

VERANTWOORD TOETSEN EN BESLISSEN IN HET HOGER BEROEPSONDERWIJS

een voorstel voor een programma van eisen voor
een basis- en seniorkwalificatie examinering
(BKE/SKE)



van de expertgroep BKE/SKE
in opdracht van de Vereniging Hogescholen



Oktober 2013

Inhoudsopgave

Voorwoord	3
1. Aanleiding, context en opdracht.....	4
2. Samenstelling en werkwijze expertgroep.....	6
3. Operationalisering van de opdracht	8
3.1 Uitwerking van de opdracht en begripsbepaling.....	8
3.2 De rationale voor het vormgeven van een SKE	8
4. Conceptueel kader voor de BKE en SKE	9
4.1 De kwaliteitspiramide voor eigentijds toetsen en beoordelen	9
4.2 De toetscyclus.....	10
4.3 De kwaliteitspiramide en de toetscyclus in relatie tot de BKE en de SKE	12
5. Het onderscheid tussen de BKE en de SKE	13
6. Leeruitkomsten van de BKE en de SKE	14
7. Criteria voor de bewijslast van de BKE en de SKE	16
Bijlage 1. Toelichting op de toetscyclus.....	17
Bijlage 2. Aanbevolen toetsbronnen.....	19

Voorwoord

Voor u ligt het voorstel van de expertgroep basis- en seniorkwalificering examinatoren (BKE/SKE). Deze groep heeft zich in opdracht van de Vereniging Hogescholen gebogen over de formulering van een programma van eisen voor een module basis- en seniorkwalificatie voor examinatoren in het hoger beroepsonderwijs, naar aanleiding van het rapport *Vreemde ogen dwingen*. In dat rapport adviseert de Commissie Bruijn om te komen tot verdere professionalisering van examinatoren.

De expertgroep heeft het begrip 'module' geïnterpreteerd als de *context* waarbinnen hogescholen gaan zorgen dat aan de leeruitkomsten en bewijslast voor het behalen van een BKE of SKE wordt voldaan. Uitgangspunt van dit voorstel is de *functie* van een BKE en SKE: wat beoogt een BKE/SKE en welke bewijslast is nodig om te bepalen of deze functie wordt geborgd? De expertgroep doet dus geen uitspraken over de *vorm* waarin dat moet gaan gebeuren. Hogescholen bepalen zelf of dit aanbodgestuurd, zelfsturend, informeel of via afstandsleren gebeurt. De expertgroep sluit hiermee aan op de strekking van het rapport *Vreemde ogen dwingen*, waar steeds de voordelen van bottom-up werken benadrukt worden en op de opdracht vanuit de Vereniging Hogescholen om 'quick wins' te realiseren, door gebruik te maken van ontwikkelde expertise. De brede opvatting van het begrip 'module' laat hogescholen de ruimte om in de BKE/SKE aan te sluiten op de eigen toetspraktijk en gebruik te maken van ontwikkelde cursussen. Gezien de expertise die op hogescholen aanwezig is, zoals blijkt uit al ontwikkelde cursussen op gebied van toetsing, zullen de hogescholen de vorm waarin de leeruitkomsten verwezenlijkt worden zeker kunnen borgen. Samenvattend heeft de expertgroep zich gericht op de *wat*-vraag en niet op de *hoe*-vraag.

In deze rapportage wordt in hoofdstuk 1 de aanleiding, context en opdracht toegelicht. In hoofdstuk 2 en 3 worden achtereenvolgens de samenstelling en werkwijze van de expertgroep gepresenteerd en de operationalisering van de opdracht. Vervolgens wordt in hoofdstuk 4 verantwoord welke conceptuele kaders zijn gebruikt als basis voor het programma van eisen. Het onderscheid tussen de BKE en SKE wordt geduid in hoofdstuk 5. In hoofdstuk 6 worden de leeruitkomsten voor de BKE en SKE gepresenteerd. De rapportage besluit hoofdstuk 7 waarin de criteria voor de bewijslast voor het behalen van een BKE of SKE worden geformuleerd.

De expertgroep is van mening dat examinering in het hoger onderwijs door de invoering van BKE/SKE een stevige kwaliteitsimpuls zal krijgen. Docenten spelen immers een cruciale rol in het examen- en toetsproces. De kwaliteit van examinering zal op alle fronten toenemen, wanneer docenten zich kwalificeren als BKE of SKE en wanneer hogescholen toezien op het deskundigheidsniveau van hun examinatoren.

Wij wensen u plezier bij het lezen van dit voorstel en stellen kritische feedback op prijs.

Namens de expertgroep,

Dominique Sluijsmans
Voorzitter expertgroep

Dineke Behrend
Secretaris expertgroep

Oktober 2013

1. Aanleiding, context en opdracht

In mei 2012 is naar aanleiding van ongerustheid over en kritiek op de kwaliteit van hbo-diploma's het rapport van de Commissie Bruijn verschenen: *Vreemde ogen dwingen*. De commissie Bruijn heeft verschillende maatregelen voorgesteld die ervoor moeten zorgen dat toets(kwaliteits)beleid steviger wordt verankerd in het hart van de opleiding. Eén van de maatregelen die in het rapport wordt genoemd betreft de toetsdeskundigheid van docenten:

"Een docent dient aantoonbaar over voldoende onderwijskundige/didactische deskundigheid te beschikken, en daarbinnen over toetsdeskundigheid om zich te kwalificeren als (zelfstandig) examiner, (extern) lid van een examencommissie bij de eigen of een andere hogeschool, assessor of gecommiteerde" (p. 41).

Op basis van deze maatregel beveelt de commissie (verdere) professionalisering van examinatoren aan en een externe validering via certificering van examinatoren en opleiding van docenten. De volgende mogelijkheden hiertoe worden in het rapport benoemd volgens het principe 'pas toe of leg uit' (p. 61-62):

- De opzet van een systeem van certificering van examinatoren, via de invoering van een basis- en senior kwalificatie examinering (BKE/SKE) systematiek,
- Scholing van docenten die bij toetsing betrokken zijn tenminste tot op het niveau van hun betrokkenheid,
- De instelling van een wettelijk register voor bij- en nascholing van hbo-docenten, waarin ook de toetsexpertise (BKE of SKE) van de docent wordt opgenomen.

