

“We moeten studenten voorbereiden op hun toekomst en niet op ons verleden”

Mission Statement East Valley Institute of Technology, Mesa, Az

Kwaliteit van toetsing in het Hoger Onderwijs

naar een versterkte rol van examencommissies

Dienst Onderwijs en Onderzoek

Mieke Jaspers
Els van Zijl
2011

Kwaliteit van toetsing in het Hoger Onderwijs

naar een versterkte rol van examencommissies

Dienst Onderwijs en Onderzoek

Mieke Jaspers
Els van Zijl
2011

Colofon

Uitgave	Fontys Hogescholen, Dienst Onderwijs en Onderzoek
Titel	Kwaliteit van toetsing in het Hoger Onderwijs
Subtitel	Naar een versterkte rol van examencommissies
Auteurs	Mieke Jaspers & Els van Zijl
Eindredactie	Mieke Jaspers
Met dank aan	Marlies Hesseling, Desirée Joosten-ten Brinke
Lay-out	Grafische Producties, Dienst Marketing en Communicatie
Bestellingen	Secretariaat Dienst Onderwijs en Onderzoek
Telefoon	0877 878 833
Emailadres	dienstonderwijsenonderzoek@fontys.nl

© 2011 Dienst Onderwijs en Onderzoek. Alle rechten voorbehouden. All rights reserved.

Hoewel dit materiaal met zorg is samengesteld, aanvaarden auteurs, noch de Dienst Onderwijs en Onderzoek, noch Fontys enige aansprakelijkheid voor schade ontstaan door eventuele fouten en/of onvolkomenheden in de publicatie.

Inhoudsopgave

Inleiding	5
1. Rol examencommissie	7
1.1 Taakverschuiving van procedure naar kwaliteit	7
2. Het brede perspectief van toetsen en beoordelen	11
2.1 Consistentie visie op leren en toetsen	11
2.2 Consistentie toetsprogramma en onderwijsprogramma	12
2.3 Bevordering studiesucces	15
3. Functies van toetsing	18
4. Kwaliteit van toetsing en examinering	21
4.1 Beleid	22
4.2 Kwaliteit van het toetsprogramma	23
4.2.1 Competentie Assessment Programma (CAP)	24
4.2.2 Methodemix	25
4.3 Kwaliteit van toets- en beoordelingsinstrumenten	27
4.4 Kwaliteit van taken, opdrachten of items	27
4.5 Digitale toetsing	28
4.6 Cultuur van gezamenlijk leren en ontwikkelen	29
5. Kwaliteitssysteem toetsing en examinering	30
5.1 PDCA-cyclus	30
5.2 In elkaar grijpende PDCA-cycli	31
5.3 Voortdurend verbeteren en vernieuwen	32
6. Toetsvormen in een competentiegerichte leeromgeving	35
6.1 Indeling in toetsvormen: een werkmodel	35
6.2 Beschrijving van toetsvormen	37
Bijlage 1 Format Jaarverslag examencommissie Fontys Hogescholen	56
Bijlage 2 Competentie Assessment Programma	59
Bijlage 3 Checklist Geautomatiseerd toetsen	62
7. Referenties	63

Inleiding

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. In januari 2011 is tevens het nieuwe accreditatiestelsel in werking getreden waarbij 'Toetsing en gerealiseerde eindkwalificaties' als aparte standaard is opgenomen. De beoordeling van deze standaard moet voldoende zijn om voor accreditatie in aanmerking te komen. Een onvoldoende beoordeling voor dit onderwerp leidt tot verlies van accreditatie¹. Tevens moet de examencommissie zich bij accreditatie kunnen verantwoorden over het aspect 'toetsen en beoordelen'. Een en ander vraagt inhoudelijke versterking van de rol van de examencommissie ten aanzien van toetsing en examinering.

Ondanks de verscherpte aandacht voor de kwaliteit van toetsing en examinering merkt het Interstedelijk Studenten Overleg (ISO, 2009) echter 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. Toetsen en beoordelen dienen meer te zijn dan het geven van een cijfer. Het is eerder een 'front end' probleem dan een 'back end' probleem. De waarde van de toetsuitslag ligt niet uitsluitend in de vraag: 'voldoende of onvoldoende?' maar in de informatie die de student krijgt om het eigen leerproces te sturen. Toetsing is tegelijkertijd een van de kwetsbaarste onderdelen van het onderwijsproces: het onderwijs is net zo goed als de toetsen zijn. Zolang er echter getuigschriften met wettelijk civiel effect worden verstrekt en zolang er mensen in een leersituatie zitten die zicht willen krijgen op hun leerproces, zullen er toetsen zijn. Toetsing vormt het kloppend hart en de ruggengraat van het curriculum en is daarmee een sterk bepalende factor voor de kwaliteit van het onderwijs.

