opleveren op basis waarvan kleine en grote beslissingen over de voortgang van de student worden genomen, is het van belang in goede toetsing te investeren. Daarbij omvat toetsen in zowel praktische als theoretische zin meer dan het construeren van een toets alleen: het vormt de ruggengraat in het curriculum van de opleiding en is daarmee een sterk sturende factor in het bepalen van kwaliteit van het onderwijs.

Implicaties voor het systeem van toetsing

Het Hoger onderwijs, en met name de Hogescholen, heeft onder invloed van alle ontwikkelingen niet stilgestaan op toetsgebied. Het besef dat de kwaliteit van toetsing voor een groot deel de kwaliteit van het onderwijs representeert en dus leren en toetsen goed op elkaar aan moet sluiten, heeft drie ingrijpende implicaties gehad voor de inrichting van het toetsbouwwerk van opleidingen en de actoren hierin.

Een eerste implicatie is een sterke uitbreiding aan toetsvormen – en instrumenten. De opkomst van bijvoorbeeld het competentiegericht leren, heeft geresulteerd in een mix van toetsvormen die doorgaans worden bestempeld als 'innovatief'. Voorbeelden zijn de generieke toepassing van (digitale) portfolio's, proeven van bekwaamheid en simulaties.

Een tweede implicatie is dat toetsing anno 2012 een herijking van de kwaliteitscriteria vraagt. Omdat toetsing integraal onderdeel wordt van een leertraject zal naast de gangbare psychometrische criteria als betrouwbaarheid en validiteit een uitbreiding van criteria nodig zijn. De laatste jaren is dan ook (inter)nationaal geïnvesteerd in deze uitbreiding waarbij aansluiting met het leerproces een prominentere plek krijgt. Drie voorbeelden hiervan zijn de zeven principes van Assessment 2020 (Boud and associates, 2010; zie bijlage 1) en de tien kenmerken van duurzaam toetsen (Sluijsmans, 2010; zie bijlage 2) en het kwaliteitskader van Baartman en anderen (2006; zie bijlage 3).

Een derde implicatie is dat de inzichten over de kwaliteit van toetsing een uitbreiding van kennis en kunde over toetsing vraagt die opleiders vanuit hun eigen opleiding niet (voldoende) hebben vergaard. Omdat toetsing niet meer los staat van het onderwijs en de kwaliteit van de toetsing steeds meer moet worden verantwoord, zal de opleider zich moeten professionaliseren in de nieuwe benaderingen van toetsing. Ook zal de opleider moeten worden voorzien van handvatten om de toetsing beheersbaar, transparant en eerlijk te houden.

Borgen van toetskwaliteit: de specifieke taken en verantwoordelijkheden van examencommissies

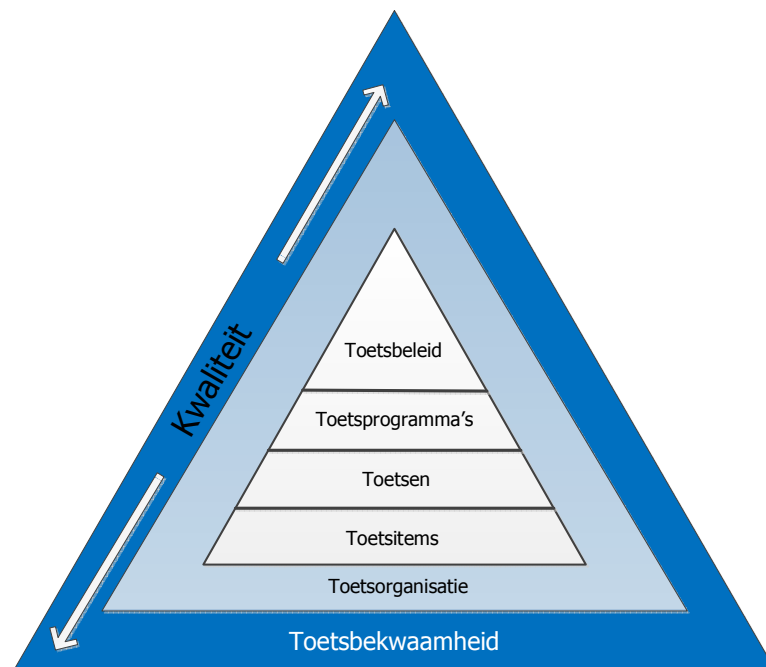
Om kwaliteit van toetsen op een hoger plan te trekken hebben examencommissies door de wijzigingen in de Wet op het Hoger onderwijs een nieuwe rol gekregen: naast hun aanvankelijke procedurele verantwoordelijkheid krijgen zij nu ook een inhoudelijke verantwoordelijkheid.