Naar aanleiding van het rapport heeft de Vereniging Hogescholen een voorstel tot implementatie naar de toenmalige staatssecretaris van Onderwijs, Cultuur en Wetenschappen (OCW) Halbe Zijlstra gestuurd¹. De minister van OCW, Jet Bussemaker, heeft haar goedkeuring aan dit voorstel gegeven. Het voorstel tot implementatie heeft instemming van de Algemene Vergadering van de leden van de Vereniging Hogescholen. Met betrekking tot BKE/SKE behelst het voorstel het volgende:

¹ D.d. 29 oktober 2012, referentienummer 12-2124 O&S

"Besloten is om conform het advies van de commissie Bruijn de ontwikkeling van een basis- resp. seniorkwalificatie examinering (BKE/SKE) centraal te stellen. Hierbij gaat het niet om het verzorgen van een dergelijke scholing als zodanig, maar om de ontwikkeling van een *programma van eisen* waaraan scholing moet voldoen. Certificering kan ook aan de orde zijn, in de zin dat het bedoelde programma van eisen betrekking heeft op de minimumvoorwaarden voor certificering.

Omdat toetsing de verantwoordelijkheid van de individuele hogescholen is, is deze tweede lijn (mede vanwege het ontwikkelingskarakter) vooral een expertprogramma. Daartoe zal het bestuur van de HBO-raad² een expertgroep benoemen die tot opdracht krijgt om een module BKE te ontwikkelen, met als doel deze op te nemen in het (eveneens in ontwikkeling zijnde) deskundigheidsbevorderingsprogramma basisvaardigheden didactische bekwaamheid (BDB). Tevens wordt deze expertgroep gevraagd het bestuur te adviseren over de noodzaak om een separate module SKE te ontwikkelen gericht op specifieke functies die senioriteit vereisen. Op het terrein van deskundigheidsbevordering t.a.v. toetsing en examinering gebeurt er binnen diverse hogescholen al erg veel. Daarom is in deze programmaliijn op korte termijn veel te realiseren ('quick wins') wanneer de expertgroep de bestaande professionaliseringstrajecten transparant(er) maakt en de beste elementen daaruit samensmeedt tot één plan van eisen.

Eén en ander moet er toe leiden dat uiterlijk 1 september 2013 een voorstel is opgesteld voor de (verdere) professionalisering van examinatoren, in de vorm van een programma van eisen voor de module BKE van het algemene deskundigheidsbevorderingsprogramma BDB, alsmede een advies aan het bestuur van de HBO-raad over de noodzaak om een separate module SKE te ontwikkelen.

Indien de expertgroep adviseert een dergelijke module te ontwikkelen, bevat het advies tevens een programma van eisen voor deze module alsmede een inventarisatie van de functie waarvoor het hebben behaald van een SKE-kwalificatie noodzakelijk is."

(Brief HBO-Raad aan Staatssecretaris Zijlstra, referentienummer 12-2124, O&S p. 4)

Opdracht aan de expertgroep

Uit de brief aan de minister kan de opdracht aan de expertgroep BKE/SKE als volgt worden benoemd:

- Stel een advies op aan het bestuur van de Vereniging Hogescholen over de noodzaak om een separate module SKE te ontwikkelen.
- Stel een voorstel op voor de (verdere) professionalisering van examinatoren, in de vorm van een *programma van eisen* voor de module BKE van het algemene deskundigheidsbevorderingsprogramma Basiskwalificatie Didactische Bekwaamheden. Tijdens het werken aan dit voorstel is tegelijkertijd een commissie aan de slag om een protocol rondom de wederzijdse erkenning van professionaliseringsarrangementen van hogescholen op te leveren. In de conceptversie van het *Protocol inzake wederzijdse erkenning basiskwalificatie didactische bekwaamheid* van september 2013 is de relatie tussen BDB en BKE als volgt benoemd:

"De docent neemt bestaande toetsen af voor zijn vakgebied en beoordeelt de resultaten van zijn studenten. Toetsen en beoordelingsformulieren worden voorzien van toetstechnisch commentaar en toetsuitslagen worden geanalyseerd. Er wordt zowel formatief als summatief getoetst. De Basiskwalificatie Examinering (BKE), die voor het HBO in ontwikkeling is, vormt hierbij de leidraad zodra deze klaar is." (p.3)

² Inmiddels is deze naam gewijzigd in Vereniging Hogescholen.

2. Samenstelling en werkwijze expertgroep

Om de opdracht uit te voeren is een expertgroep BKE/SKE samengesteld. De expertgroep bestaat uit een kerngroep en een klankbordgroep (zie Tabel 1).

Tabel 1.

Samenstelling expertgroep BKE/SKE

Naam	Rol	Functieomschrijving
Dominique Sluijsmans	Voorzitter expertgroep - expert - kerngroep	Zuyd Hogeschool Lector Professioneel Beoordelen
Remko van der Lei	Expert - kerngroep	Hanzehogeschool Groningen Expert hoger onderwijs / onderwijskundig adviseur Project Kwaliteit van Toetsen en Beoordeling
Jetske ten Caat - Schokker	Extern deskundige - kerngroep	Politieacademie Toetskundige, Staf Onderwijs
Paula Bouw	Docent - kerngroep	Hogeschool Leiden Lerarenopleidingen Gezondheidszorg en Welzijn en Omgangskunde Lid expertgroep toetsing en examinering
Dineke Behrend	Secretaris - kerngroep	Hogeschool Windesheim Voorzitter examencommissie SPH-MWD-CMV; coördinator jaarstage SPH
Anton Béguin	Extern deskundige - klankbordgroep	CITO Wetenschappelijk directeur CITO
Desirée Joosten - ten Brinke	Extern deskundige - klankbordgroep	Fontys Lerarenopleiding Tilburg Lector Eigentijds toetsen en beoordelen Open universiteit

Voorafgaand aan de eerste bijeenkomst is een inventarisatie gemaakt van bestaande toetscursussen³. Vanuit deze inventarisatie van deskundigheidsbevordering op het gebied van toetsing heeft de expertgroep een aantal conclusies getrokken:

- Er is veel aanbod op het gebied van toetsen vanuit de hogescholen, dat elkaar grotendeels overlapt. Er is echter nog zeer weinig BKE/SKE-aanbod.
- Doelgroepen van het cursusaanbod zijn docenten die examens ontwikkelen, afnemen en beoordelen en leden van toets- en examencommissies.
- Het cursusaanbod bestaat vooral uit cursussen of scholingsdagen, maar is in veel mindere mate een integraal onderdeel van de dagelijkse onderwijspraktijk.
- Het is niet altijd helder wie de professionalisering aanbiedt c.q. uitvoert.
- Het aanbod is grotendeels gericht op toetsontwerp, minder op het niveau van toetsprogramma en toetsbeleid⁴.
- Het aanbod richt zich vooral op het ontwerp van verschillende soort toetsvormen (meerkeuze-toetsen, open vragen toetsen, proeven van bekwaamheid, criteriumgericht interview, enzovoort).
- Het is onvoldoende helder of de cursussen van hogescholen vergelijkbaar zijn in kwaliteit en verwachte leeruitkomsten.
- Het is niet helder of en hoe certificering van leeruitkomsten plaatsvindt.
- Het is niet helder hoe en door wie de effecten van de scholing worden geborgd.

Deze conclusies heeft de expertgroep als basis gebruikt voor een brainstorm in de eerste werksessie.

³ De volledige inventarisatie is op te vragen bij de secretaris van de expertgroep

⁴ Deze constatering heeft geleid tot de keuze een SKE te ontwikkelen.

De expertgroep heeft zich vervolgens in een aantal intensieve werksessies in de periode januari-september 2013 op de volgende manier gewerkt aan haar opdracht:

- In de startbijeenkomst is er kennisgemaakt en is de opdracht van de Vereniging Hogescholen verder aangescherpt. Tevens is de inventarisatie besproken en zijn er werkafspraken gemaakt.
- Vervolgens is de stand van zaken opgemaakt en is een eerste aanzet gemaakt van leeruitkomsten en een overzicht van onderscheidende factoren tussen BKE en SKE. Met leeruitkomsten worden de verwachte doelen bedoeld die de BKE en SKE examinator dient te behalen⁵. De leeruitkomsten kunnen via formeel en informeel leren worden behaald.
- Tussen de bijeenkomsten hebben de leden van de expertgroep informatie uitgewisseld, per mail of via de Dropbox. De secretaris was beschikbaar voor overleg en afstemming. Na de tweede bijeenkomst heeft de expertgroep via het 'estafettemodel' gewerkt: het voorstel voor het advies rond BKE en SKE werd opeenvolgend langs de leden van de expertgroep gestuurd. De secretaris voegde bestaande ideeën samen tot een gespreksdocument. Dit tussenproduct is door alle leden van de expertgroep voorgelegd aan een aantal toetsdeskundigen uit hun netwerk ter validatie.
- Het laatste overleg heeft op 9 september plaats gevonden. In dit overleg zijn definitieve beslissingen genomen over de leeruitkomsten en het onderscheid BKE/SKE.

Het conceptvoorstel dat er na de derde bijeenkomst van de expertgroep lag, alsmede de voorlaatste versie van de rapportage, zijn voorgelegd aan de Projectgroep Externe Validering van de Vereniging Hogescholen, zodat ook daar instemming op hoofdlijnen is verkregen over werkwijze en inhoud.

De definitieve versie van het advies is op dezelfde manier tot stand gekomen. De secretaris heeft gezorgd voor afstemming op de ontwikkelingen rond BDB. Het was de bedoeling om op 1 september het advies aan de Vereniging Hogescholen op te leveren. In juni is die datum vervangen door 1 oktober, om meer tijd te hebben voor het afronden van het voorstel.

⁵ Zie voor een definitie van leeruitkomsten Lokhoff, J., Wegewijs, B., Durkin, K., Wagenaar, R., González, J., Isaacs, A. K., Donà dalle Rose, L. F., & Gobbi, M. (Eds) (2010). *Tuning educational structures. A Guide to Formulating Degree Programme Profiles*. NUFFIC: Bilbao, Groningen and The Hague.

3. Operationalisering van de opdracht

3.1 Uitwerking van de opdracht en begripsbepaling

De opdracht aan de expertgroep is een *programma van eisen* op te stellen voor een module BKE en tevens een advies te geven over de noodzaak om een separate module SKE te ontwikkelen.

De expertgroep heeft het begrip *programma van eisen* als volgt geoperationaliseerd:

- De expertgroep doet een voorstel over de vraag wat examinatoren moeten kennen en kunnen en hoe opleidingen dat kunnen borgen in termen van **leeruitkomsten en bijbehorende indicatoren**. Om tot een gedegen set van leeruitkomsten te komen, heeft de expertgroep een aantal uitgangspunten geformuleerd (zie hoofdstuk 4). Belangrijkste uitgangspunten zijn het gebruik van de kwaliteitspiramide eigentijds toetsen en beoordelen en de toetscyclus.
- De expertgroep doet een voorstel welke **beoordelingscriteria** kunnen worden gesteld aan de bewijslast die examinatoren moeten leveren om als examinator te kunnen worden gekwalificeerd.
- De expertgroep heeft bij de formulering van leeruitkomsten en indicatoren de volgende definitie van examinator aangehouden:

Met de *examinator* in het hoger beroepsonderwijs wordt iedere docent / beoordelaar bedoeld die eindverantwoordelijk is voor één of meer fasen van de toetscyclus.