Deze publicatie is bedoeld om examencommissies en onderwijsgevend handvatten aan te reiken om de kwaliteit van toetsing en examinering te verbeteren en te borgen. Hierdoor kunnen zij een bijdrage leveren aan het bevorderen van studiesucces van studenten en aan een duurzaam en breed gedragen vertrouwen in de kwaliteit van de getuigschriften.

¹ Beoordelingskaders accreditatiestelsel hoger onderwijs 6 december 2010 p. 57

1. Rol examencommissie

De Inspectie van het Onderwijs en de Onderwijsraad hebben de afgelopen jaren kritische rapporten uitgebracht over het functioneren van examencommissies, de kwaliteit rond de afgifte van getuigschriften en de examinering. De Inspectie van het Onderwijs (2009) en de Onderwijsraad (2004; 2006) constateerden dat de examinering en tentaminering in het hoger onderwijs in bepaalde gevallen te weinig kwaliteitsgarantie bieden en niet transparant genoeg zijn. Over de taak van de examencommissie en haar rol bij de kwaliteit van de examens stelde de inspectie vast dat examencommissies in de praktijk vooral getuigschriften afgeven, vrijstellingen geven en klachten over examens behandelen. Ze besteden te weinig aandacht aan kwaliteitsborging en kwaliteitsbeleid rond examens en toetsen. Uit een brede enquête van de inspectie blijkt dat examencommissies zelf vinden dat de kwaliteit van examinering en toetsing veel meer aandacht verdient. De Onderwijsraad concludeert dat examinering een ‘opleidingsafhankelijke aangelegenheid is met weinig kwaliteitsborging en een geringe betrokkenheid van externe deskundigen’.

Naast deze rapporten zijn er nog meer argumenten om de positie van de examencommissie te versterken. In de eerste plaats wordt het belang van meer variëteit en maatwerk in het hoger onderwijs onderschreven vanuit de visie van een leven lang leren. In de tweede plaats krijgt de EVC-procedure een steeds nadrukkelijker plaats. Door de erkenning van verworven competenties (EVC) neemt naar verwachting het aantal flexibele leerroutes toe. De combinatie van deze factoren vraagt om het versterken van de positie van de examencommissie. Het is immers de examencommissie die beoordeelt of het niveau van de afgestudeerde conform het gestelde eindniveau is gerealiseerd.

Als gevolg van de wetswijziging heeft de examencommissie een autonome verantwoordelijkheid voor de borging van de kwaliteit van de toetsing gekregen. In die verantwoordelijkheid spreekt het visitatiepanel tijdens een accreditatie met leden van de examencommissie.

1.1 Taakverschuiving van procedure naar kwaliteit

Een van de uitgangspunten bij de Wet Versterking Besturing is het vergroten van de onafhankelijkheid en de deskundigheid van de examencommissie. Instellingen zijn verplicht om een examencommissie in te stellen, te handhaven en de positie ervan te versterken. De examencommissie kan zich uitstrekken over één of meer opleidingen. De drie belangrijkste aspecten die in de wet zijn opgenomen over de examencommissie zijn:

1. Onafhankelijkheid en deskundigheid.
2. Breder takenpakket.
3. Verhouding examencommissie- instellingsbestuur.

Ad 1. Onafhankelijkheid en deskundigheid

De wet schrijft dat de examencommissie onafhankelijk en deskundig moet zijn. Aan het instellingsbestuur, (bij Fontys gemandateerd aan de instituutsdirecteuren), is opgedragen ervoor zorg te dragen dat de examencommissie onafhankelijk en deskundig kan functioneren. Dit impliceert dat de directeur de examencommissie daartoe in staat stelt en zich moet onthouden van inmenging in de taken van de examencommissie. Het gaat om een functioneel onafhankelijk orgaan dat niet in een hiërarchische relatie staat ten opzichte van de directeur.