In het nieuwe takenpakket is veel nadrukkelijker beschreven waarvoor zij verantwoordelijk zijn. Aan de ene kant zijn dat beleidsmatige taken, zoals het vaststellen van regels voor borgen van de kwaliteit van tentamens en examens en het vaststellen van richtlijnen en aanwijzingen binnen kader van de Onderwijs- en Examen Regeling (OER) om uitslag van tentamens te beoordelen en vast te stellen. Anderzijds zijn dat uitvoerende taken zoals het uitreiken van getuigschriften, het verlenen van vrijstellingen, het opleggen van sancties bij fraude, aanwijzen van examinatoren. Bij deze taken staat

het borgen van de kwaliteit voorop. Dit vereist verschillende deskundigheden op het gebied van wet- en regelgeving, opleiding, kwaliteitszorg, organisatie en toetsing. Examencommissies moeten op objectieve en deskundige wijze vaststellen en waarborgen of een student voldoet aan de voorwaarden die de OER stelt ten aanzien van kennis, inzicht en vaardigheden die nodig zijn voor het verkrijgen van de graad. Om die kwaliteit te kunnen waarborgen moet het voor examencommissies duidelijk zijn hoe de gewenste kwaliteit kan worden bereikt.

De kwaliteitspiramide voor eigentijds toetsen en beoordelen

Voordat we overgaan tot een beschrijving van de deskundigheden van examencommissies die nodig zijn om bovenstaande taken te kunnen uitvoeren, is het zinvol om eerst een invulling te geven aan het begrip toetskwaliteit. Hiervoor hanteren we de kwaliteitspiramide van eigentijds toetsen en beoordelen van Joosten-ten Brinke, Sluijsmans & Peeters (2012), zie Figuur 1.



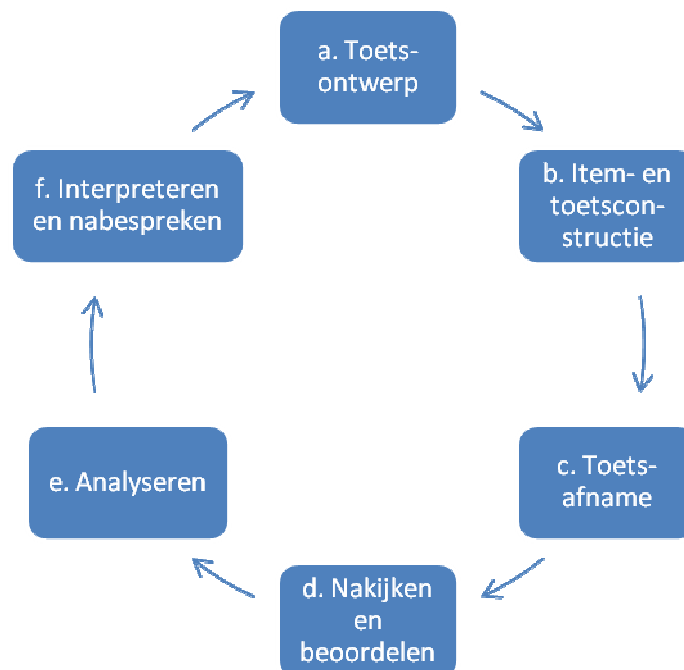
Figuur 1. *De kwaliteitspiramide voor eigentijds toetsen en beoordelen (Joosten-ten Brinke, 2011; Joosten-ten Brinke, Sluijsmans, & Peeters, 2012)*

In de kwaliteitspiramide in het witte gedeelte worden vier lagen onderscheiden:

- *Kwaliteit van toetsen en toetsvragen.* Elke taak, opdracht of vraag in een toets moet voldoen aan de criteria relevantie, objectiviteit, efficiëntie en moeilijkheid;
- *Kwaliteit van toetsen.* Hiermee wordt de kwaliteit van de hele toets bedoeld, inclusief het ontwerp van de beoordelingsmodellen en student- en beoordelaarsinstructies. Voorbeelden van kwaliteitscriteria zijn validiteit, betrouwbaarheid en bruikbaarheid;