3.2 De rationale voor het vormgeven van een SKE

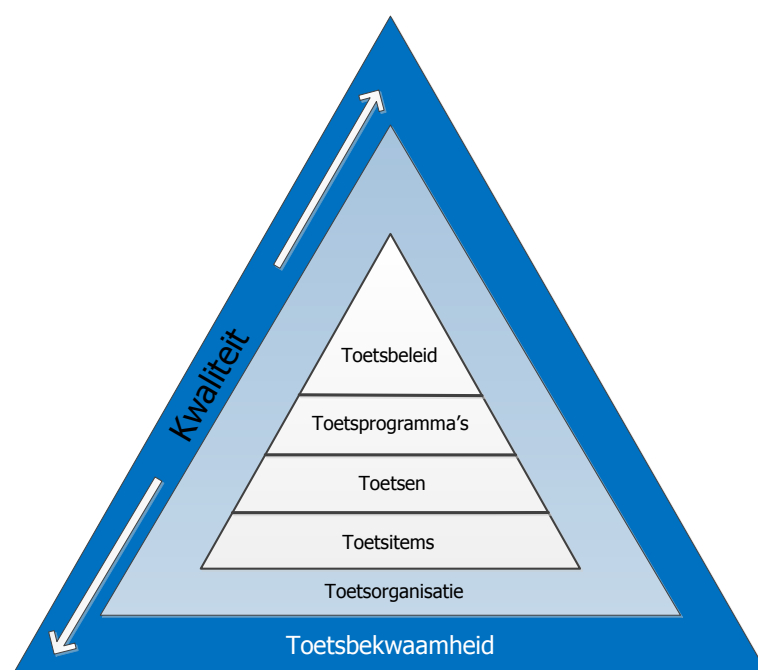
Gezien het complexe domein van toetsen en beoordelen en de noodzaak tot deskundigheid in opleidingen op het gebied van programmatisch toetsen en toetsbeleid, adviseert de expertgroep ook eisen te formuleren waaraan een SKE moet voldoen. Het programma van eisen voor een SKE is om deze reden in dit voorstel meegenomen. In het volgende hoofdstuk wordt dit advies verder uitgewerkt en onderbouwd.

4. Conceptueel kader voor de BKE en SKE

De expertgroep heeft voor de omschrijving van het programma van eisen voor een BKE en SKE twee belangrijke constructen als uitgangspunt genomen: de kwaliteitspiramide van eigentijds toetsen en beoordelen (Joosten-ten Brinke, 2011)⁶ en de toetscyclus. Deze worden in het volgende toegelicht.

4.1 De kwaliteitspiramide voor eigentijds toetsen en beoordelen

Voordat kan worden overgegaan tot de formulering van leeruitkomsten voor een BKE en SKE is het zinvol om eerst een invulling te geven aan het begrip toetskwaliteit. Hiervoor hanteren we de kwaliteitspiramide van eigentijds toetsen en beoordelen (Joosten-ten Brinke, 2011; Sluijsmans, Peeters, Jakobs & Weijzen, 2012⁷; zie Figuur 1).



Figuur 1. De kwaliteitspiramide van eigentijds toetsen en beoordelen

In deze piramide wordt kwaliteit van toetsing in zes belangrijke entiteiten vertaald:

- *Kwaliteit van toetsitems.* Hiermee wordt de kwaliteit van een taak, opdracht of vraag in een toets bedoeld. Belangrijke kwaliteitscriteria zijn 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. Deze toetsen hebben tot doel om de prestaties van de student te beoordelen, om daarmee een beslissing te nemen ten aanzien van het leren, selectie of certificering van studenten. Voorbeelden van kwaliteitscriteria waaraan toetsen dienen te voldoen 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

⁶ Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen* (lectorale rede). Tilburg: Fontys Hogescholen.

⁷ Sluijsmans, D. M. A., Peeters, A., Jakobs, L., & Weijzen, S. (2012). Kwaliteit van toetsing onder de loep. *Onderwijsinnovatie*, 4, 17-26.

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 bedoeld. 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, 2006⁸). In hoeverre draagt het toetsprogramma bij aan een duurzaam leerresultaat? Hoe is de toetsing verbonden aan kennis en vaardigheden voor de lange termijn?
- *Kwaliteit van toetsorganisatie.* Vervolgens is de hele organisatie van het toetsbouwwerk belangrijk om toetskwaliteit te kunnen garanderen. Dit komt terug in de criteria als kosten en haalbaarheid (zie Baartman, 2008⁹). Toetsorganisatie gaat over zaken als het vaststellen van de rollen, taken en verantwoordelijkheden van betrokkenen bij toetsing, waaronder toets- en examencommissies, de werkprocessen rondom de toetsing, de logistieke organisatie van de toetsing en het vaststellen van toetsroosters.
- *Kwaliteit van toetsbekwaamheid.* De toetsbekwaamheid van docenten, examencommissies en andere betrokkenen is van essentieel belang, aangezien kwaliteit van toetsing in hoge mate wordt bepaald door de kwaliteit van de beoordelaar. Elke docent moet in staat zijn om toetsinformatie te interpreteren en te gebruiken om vast te stellen waar studenten staan en hoe hij verder kan bijdragen aan het leren van de student (Straetmans, 2006¹⁰).

Kwaliteit van toetsing wordt geborgd door een sterke samenhang van alle entiteiten (aangegeven door de zijdelingse pijl). Is de kwaliteit op een bepaalde entiteit niet op orde, dan heeft dat direct implicaties voor de kwaliteit van de andere entiteiten. Als bijvoorbeeld het toetsbeleid niet transparant is, zal dit zijn uitwerking naar beneden hebben op de kwaliteit van toetsprogramma's, toetsen en de onderliggende items.

4.2 De toetscyclus

De kwaliteitsborging van verantwoord toetsen en beoordelen is afhankelijk van de kwaliteit van elke fase in het toetsproces (Joosten-ten Brinke & Sluijsmans, 2012¹¹). Dat wil zeggen: de kwaliteit van de gehele toets is zo hoog als de kwaliteit van de zwakste schakel in het toetsproces. Als bijvoorbeeld de eerste fase niet goed genoeg is doorlopen, heeft dat implicaties voor de kwaliteit van de andere fasen. De taken van de examinerator zijn verweven met de fasen van dit toetsproces, zoals beschreven in de toetscyclus (zie Figuur 2¹²). In deze toetscyclus worden zeven fasen¹³ in het toetsproces onderscheiden die op iedere toetsvorm kunnen worden toegepast, namelijk:

1. Basisontwerp
2. Construeren van onderdelen
3. Samenstellen en normeren
4. Afnemen
5. Beoordelen, verwerken en analyseren
6. Resultaat registreren en communiceren
7. Evalueren

In elke fase levert een examinerator producten op. De mogelijke producten zijn tevens in Figuur 2 weergegeven. Het doorlopen van deze fasen moet zoveel mogelijk worden gezien vanuit een *hele-taakbenadering*¹⁴ in termen van het doorlopen van een volledige toetscyclus of meerdere toetscycli,

⁸ 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.