Vanuit het oogpunt van transparantie, onafhankelijkheid én professionalisering is het aan te bevelen dat gebruik wordt gemaakt van externe deskundigheid, niet alleen wat betreft de examinatoren maar ook ten aanzien van de examencommissie. Dat kan bijvoorbeeld door externe leden op te nemen of een externe deskundige te consulteren. Daarmee wordt de externe legitimatie van toetsing bovendien vergroot. Een wettelijke verplichting tot het opnemen van externen is niet opgenomen. Daarvoor zijn opleidingen te verschillend. Wel zal door het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) worden verkend of en hoe, naar Brits model, gewerkt kan worden met een zogenaamde ‘pool’ van docenten die bij andere instellingen als externe deskundige optreden. Een onafhankelijke en deskundige examencommissie draagt er uiteindelijk aan bij dat de professional, zijnde de inhoudelijk deskundige, een stevige positie inneemt binnen de instelling. Daarbij is het niet aan te bevelen dat personen met een managementfunctie, die verantwoordelijk zijn voor de financiële situatie van de instelling, zitting hebben in de examencommissie. De onafhankelijkheid en de transparantie kunnen daarmee in het geding komen.

Om het onafhankelijke karakter van de examencommissie ten opzichte van het instellingsbestuur te benadrukken, is niet alleen gekozen voor het expliciet formuleren van het belang en de exclusieve taak van de examencommissie. Er is ook uitdrukkelijk gekozen voor het in stand houden van het College van beroep voor de examens als intern orgaan waar geschillen naar aanleiding van beslissingen van de examencommissie en examinator worden beoordeeld.

Vanuit de verantwoordelijkheid voor het waarborgen van het onafhankelijk en deskundig functioneren van de examencommissie moet het instellingsbestuur bij de benoeming van de leden van de examencommissie er op toe te zien dat de examencommissie als geheel over voldoende deskundigheid beschikt ten aanzien van toetsing, regelgeving (o.a. in de onderwijs- en examenregeling: OER) en interne en externe kwaliteitszorg en daar waar nodig, professionalisering faciliteert.

Ten behoeve van de transparantie is in de Wet Versterking Besturing opgenomen dat de examencommissie jaarlijks verslag doet van de werkzaamheden aan het instellingsbestuur. Daarmee wordt benadrukt dat de examencommissie een onafhankelijk opererend orgaan is binnen de instelling. Zie bijlage 1: Format Jaarverslag examencommissie.

Ad 2. Breder takenpakket

Tot nu toe was de examencommissie vooral verantwoordelijk voor het vaststellen van regels die een goede gang van zaken rond toetsing waarborgen: examinatoren aanwijzen, tentamens en examens organiseren en coördineren. Tot 1 september 2010 was het zo dat de wet bepaalde dat de examencommissie optioneel richtlijnen en beoordelingsnormen (voor tentamens) kon vaststellen. Hierbij ging het om bindende instructies van algemene of concrete aard. Sinds september 2010 is het een taak van de examencommissie om, binnen het kader van de onderwijs- en examenregeling, richtlijnen en aanwijzingen voor tentamens en examens vast te stellen.

Om invulling te kunnen geven aan hun inhoudelijke taak is het van belang dat de examencommissie niet alleen conform het toetsbeleid van de instelling handelt maar ook in de gelegenheid wordt gesteld actief mee te denken over de manier waarop het toetsbeleid wordt vormgegeven. Er zijn verschillende vormen denkbaar. De instelling kan de examencommissie raadplegen bij instellingsbreed toetsbeleid. Ook kan de examencommissie door de directeur worden betrokken bij de formulering van zaken betreffende toetsen en examens voor een bepaalde opleiding in de OER. De OER moet in ieder geval ter advisering aan de examencommissie worden voorgelegd.

Ad 3. Verhouding instellingsbestuur – examencommissie

De verhouding tussen instellingsbestuur en examencommissie verandert niet wezenlijk, maar er is wel sprake van een aantal aanscherpingen in de wet, die de positie van de examencommissie binnen de instelling verheldert en versterkt. De functionele onafhankelijkheid van de examencommissie ten opzichte van het instellingsbestuur betekent dat de examencommissie weliswaar wordt ingesteld door het instellingsbestuur, maar dat het instellingsbestuur er tevens voor zorgt dat de examencommissie binnen de instelling onafhankelijk haar werk kan verrichten. Dat betekent bijvoorbeeld dat het instellingsbestuur geen verplichtingen kan opleggen aan de examencommissie inzake de beoordeling van studenten. Het instellingsbestuur blijft eindverantwoordelijk voor de kwaliteit van het onderwijs en de graadverlening. De examencommissie moet handelen binnen de grenzen van de onderwijs- en examenregeling en stelt vast of de student het afsluitend examen van de opleiding heeft behaald. Aangezien de examencommissie daarmee vaststelt dat het eindniveau voldoende is om de graad te verlenen en bovendien tot taak heeft het getuigschrift uit te reiken, is het goed denkbaar dat de examencommissie de graad aan een student feitelijk verleent. Het instellingsbestuur kan deze verantwoordelijkheid tot het verlenen van een graad aan een student desgewenst mandateren aan de examencommissie.