- *Kwaliteit van toetsprogramma's*. Hiermee wordt de kwaliteit van het toetsprogramma als geheel bedoeld. Het gaat om de heldere samenhang tussen toetsen en de borging van de verschillende functies van toetsing. Kwaliteitscriteria zijn gericht op toetsprogramma's als geheel;
- *Kwaliteit van het toetsbeleid*. Hiermee wordt de kwaliteit van de toetsing vanuit de visie van de opleiding. Kwaliteit van toetsing begint bij de visie die opleidingen hebben ten aanzien van toetsen en beoordelen. Het inrichten van de toetsing is een ontwerpvoorbeeld, waarbij de aansluiting met het leren essentieel is (zie ook Schuwirth & Van der Vleuten, 2006). In hoeverre draagt het toetsprogramma bij aan een duurzaam leerresultaat? Hoe is de toetsing verbonden aan kennis en vaardigheden voor de lange termijn?

De kwaliteitsborging per laag is afhankelijk van de fase van het toetsproces (zie Figuur 2). In de ontwerpfase zullen vragen over de functie van toetsen, de keuze voor toetsvormen, de beoordelaars, de uitlevervorm belangrijk zijn, terwijl bij de fase van toetsvraag- en toetsconstructie deskundigheid nodig is in het beoordelen van de kwaliteit van de geconstrueerde vragen en samengestelde toetsen. Voor het borgen van de kwaliteit is het belangrijk dat het hele proces doorlopen wordt. Na een toetsafname moeten toetsen en de toetsvragen geanalyseerd worden en moeten de resultaten geïnterpreteerd en nabesproken worden. Studenten moeten zodanig feedback op hun toetsresultaten krijgen dat zij daarvan kunnen leren voor een toekomstige situatie. Maar ook docenten moeten weten wat de kwaliteit is van de toets en de afzonderlijke toetsvragen om hiermee conclusies te trekken voor een volgend toetsontwerp.



Figuur 2. *Het toetsproces*

De lagen *toetsorganisatie en toetsbekwaamheid* (zie Figuur 1) worden beschouwd als twee cruciale condities voor het realiseren van een goede kwaliteit van toetsing. Opleiders – en dus ook de leden van de examencommissies - zijn echter niet of nauwelijks opgeleid in goede toetsing en hebben onvoldoende kennis van de basisprincipes van toetsen en beoordelen (ISO, 2009; Straetmans, 2006). Dat is zorgelijk immers het vaststellen van een betrouwbaar beeld van de kennis en kunde van de lerende en het nemen van vaak ingrijpende beslissingen over de studieloopbaan van studenten vraagt om kwalitatief goede beoordelaars. Naast de toetsbekwaamheid is de hele toetsorganisatie belangrijk om toetskwaliteit te kunnen garanderen. Het gaat hier om criteria als kosten en haalbaarheid van de wijze van toetsing (Baartman, 2008), maar ook over zaken als het vaststellen van taken en verantwoordelijkheden van toets- en examencommissies, het komen tot vormafspraken en het vaststellen van toetsroosters.

Toetsdeskundigheid in examencommissies

De samenstelling van de examencommissie moet zodanig zijn dat de toetsdeskundigheid op alle niveaus van de kwaliteitspiramide (zie Figuur 1) en in alle fasen van het toetsproces kan worden gegarandeerd. Volgens de wet moeten leden van de Examencommissie benoemd worden op basis van hun deskundigheid op het terrein van de desbetreffende opleiding en ten minste één lid moet als docent verbonden zijn aan de opleiding. Hierbij wordt er dan vanuit gegaan dat docenten toetsdeskundig zijn. Een aanname die vanuit de functieomschrijving van docenten, waarin opgenomen is dat zij studenten moeten kunnen beoordelen, ook niet vreemd is. Docenten zijn vaak wel vakinhoudelijk bekwaam, maar niet altijd toetsbekwaam. Om als docent, en daarmee als lid van de examencommissie, voldoende toetsbekwaam te zijn, moeten zij de kennis, vaardigheid en motivatie hebben om verantwoord de ontwikkeling van studenten te beoordelen. Elke docent moet in staat zijn om toetsinformatie te interpreteren en te gebruiken om vast te stellen waar onze studenten staan en hoe zij verder kunnen bijdragen aan het leren van de student (Absolum, Flockton, Hattie, Hipkins, & Reid, 2009).