⁹ Baartman, L. K. J. (2008). *Assessing the assessment. Development and use of quality criteria for Competence Assessment Programmes*. Utrecht: Academisch proefschrift.

¹⁰ Straetmans, G. (2006). *Bekwaam beoordelen en beslissen*. Deventer: Saxion Hogescholen.

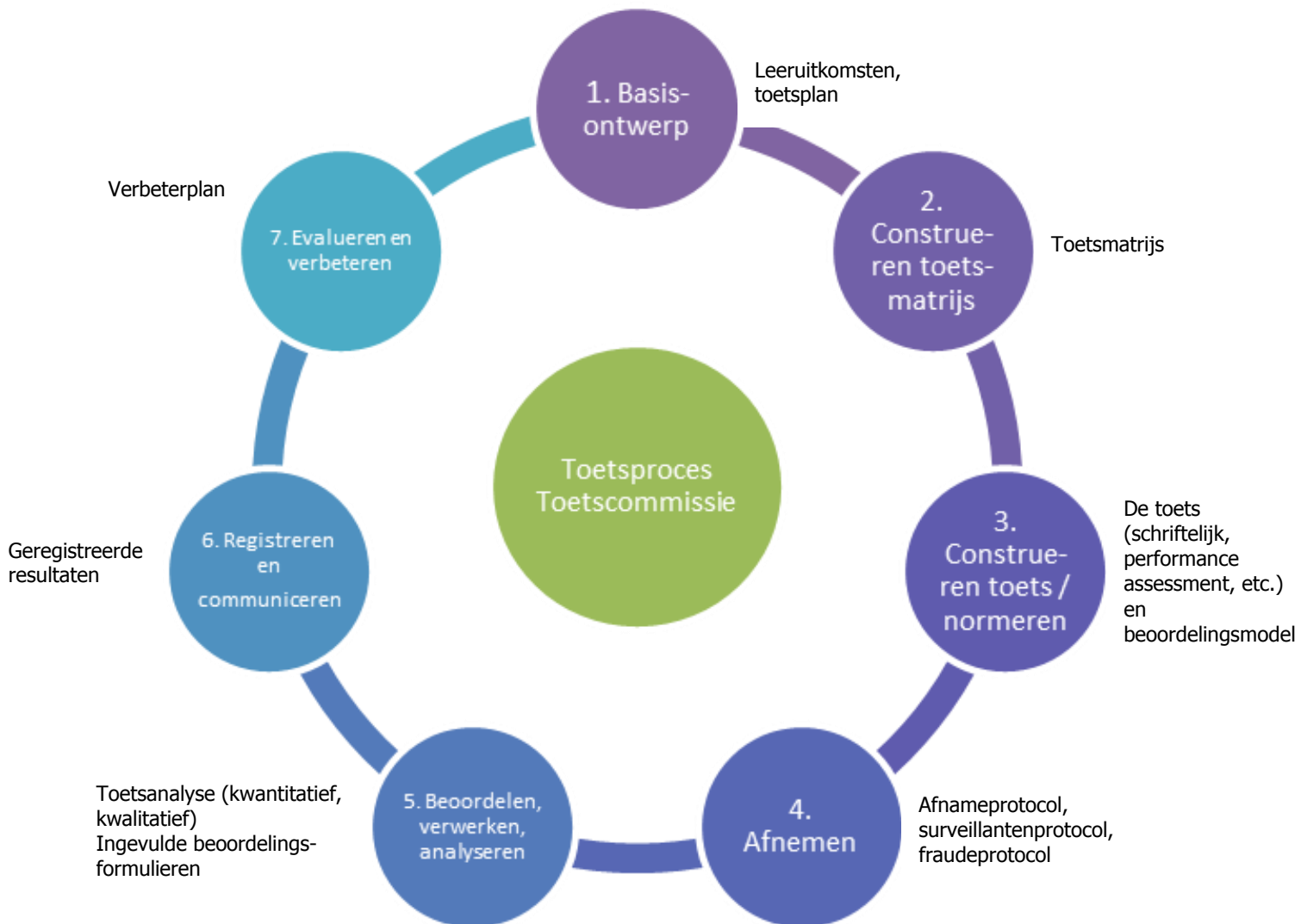
¹¹ Joosten, D., & Sluijsmans, D. M. A. (2012). Tijd voor toetskwaliteit: het borgen van toetsdeskundigheid van examencommissies. *Thema*, 16-21.

¹² Overgenomen van de Hanzehogeschool Groningen.

¹³ Zie bijlage 1 voor een toelichting.

¹⁴ Jansen-Noordman, A. M. B. & van Merriënboer, J. J. G. (1997). *Innovatief onderwijs ontwerpen. Van leertaken naar complexe vaardigheden*. Noordhoff Uitgevers B.V.

omdat het fragmentarisch aanleren van fasen uit de toetscyclus immers geen garantie geeft voor volledige borging van toetskwaliteit.



Figuur 2. De toetscyclus en bijbehorende producten.

Het cyclisch karakter impliceert dat na de fase van evalueren een nieuw basisontwerp voor een toets gemaakt kan worden of het basisontwerp wordt bijgesteld.

De taken van de examinator zijn als volgt samen te vatten:

De examinator ontwerpt en construeert (fase 1 t/m 3) een toets, neemt deze af (fase 4), beoordeelt, analyseert en registreert (fase 5 en 6). Essentieel voor de borging van de kwaliteit van het toetsproces is fase 7: de examinator evalueert het doorlopen van alle fasen en benoemt verbeteracties voor een volgende cyclus.

De toetscyclus maakt het proces van toetsing inzichtelijk: hierdoor wordt duidelijk wat de taken zijn die een examinator moet uitvoeren en uit welke producten dit blijkt.