Kanttekening

De examencommissie hoeft de kwaliteitsborging ten aanzien van toetsing niet geheel zelf te doen: er kan een toetscommissie worden ingesteld die deze taken geheel of gedeeltelijk uitvoert. Daarbij is van belang dat de examencommissie formeel verantwoordelijk blijft, ook al doet de toetscommissie het feitelijke werk. De examencommissie moet zich er daarbij van vergewissen dat de wijze waarop de toetscommissie

dit doet, voldoet aan de kwaliteitseisen, die door de examencommissie zijn vastgesteld. Daartoe moet de examencommissie zelf beschikken over expertise op het gebied van toetsing. Bovendien moet de examencommissie de toetscommissie kunnen aansturen als de werkwijze van de toetscommissie naar het oordeel van de examencommissie niet aan de kwaliteitseisen voldoet. Bij de accreditatie moet de examencommissie zich kunnen verantwoorden over het aspect ‘toetsen en beoordelen’.

De aangescherpte rol van de examencommissie én de behoefte van studenten aan verbetering van de toetskwaliteit, vereisen expertise op het gebied van toetsing en examinering. In de volgende hoofdstukken worden handvatten aangereikt om deze expertise te vergroten. Hierdoor krijgen toetsing en examinering de noodzakelijke kwaliteitsimpuls.

Voor de uitgebreide taakstelling van de examencommissies.
Zie 8 Geslaagd: Handreiking examencommissies HBO raad, april 2011.

2. Het brede perspectief van toetsen en beoordelen

Toetsen en beoordelen zijn basisconcepten in het onderwijs waar een hele wereld van ontwerpen, ontwikkelen en evalueren achter zit. Bij het ontwerpen van nieuwe curricula of herziening van bestaande curricula wordt toetsing en examinering nog wel eens ‘over het hoofd’ gezien. Geheel ten onrechte want de kwaliteit van toetsing en examinering is een van de belangrijkste onderdelen van de kwaliteit van het onderwijs.

Een kwalitatief goede opleiding heeft een evenwichtig samengesteld en weloverwogen toetsprogramma dat correspondeert met de eindkwalificaties. In het hoger beroepsonderwijs zijn deze eindkwalificaties vaak beschreven in de competentieprofielen. Voor het bereiken van de eindkwalificaties worden alle leerdoelen getoetst. De leerdoelen hebben betrekking op kennis, vaardigheden en waarneembaar gedrag en de integratie daarvan. Voor het niveau worden de Dublin descriptoren als uitgangspunt gebruikt. De Dublin descriptoren onderscheiden de volgende vijf facetten: kennis en inzicht, toepassen van kennis en inzicht, oordeelsvorming, communicatie en leervaardigheden. Ieder niveau, Associate Degree, Bachelor en Master kent een andere invulling van de Dublin descriptoren. Deze invulling moet in de eindkwalificaties en de leerdoelen van de desbetreffende opleiding herkenbaar terug te vinden zijn.

Om de kwaliteit van toetsing en examinering te borgen, zijn een aantal ontwerpeisen voorwaardelijk:

1. Consistentie visie op leren en toetsen.
2. Consistentie toetsprogramma en onderwijsprogramma.
3. Bevordering studiesucces.

2.1 Consistentie visie op leren en toetsen

De visie op leren is van invloed op de manier van opleiden, toetsen en beoordelen en daarmee op de manier waarop getoetst wordt. In het hoger beroepsonderwijs kan de visie op leren gekarakteriseerd worden als een competentiegerichte visie op leren.

Eigenschappen van de competentiegerichte visie op leren zijn volgens Cluitmans (2002):

1. Er is een sterke relatie met de beroepspraktijk.
2. Er wordt een krachtige leeromgeving ontwikkeld.
3. Persoonlijke ontwikkeling en competentiegroei van de student staan centraal.