Elf toetscompetenties voor leden van examencommissies

Brookhart (2009) heeft de volgende 11 standaarden voor toetscompetenties van docenten beschreven die ook voor leden van de examencommissies gelden. Deze zijn:

1.	Docenten moeten begrijpen wat 'leren' is in de context waarin ze les geven.
2.	Docenten moeten in staat zijn om toegankelijke en meetbare leerdoelen te formuleren die aansluiten bij de gewenste inhoud en beheersingsniveau volgens de curriculum doelen en visies.

3.	Docenten moeten beschikken over verschillende strategieën om met studenten te communiceren over de leerdoelen en hoe deze bereikt kunnen worden.
4.	Docenten moeten de bedoelingen en het gebruik van verschillende beoordelingsmethodes kennen en vaardig zijn om deze te gebruiken.
5.	Docenten moeten vragen in klassensituaties, toetsvragen en performance assessmenttaken kunnen analyseren om er zeker van te zijn dat ze passen bij het meten van de kennis en vaardigheden van hun studenten.
6.	Docenten moeten effectieve en bruikbare feedback op studentenwerk kunnen geven.
7.	Docenten moeten in staat zijn om beoordelingsmodellen te construeren die leiden tot bruikbare informatie om beslissingen te kunnen nemen over studenten en/of klassen. Deze beslissingen moeten leiden tot verbeteringen van het leren, groeien of ontwikkelen van de student.
8.	Docenten moeten externe toetsresultaten kunnen gebruiken en op de juiste manier interpreteren ten behoeve van de juiste beslissingen.
9.	Docenten moeten hun interpretaties van resultaten en beslissingen kunnen communiceren naar de studenten, schoolleiding en ouders.
10.	Docenten moeten hun studenten kunnen helpen om op basis van toetsinformatie beslissingen te nemen.
11.	Docenten zijn zich bewust van hun wettelijke en ethische verantwoordelijkheden bij het beoordelen van studenten.

De standaarden passen bij de vernieuwde visie op toetsen en beoordelen waarbij de rol van de student meer centraal staat. Een student heeft op een bepaald moment, bepaalde informatie nodig om beslissingen te kunnen nemen over zijn eigen leren (leerproces). Deze beschrijving sluit aan bij het volledige toetsproces en stelt communicatie over toetsing vanuit een visie op toetsen naar verschillende betrokkenen centraal. Docenten die voldoen aan deze competenties kunnen in de examencommissie plaatsnemen. Daarnaast is het examencommissies toegestaan om deskundigen van buiten de opleiding of zelfs de instelling lid te maken van de Examencommissie om op die manier in de toetsdeskundigheid van de commissie te voorzien. Om echter de kwaliteit van toetsing te borgen en daarover in gesprek te gaan met bijvoorbeeld examinatoren is het wenselijk dat alle leden in ieder geval beschikken over een basis toetsdeskundigheid. Dat betekent dat voor specifieke zaken een beroep gedaan kan worden op een specialist, maar dat de meest voorkomen kwaliteitsaspecten door alle leden overzien moeten kunnen worden.

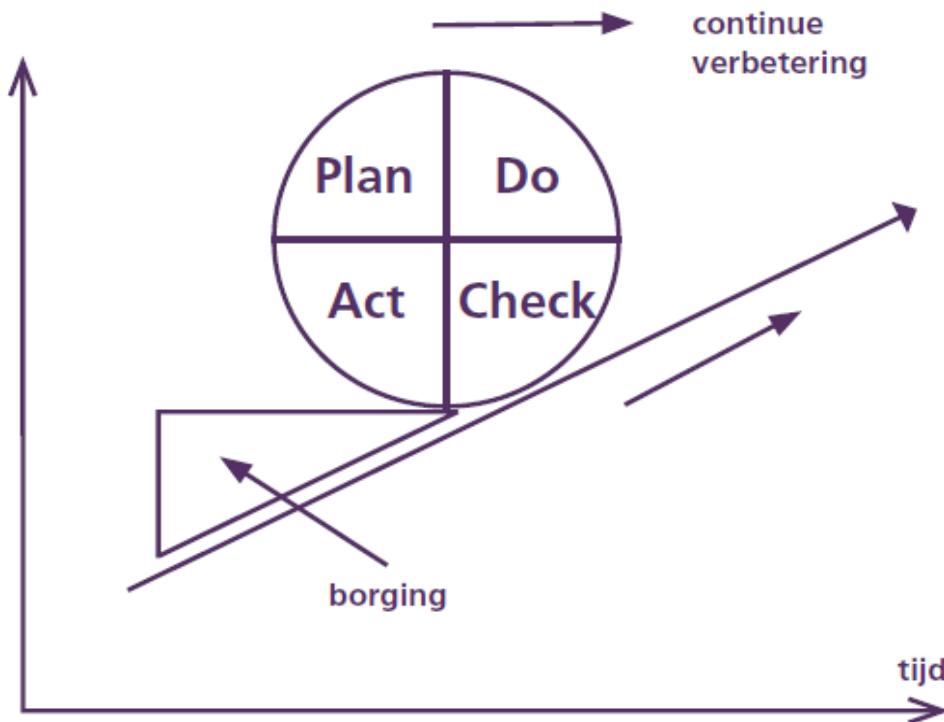
Borging van het eindniveau: PLAN-DO-ACT-CHECK