4.3 De kwaliteitspiramide en de toetscyclus in relatie tot de BKE en de SKE

De twee conceptuele kaders beschreven in het voorgaande hebben geleid tot de volgende keuzes door de expertgroep:

- De expertgroep heeft de vier entiteiten gekozen in het *witte* vlak van Figuur 1 om de leeruitkomsten van de BKE en SKE te kaderen. Deze vier entiteiten (toetsitems, toetsen, toetsprogramma's en toetsbeleid) bieden de mogelijkheid om op een gestructureerde manier een kennisbasis te beschrijven voor zowel de BKE als de SKE, waarbij toetsitems en toetsen het accent krijgen in een BKE en toetsprogramma en toetsbeleid het accent krijgen in een SKE.
- De expertgroep heeft de toetscyclus zoals weergegeven in Figuur 2 gekozen om de leeruitkomsten van de BKE en SKE te kaderen, waarbij de BKE zich richt op het doorlopen van de toetscyclus binnen het eigen vak/module voor een specifieke toetsvorm, de SKE richt zich op de samenhang tussen meer toetsvormen op programmaniveau.

5. Het onderscheid tussen de BKE en de SKE

Voordat er concrete leeruitkomsten met bijhorende indicatoren kunnen worden geformuleerd is het nodig het onderscheid tussen het BKE- en het SKE-niveau te duiden. Het verschil tussen de BKE en SKE kan worden geduid met vier complexiteitsfactoren:

1. *Het betrekkningsniveau van de toetsdeskundigheid:* binnen welke context moet de toetsdeskundigheid van de examiner worden aangetoond?
2. *Niveau van begeleiding:* welke mate van begeleiding wordt van de examiner verwacht?
3. *Het verantwoordelijkheidsniveau:* wat is de zwaarte van beslissingen die door de examiner moet worden genomen?
4. *De rol:* welke rol vervult de examiner in de organisatie?

In de onderstaande tabel worden de verschillen tussen de BKE en SKE uit een gezet. In Tabel 2 wordt op vier factoren aangegeven waarin de BKE en SKE van elkaar verschillen.

Tabel 2.
Onderscheidende factoren BKE en SKE

Factor	BKE	SKE
Betrekkningsniveau van de toetsdeskundigheid	Specifieke deskundigheid voor toetsing binnen de eigen onderwijseenheid (c.q. alle toetsen binnen de betreffende onderwijseenheid). Deskundig op de piramidelagen toetsitems en toetsen.	Bredere deskundigheid voor toetsing dan de toetsing van het eigen vak. Deskundig op alle vier de piramidelagen, met de nadruk op toetsprogramma en toetsbeleid.
Begeleiding	De examiner raadpleegt collega's bij alle fasen van de toetscyclus.	Collegiale intervisies en begeleiden BKE'ers.
Verantwoordelijkheidsniveau	Verantwoordelijk voor beslissingen over studenten op niveau van vak/onderwijseenheid.	Verantwoordelijk voor beslissingen over studenten op niveau van curriculum (propedeuse, hoofdfase, afstudeerfase).
Rol	Examinator, docent, onderwijsontwikkelaar, lid toetscommissie.	Examinator, kwaliteitsborger, senior docent, lid toetscommissie/examencommissie.

6. Leeruitkomsten van de BKE en de SKE

Op basis van het onderscheid tussen BKE en SKE zoals geformuleerd in Tabel 2 kunnen leeruitkomsten en bijbehorende indicatoren worden geformuleerd. In Tabel 3 en 4 staan de leeruitkomsten en indicatoren voor respectievelijk de BKE en de SKE beschreven.

Tabel 3.

Leeruitkomsten en indicatoren voor de BKE

Piramidelaag	NR	Beschrijving leeruitkomst	Indicatoren*
Toetsitems en toetsen	BKE-1	Is in staat de toetscyclus toe te passen voor het toetsen van leerresultaten van studenten binnen een vakgebied / module / onderwijseenheid voor een specifieke toetsvorm en levert hiervoor adequaat bewijs in de vorm van een toetsontwerp. Onderbouwt per fase uit de toetscyclus de gemaakte keuzes aan de hand van relevante toetskundige literatuur.	<ul style="list-style-type: none"> Beschrijft per fase van de toetscyclus doel en producten. Bepaalt in overleg met collega's (examinatoren BKE en SKE) de te ontwerpen toets en maakt een planning voor de daarbij behorende werkzaamheden. Verzamelt de kaders voor de te ontwerpen toets: leeruitkomsten, toetsbeleid, toetsplan en toetsmatrijs. Stelt de leeruitkomsten waar nodig bij. Kiest voor de te ontwerpen toets een passende toetsvorm gegeven de leeruitkomsten. Stelt de toetsmatrijs op (of past een bestaande toetsmatrijs aan), gegeven leeruitkomsten en toetsvorm. Stelt volgens het vier-ogenprincipe de toets samen op basis van de toetsmatrijs cf. kwaliteitseisen. Stelt de normering vast en maakt een beoordelingsmodel bij de toets in overleg met collega's. Neemt de toets af. Beoordeelt de toets aan de hand van het beoordelingsmodel, verwerkt en analyseert de resultaten. Geeft adequate, informatierijke feedback aan de student naar aanleiding van de toetsresultaten. Voert de cijfers in van de beoordeelde toets. Evalueert alle fasen van de toetscyclus.
Toetsprogramma	BKE-2	Kan de positie van de toets in het totale toetsprogramma aangeven.	<ul style="list-style-type: none"> Geeft aan welke toetsvormen in het programma worden gegeven. Geeft aan welke positie het vak / de onderwijseenheid in het programma heeft. Bepaalt consequenties van het programma voor de inhoud, de vorm en het niveau van de toets.
Toetsbeleid	BKE-3	Kan consequenties van het toetsbeleid benoemen voor de eigen toets.	<ul style="list-style-type: none"> Kan de visie op toetsing in de organisatie benoemen. Vertaalt de kernpunten van de visie naar de eigen toets.