Bij deze ontwikkelingsgerichte benadering op leren staan begeleiden en beoordelen in dienst van het leren en niet omgekeerd. Dit betekent dat toetsen en leren elkaar afwisselen, de student de meeste verantwoordelijkheid heeft in het leerproces en feedback geven en krijgen structureel onderdeel in het leerproces zijn. Zo worden toetsvormen als self/peer/co-assessment en (360 graden) feedback gehanteerd in het toets- en onderwijsprogramma om consistentie tussen visie op leren en toetsen te borgen.

Als het onderwijsprogramma vanuit deze visie op leren competentiegericht is maar het toetsprogramma onveranderd blijft, dan richten studenten zich op de onveranderende eisen van de toetsing en wordt het gewenste effect- het verwerven van competenties- niet bereikt. Toetsen en beoordelen zijn onderdeel van het onderwijsconcept en moet daarom ook optimaal passen bij het onderwijsconcept (Linn, 1990; Shepard, 2000).

2.2 Consistentie toetsprogramma en onderwijsprogramma

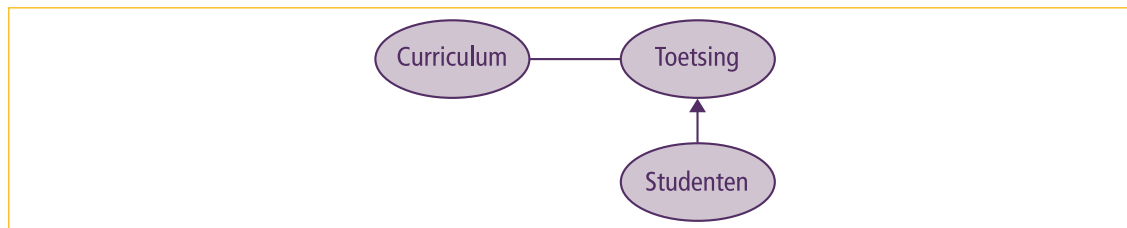
‘Wanneer het toetsprogramma [...] niet aansluit bij het onderwijsprogramma zal - alle idealen ten spijt - het toetsprogramma domineren. Er bestaat geen uitzondering op deze wetmatigheid en ontsnappen is niet mogelijk’ (Van der Vleuten, 1997).

Deze uitspraak geeft aan hoe dominant toetsing is voor het onderwijsprogramma en daarmee voor de opleiding en de student. De manier van toetsen dicteert wat en hoe er geleerd wordt. ‘Students learn what you inspect, not what you expect’². Bijna vijftien jaren na dato wint deze wetmatigheid nog meer aan kracht doordat toetsing en examinering bij accreditaties een groter gewicht krijgen en de versterkte rol van examencommissies in de Wet op het Hoger Onderwijs en Wetenschappelijk Onderzoek (WHW). Indien het toetsprogramma kwalitatief goed is geconstrueerd en consistent met de visie op leren, gecommuniceerde leerdoelen en met het uitgevoerde onderwijsprogramma dan horen goed presterende studenten te slagen. Het ontwerpen van een consistent toets- en onderwijsprogramma blijkt niet eenvoudig te zijn.

Tot de dag van vandaag worden leren en beoordelen als twee gescheiden werelden ervaren zowel door studenten als docenten. En zolang onderwijstijd en beoordelingstijd zwak zijn geïntegreerd, zullen studenten zich vooral inspinnen in de periodes vlak voor de beoordeling met het bekende uitstelgedrag als gevolg (Crooks, 1988).

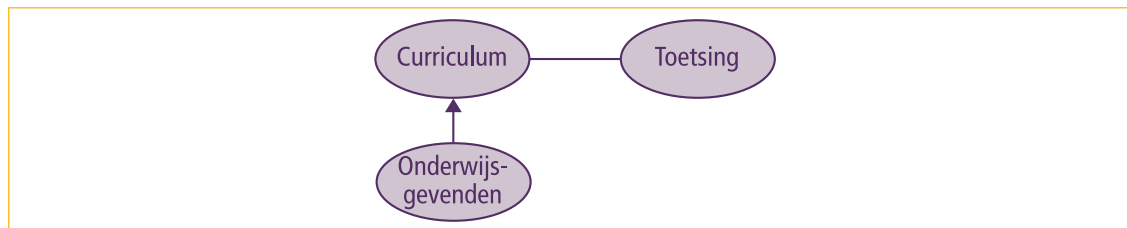
² Uitspraak Ian Hart op conferentie in Praag 1998

Uit onderzoek blijkt dat, wanneer toetsing en onderwijsinhoud niet op elkaar aansluiten, studenten zich laten leiden door de toetsing – alle inspanningen van onderwijsgegenden en onderwijsorganisatie ten spijt. Voor de student wordt succes gedefinieerd door de uitslag van de toets waarbij het behalen van het studiepoint een hogere prioriteit lijkt te hebben dan het behalen van een goed cijfer. Voor de student is toetsing het onderwijsprogramma geworden. Kort gezegd: als feiten worden beoordeeld, leren ze vooral feiten! (zie figuur 1).