De examencommissie heeft volgens de WHW tot taak *'op objectieve en deskundige wijze'* (artikel 7.12 lid 2) te beoordelen of het gerealiseerde eindniveau van een afgestudeerde voldoet aan de eisen die

daaraan gesteld zijn. Het beoordelen van dit gerealiseerde eindniveau zoals beschreven door de NVAO (2010) sluit aan bij de verschillende niveaus van de kwaliteitspiramide (Figuur 1). Op het onderste twee niveaus wordt gekeken naar de kwaliteit van de zogenoemde *tussentijdse en afsluitende toetsen* en de *afstudeerwerken*. De toetsing en de beoordeling moeten valide en betrouwbaar zijn en voor de studenten transparant. Op het derde niveau is het van belang dat er sprake is van samenhang in het programma en tot slot wordt vanuit het duurzaamheidsprincipe gekeken naar de wijze waarop afgestudeerden in de praktijk of in een vervolgopleiding functioneren.

Het te behalen eindniveau is afhankelijk van het type opleiding. Op HBO-niveau onderscheiden we het associate degree-niveau, het bachelor-niveau en het master-niveau; Op WO-niveau wordt het bachelor- en het master-niveau onderscheiden. De Dublin descriptoren bieden de mogelijkheid om deze niveaus verder te beschrijven. De laatste jaren is echter in het HBO ook veel nadrukkelijk het uitvoeren van onderzoek als vereiste opgevoerd.

Voor de examencommissie is het van belang om dit eindniveau te garanderen en te borgen. Een hulpmiddel hierbij is de kwaliteitscirkel van Deming (1986), In deze kwaliteitscirkel wordt door middel van vier stappen, te weten Plan, Do, Check, Act, de kwaliteit van processen gestimuleerd. Op basis van hun expertise bepalen examencommissies welke processen of activiteiten aangepakt moeten worden om het eindniveau van de afgestudeerden te kunnen garanderen. In het geval dat een examencommissie geconstateerd heeft dat een deel van de leerdoelen niet bereikt wordt door de studenten maakt de commissie een Plan. Ze kijken naar de aansluiting van de leerdoelen op de instructies en de gebruikte beoordeling (= '*constructive alignment*', Biggs, 1996) en bedenken hoe deze aansluiting kan worden verbeterd. Dit plan wordt uitgewerkt (DO) en daarna controleert de examencommissie of de uitwerking van het plan het oorspronkelijke probleem oplost (CHECK). Als dat het geval is wordt de uitwerking geïmplementeerd (Act) en worden nieuwe plannen gemaakt. De borging door de examencommissie zit hem er dan in dat zij voorkomen dat de kwaliteitsverbetering die gemaakt is behouden blijft en niet in een volgende aanpassing weer teruggedraaid wordt. Dit proces van kwaliteitsborging wordt weergegeven in de Figuur 3. Waarbij de helling staat voor de toename van kwaliteit. Op het moment dat de borgpen (lees: het werk van de examencommissie) er onderuit wordt gehaald, rolt de kwaliteit achteruit.



Figuur 3. De kwaliteitscyclus van Deming (uit: Jaspers & Van Zijl, 2011, pagina 31).