Tabel 4.
Leeruitkomsten en indicatoren voor de SKE

Piramidelaag	NR	Beschrijving leeruitkomst	Indicatoren*
Toetsitems en toetsen	SKE-1	Is in staat de toetscyclus toe te passen voor het toetsen van leerresultaten van studenten binnen verschillende vakoverstijgende modules/onderwijseenheden met verschillende toetsvormen en levert hiervoor adequaat bewijs in de vorm van meerdere toetsontwerpen.	<ul style="list-style-type: none"> Voor het doorlopen van de toetscyclus gelden in elk geval dezelfde indicatoren en criteria als voor de BKE (leeruitkomst B1). Als extra criterium voor de SKE geldt dat de examinerator zelf kan ontwerpen en redeneert vanuit de samenhang van het programma en het toetsbeleid van de opleiding. Stelt voor collega's richtlijnen op voor het hanteren van beoordelingsmodellen. Kan toetsen van collega's beoordelen op kwaliteitscriteria.
Toetsprogramma	SKE-2	Is in staat de kwaliteit van het toetsprogramma te analyseren en verbeteracties te formuleren.	<ul style="list-style-type: none"> Stelt in samenwerking met collega's een toetsplan voor de gehele opleiding op. Draagt zorg voor de kwaliteitsborging van tentamens en examens. Evalueert (de samenhang van) toetsresultaten en beslist over verbetermaatregelen. Stelt vast welke scholingsbehoeften m.b.t. examineren in het team zijn en formuleert hier leeruitkomsten voor.
Toetsbeleid	SKE-3	Levert een bijdrage aan het toetsbeleid van de gehele opleiding.	<ul style="list-style-type: none"> Stelt richtlijnen op voor het ontwikkelen, afnemen en beoordelen van tentamens en examens.
	SKE-4	Is in staat de kwaliteit van toetsen van collega-examinatoren te evalueren en aanbevelingen te doen gerelateerd aan de visie en het beleid op toetsen van de instelling. Kan collega's op deze manier begeleiden in het verbeteren van toetsen.	<ul style="list-style-type: none"> Ondersteunt en begeleidt collega-examinatoren (BKE) bij het doorlopen van de toetscyclus. Geeft de collega-examinator feedback op de kwaliteit van de toets en bijbehorende toetsproducten (zie Figuur 2).

***De genoemde indicatoren kunnen mogelijke bewijzen leveren van beheersing van de gewenste leeruitkomst. Niet alle indicatoren hoeven noodzakelijk zichtbaar te zijn, maar wel geldt dat ze adequaat uitgevoerd zouden kunnen worden. De wel zichtbare indicatoren moeten aannemelijk maken dat de leeruitkomst in voldoende mate wordt beheerst.**

7. Criteria voor de bewijslast van de BKE en de SKE

De leeruitkomsten en indicatoren zijn richtinggevend voor de bewijslast die een examiner moet kunnen laten zien om gekwalificeerd te worden als BKE of SKE. De expertgroep stelt de volgende beoordelingscriteria voor die aan deze bewijslast moeten worden gesteld. Deze zijn afkomstig van de gangbare criteria die aan toetsing kunnen worden gesteld. De bewijslast is:

- *Valide*: de bewijslast die wordt geleverd is een adequate vertaling van de gewenste leeruitkomsten zoals gepresenteerd in Tabel 3 en 4.
- *Betrouwbaar*: de bewijslast is van voldoende omvang om een accurate beoordeling mogelijk te maken van de gewenste leeruitkomsten.
- *Transparant*: de bewijslast is duidelijk en inzichtelijk voor collega's en beoordelaars.
- *Bruikbaar*: de bewijslast geeft blijk van directe toepasbaarheid in de onderwijspraktijk.
- *Authentiek*: de bewijslast geeft blijk van voldoende aansluiting bij beroepspraktijk.
- *Actueel*: de bewijslast is van actuele waarde.

Voor de BKE betekent dit:

- Dat uit de bewijslast is af te leiden dat de BKE'er een brede basiskennis bezit over toetsen. Deze kennis is nodig om de stappen uit de toetscyclus goed te kunnen begrijpen en doorlopen. Dit betekent dat de BKE'er onderbouwde voorstellen doet voor verbetering van iedere fase in de toetscyclus. In Bijlage 2 zijn voorbeelden opgenomen van bronnen die richtinggevend zijn om de leeruitkomsten te kunnen behalen.
- Dat de bewijslast producten bevat zoals weergegeven in Figuur 2 (pagina 11).
- Dat de bewijslast aantoont dat de BKE'er een doorleefd beeld heeft van toetskwaliteit in de context van de eigen onderwijspraktijk. Dit betekent de BKE'er de leeruitkomsten aantoont aan de hand van voor hem/haar relevante toetsvormen.

Voor de SKE betekent dit:

- Dat uit de bewijslast is af te leiden dat de SKE'er een brede basiskennis bezit over toetsen, toetsprogramma en toetsbeleid. Dit betekent dat de SKE'er onderbouwde voorstellen geeft voor verbetering van toetsprogramma en toetsbeleid. In Bijlage 2 zijn voorbeelden opgenomen van bronnen die richtinggevend zijn om de leeruitkomsten te kunnen behalen.
- Dat de bewijslast producten bevat als samenhangend toetsprogramma en een toetsbeleid.
- Dat de bewijslast aantoont dat de SKE'er een doorleefd beeld heeft van toetskwaliteit in de context van de betreffende opleiding. Dit betekent dat de SKE'er de leeruitkomsten aantoont aan de hand van voor hem/haar relevant toetsprogramma en toetsbeleid.

NB: De expertgroep benadrukt dat de BKE'er mag werken vanuit bestaand toetsmateriaal om leeruitkomsten aan te tonen.

Bijlage 1. Toelichting op de toetscyclus

Fase 1: Basisontwerp

De toetscyclus start met het ontwerpen van een toets. De toets dient congruent te zijn met het gegeven onderwijs en de in dit studieonderdeel te bereiken leeruitkomsten. De inhoud van de toets, het niveau en de toetsvorm moeten passen bij de onderwijsvisie en het didactisch concept van de opleiding. Het basisontwerp van de toets geeft een eerste idee van de inhoud en de vormgeving van de toets en laat zien wat de ontwikkelaars met de toets willen bereiken.

Fase 2: Construeren onderdelen

Bij de ontwikkeling van een nieuw studieonderdeel kan gestart worden met de ontwikkeling van een eerste toets. Dit varieert van het ontwerpen van een complete case inclusief multimedia tot het ontwikkelen van een meerkeuzevraag over een bepaald onderwerp. Een nuttig instrument bij het construeren van een toets is de toetsmatrijs. Het construeren van toetsonderdelen moet uiteraard in overeenstemming zijn met het basisontwerp van de toets. De toetsonderdelen kunnen gescreend worden door vakgenoten en/of een toetscommissie. De eerste toets kan daarna uitgebreid worden tot een geïntegreerde verzameling toetsen of toetsonderdelen, behorend bij het studieonderdeel. De verzameling toetsen of toetsonderdelen wordt opgenomen in het toetsarchief.