Tot slot

Met deze bijdrage hebben we getracht de complexe taak van examencommissies in kaart te brengen. Naast een overzicht van taken en verantwoordelijkheden is een nadere invulling van het begrip toetskwaliteit gepresenteerd. Het borgen van toetskwaliteit is een enorm lastige taak. De kwaliteit wordt door veel aspecten beïnvloed. Toetsdeskundigheid is vanuit dat perspectief dan ook een vereiste voor elke examencommissie. De vraag is echter of elk lid van de examencommissie deze expertise moet hebben. Zoals in het begin van dit artikel is aangegeven, zijn er in de examencommissie verschillende deskundigheden nodig, te weten deskundigheden op het gebied van wet- en regelgeving, opleiding, kwaliteitszorg, organisatie en toetsing. Belangrijk is dat al deze deskundigheden aanwezig zijn en dat de examencommissie vertrouwen heeft in hun belangrijke taak. Het borgen van de toetskwaliteit zodanig dat er continu op toegezien wordt dat opleidingen beschikken over een adequaat systeem van toetsing waarbij studenten de opleiding verlaten aantoonbaar voldoen aan de gestelde eindkwalificaties. Een zorgvuldig kwaliteitsborgingsysteem is noodzakelijk om de kwaliteit van toetsing op elke moment, zowel intern als extern te kunnen verantwoorden.

Referenties

Absolum, M., Flockton, L., Hattie, J., Hipkins R., Reid, I., (2009) *Directions for Assessment in New Zealand (DANZ)*. New Zealand: Ministry of Education.

Baartman, L. K. J. (2008). *Assessing the assessment. Development and use of quality criteria for Competence Assessment Programmes*. Proefschrift.

Baartman, L. K. J., Prins, F. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2007). Determining the quality of competence assessment programs: a self-evaluation procedure. *Studies in Educational Evaluation*, 33, 258-281.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.

Boud, D. and associates (2010). *Assessment 2010: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council.

Gerritsen – Van Leeuwenkamp, K. (2012). *Het Relatieve Belang van Vijftig Kwaliteitskenmerken van Toetsing voor Studententevredenheid in het Hoger Beroepsonderwijs*. Masterscriptie. Heerlen: Open universiteit.

Gibbs, G. (2010). Does assessment in open learning support students? *Open Learning: The Journal of Open and Distance Learning*, 25(2), 163-166.

ISO (2009). *Toetsing een vak apart*. Utrecht: Interstedelijk studenten overleg.

Jaspers, M. & Van Zijl, E. (2011). *Kwaliteit van toetsing in het Hoger Onderwijs, naar een versterkte rol van examencommissies*. Eindhoven: Fontys Hogescholen.

Inspectie van het Onderwijs (2009). *Boekhouder of wakend oog. Verslag van een onderzoek bij examencommissies in het hoger onderwijs over de garantie van het niveau*. Inspectierapport 2009-16 (april). <http://www.onderwijsinspectie.nl/actueel/publicaties/Boekhouder+of+wakend+oog.html>.

Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen* (lectorale rede). Tilburg: Fontys Hogescholen.

Joosten-ten Brinke, D., Sluijsmans, D. M. A., & Peeters, A. (2012). *De toetsing getoetst. Rapportage van de eerste projectfase gericht op het ontwerp van een methodiek om toetskwaliteit te analyseren, verbeteren en borgen*. Nijmegen: Hogeschool van Arnhem en Nijmegen.

Ministerie van OC&W. (2012). Kamerbrief hoger beroepsonderwijs, 17 jan 2012. Den Haag: OCW.

Nederlands-Vlaamse Accreditatieorganisatie (2010). *Beoordelingskaders ten behoeve van het accreditatiestelsel*. NVAO: Den Haag.

Mager, U., & Nowak, P. (2012). Effects of student participation in decision making at school. A systematic review and synthesis of empirical research. *Educational Research Review*, 7, 38-61. doi: 10.1016/j.edurev.2011.11.001.

Onderwijsraad (2006). *Advies Examinering: draagvlakken toegankelijkheid, uitgebracht aan de staatssecretaris van Onderwijs, Cultuur en Wetenschap. Nr. 20060320/865*. Den Haag (november). http://www.onderwijsraad.nl/upload/publicaties/316/documenten/examinering__draagvlak_en_toegankelijkheid.pdf.

Schuwirth, L. W. T., & Van der Vleuten, C. P. M. (2005). A plea for new psychometric models in educational assessment. *Medical Education*, 40, 296-300.

Sluijsmans, D. M. A. (2008). *Duurzaam beoordelen in vraaggestuurd leren* (lectorale rede). Nijmegen: Hogeschool van Arnhem en Nijmegen.

Sluijsmans, D. M. A. (2010). Beoordelen met het oog op de toekomst: Tien kenmerken van duurzaam beoordelen. *Examens, tijdschrift voor de toetspraktijk*, 7, 5-9.

Smits, R. (2011). Examencommissies HBO worden belangrijker! *Examens, tijdschrift voor de toetspraktijk*, 2, 16-19.

Straetmans, G. J. J. M. (2006). *Bekwaam beoordelen en beslissen* (Lectorale rede). Deventer: Saxion Hogescholen.

Bijlage 1.

Zeven principes ten aanzien van effectiviteit van toetsing (Boud and associates, 2010):

1. Toetsing heft tot doel studenten te betrekken en productief leren te bevorderen;
2. Feedback wordt gebruikt om het actief leren van studenten te bevorderen;
3. Studenten en opleiders zijn gezamenlijk verantwoordelijk voor het leerproces;
4. Studenten worden betrokken bij de toetspraktijk en cultuur van het hoger onderwijs;
5. Toetsing om te leren staat centraal in het ontwerp van het curriculum;
6. Toetsing is focus van het hele team als lerende organisatie;
7. Toetsing representeert een waarheidsgetrouw beeld van de prestaties van de student.

Bijlage 2.

Tien kenmerken voor kwaliteit van duurzaam toetsen (Sluismans, 2010):

1. Verbetering van de student staat centraal in de toetsing.
2. Het accent ligt op leren voor de toekomst.
3. Er is sprake van betekenisvolle toetstaken.
4. Zelfsturing wordt gestimuleerd.
5. De student voelt zich eigenaar van de toetsing.
6. De mate van duurzaamheid van kennis wordt erkend.
7. Toetsen is de start van nieuwe leercyclus.
8. Er zijn heldere criteria en standaarden.
9. De resultaten van toetsen zijn informatierijk.
10. In de toetsing wordt informeel en non-formeel leren erkend.

Bijlage 3. Kader van Baartman (2008)

Baartman (2008) benoemt de volgende 12 criteria voor kwalitatief goede toetsing:

1. Geschiktheid voor onderwijsdoelen: heeft te maken met de afstemming tussen doelen van de toetsing en de doelen van het onderwijs.
2. Vergelijkbaarheid: houdt in dat de toetsing op consistente wijze moet worden uitgevoerd.
3. Herhaalbaarheid van beslissingen: houdt in dat voor de eindbeoordeling verschillende bronnen worden gecombineerd om zo tot een betrouwbare beslissing te komen.
4. Acceptatie: betekent dat alle betrokkenen bij een toetsprogramma zich kunnen vinden in de taken of opdrachten, de criteria en de uitvoering.
5. Transparantie: houdt in dat een toetsprogramma voor alle betrokkenen doorzichtig en begrijpelijk moet zijn.
6. Eerlijkheid: betekent dat alle studenten een eerlijke kans krijgen hun competenties te demonstreren.
7. Ontwikkeling van zelfsturend leren: heeft te maken met één van de doelen van competentiegericht onderwijs, namelijk dat studenten zelf hun leren vormgeven en sturen. Toetsing kan hieraan bijdragen doordat er aandacht wordt besteed aan zelf- en peerbeoordelingen;
8. Betekenisvolheid: heeft betrekking op hoe zinvol en nuttig de betrokkenen de toetsing vinden. Goede toetsing is zinvol voor studenten omdat ze er wat van leren, of omdat ze inzicht krijgen in hun sterke en zwakke kanten.
9. Authenticiteit: dit is de mate van overeenkomst van een toetsprogramma met het toekomstige beroep. Een toetsprogramma moet overeenkomen met de beroepspraktijk.
10. Cognitieve complexiteit: heeft ook te maken met de overeenkomst van een toetsprogramma met de toekomstige beroepssituatie, maar is meer gericht op de denkprocessen. Een beoordelaar moet kunnen vaststellen of iemand de cognitieve vaardigheden bezit om in een beroep te kunnen functioneren.

11. Onderwijsgevolgen: heeft te maken met de effecten van een toetsprogramma op het leren van studenten en de instructie door docenten.
12. Tijd en kosten: bepalen mede de haalbaarheid van de uitvoering van een toetsprogramma. Is er genoeg tijd en geld beschikbaar om het toetsprogramma uit te voeren?