Fase 3: Samenstellen, normeren

Als de toets in grote lijnen geconstrueerd is, kan de ontwikkeling van de eigenlijke toets plaatsvinden: samenstellen van vragen, formuleren van opdrachten, verzamelen van relevant ondersteunend materiaal, beschrijven van casuïstiek, etc.. Als de toets klaar is, moet een passende normering plaatsvinden. De cesuur moet worden vastgesteld en er moet een scoringsvoorschrift beschikbaar zijn. In geval van groepsopdrachten moet duidelijk zijn hoe individueel getoetst wordt. Indien de criteria al niet zijn geformuleerd in eerdere fasen 1 of 2 dan gebeurt dat nu. In deze fase kan ook worden nagedacht hoe studenten kunnen worden betrokken, bijvoorbeeld in het opstellen van beoordelingscriteria.

Fase 4: Afnemen

Bij het afnemen van de toets vindt de directe confrontatie met de student plaats. Het is van belang om zorgvuldig te zijn in het creëren van de juiste toetscondities, zodat het resultaat van de toets daadwerkelijk een weerspiegeling is van de kwaliteiten van de studenten. Het afnemen kan bestaan uit meerdere momenten, bijvoorbeeld wanneer er sprake is van tussentijdse formatieve momenten op tussenproducten of gedrag.

Fase 5: Beoordelen, verwerken en analyseren

De afgenomen toets wordt beoordeeld aan de hand van de normering opgesteld bij het construeren van de toets. Hierdoor ontstaat inzicht in de moeilijkheidsgraad, betrouwbaarheid en validiteit van de toets en toetsonderdelen. Dit kan aanleiding geven tot het herzien van scores en wellicht ook van toetsonderdelen in het toetsarchief. Iedere student ontvangt een eindscore, evt. na inzage of bespreking van de afgenomen toets en het geven van feedback. De examencommissie stelt de toetsresultaten formeel vast.

Fase 6: Resultaat registreren en communiceren

De toetsresultaten worden geadministreerd. De student krijgt passende feedback en feedforward op de prestatie. Het moet duidelijk zijn of deze feedback wordt begrepen door de student.

Fase 7: Evalueren

Bij het evalueren is het de bedoeling een verzameling toetsen te beoordelen op hun kwaliteit in onderlinge samenhang (o.a. moeilijkheidsgraad, betrouwbaarheid en validiteit). Een belangrijk algemeen borgingselement is de confrontatie van toetsen met empirische gegevens (kwantitatief en/of kwalitatief), zoals evaluaties van de toetsing of analyses van de toetsresultaten zelf, die er toe leiden dat onderdelen worden verbeterd en zwakke plekken worden gesignaleerd. Op grond van de inzichten in deze evaluatiefase kan men nieuwe ideeën ontwikkelen omtrent het ontwerp van een nieuwe toets / nieuwe toetsonderdelen, bijvoorbeeld door middel van collegiale consultatie en intervisie.

Bijlage 2. Aanbevolen toetsbronnen

Nederlandse bronnen

- Baartman, L., Prins, F., Kirschner, P., & Van der Vleuten, C. (2007). Kwaliteitsmeting van Competentie Assessment Programma's via zelfevaluatie. *Onderwijsinnovatie*, 1, 17-26.
- Joosten-ten Brinke, D. (2011). *Eigentijds toetsen en beoordelen*. Tilburg: Fontys Lerarenopleiding Tilburg.
- Geerts, W. & van Kralingen, R. (2011). *Handboek voor leraren* (pp. 155-201). Bussum: Couthino.
- Teelen Kennismanagement (2011). *Toetsontwikkeling in de praktijk, hoe maak ik goede vragen en toetsen?* Wilp: Teelen B.V.
- Sluijsmans, D. M. A. (2013). *Verankerd in leren. Vijf bouwstenen voor professioneel beoordelen in het hoger onderwijs*. Heerlen: Zuyd Hogeschool.
- Van Berkel, H., & Bax, A. (2002). *Toetsen in het hoger onderwijs*. Bohn Stafleu Van Loghum.
- Van der Vleuten, C., Schuwirth, L. (2011). Een model voor programmatisch toetsen. *Examens*, 3, 5-9.
- Wools, S., Sanders, P., Eggen, T., Baartman, L., & Roelofs, E. (2011). Evaluatie van een beoordelingssysteem voor de kwaliteit van competentie-assessments. *Pedagogische Studiën*, 88(1), 23-40.

Engelstalige bronnen

- Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities. *Studies in Educational Evaluation*, 35, 130-149.
- Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-73.
- Boud, D. and Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Sydney: Australian Learning and Teaching Council.
- Brown, G., & Hattie, J. (2009). *Understanding teachers' thinking about assesment: Insights for developing better educational assessments*. Paper presented at the Annual Meeting of the National Council of Measurement in Education. VS: San Diego.
- Draaijer, S., & Hartog, R. (2007). Guidelines for the design of digital closed questions for assessment and learning in higher education. *e-Journal of Instructional Science and Technology*, 10(1).
- Haladyna, T. M., & Downing, S. M. (2006). *Handbook of Test Development*. Mahwah: NJ: Lawrence Erlbaum.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders*. Needham, MA: Allyn and Bacon.
- Schmeiser, C.B. & Welch, C.J. (2006). Test development. In R. L. Brennan (2006). *Educational Measurement* (4th Ed). American Council on Education/Oryx Press Series on Higher Education.
- Wiggins, G. (1998). *Educative Assessment. Designing Assessments To Inform and Improve Student Performance*. San Francisco: Jossey-Bass Publishers.

Aanbevolen links

- www.toetsenopschool.nl
- <http://toetswijzer.kennisnet.nl/html/toetsconstructie/home.htm>
- <http://testing.wisc.edu/Handbook%20on%20Test%20Construction.pdf>