Figuur 1 Studenten richten zich op het toetsprogramma

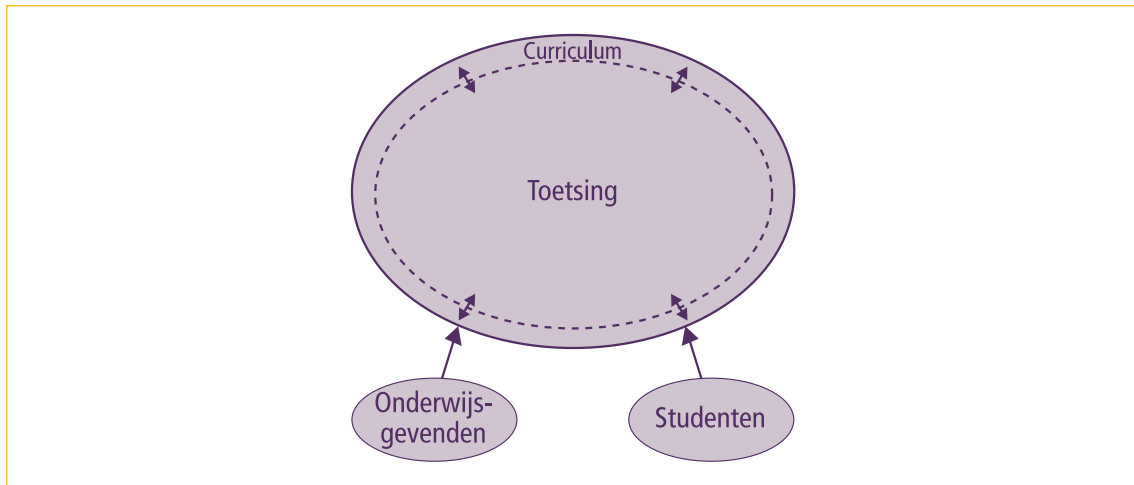
Voor onderwijsgegenden is deze constatering vaak moeilijk te accepteren. Onderwijsgegenden richten zich vooral op het onderwijsprogramma en houden zich minder bezig met toetsing. Zij ervaren toetsing eerder als last en zien toetsing als sluitstuk van het onderwijs (zie figuur 2).



Figuur 2 Onderwijsgegenden richten zich op het onderwijsprogramma

Bij onderwijsinnovaties en verbeteringen wordt het toetsprogramma met regelmaat als laatste aangepakt en soms zelfs helemaal niet. Daardoor moet gaandeweg het onderwijsprogramma drastisch worden bijgesteld omdat leren en toetsing niet in elkaars verlengde liggen en hierdoor fricties opleveren. Een curriculum dat niet aansluit bij een consistent toetssysteem misleidt zowel onderwijsgegenden als studenten: deze laatste studeren al naargelang het toetsprogramma is opgesteld. Hoe verleidelijk en goed het onderwijsprogramma is vormgegeven met didactische werkvormen zoals probleemgestuurd werken en projectonderwijs als het toetsprogramma daar niet op aansluit, worden de doelstellingen niet bereikt.

Onderwijsgegenden zijn voortdurend teleurgesteld omdat studenten niet op de beoogde wijze aan het onderwijs deelnemen. Presentieplicht, ingangseisen, selectie aan de poort, bindend studieadvies en veelvuldig toetsen worden in dat geval als zwakbodem geïntroduceerd om studenten te dwingen tot het volgen van het onderwijsprogramma. 'Deze mismatch tussen leren, toetsen en beoordelen, blijkend uit gebrekkige aansluiting tussen oefentaken ('lessen') en beoordelingen ('summatieve toetsen'), belemmeren een continu leerproces, het beklijven van kennis en vaardigheden en het ontwikkelen van een professionele identiteit' (Sluijsmans 2008).



Figuur 3 Toetsing als integraal onderdeel van het curriculum

In de ideale situatie maakt toetsing integraal deel uit van het curriculum en is er geen discrepantie tussen de leeroriëntaties van studenten en docenten. Het onderwijsprogramma en het toetsprogramma zijn twee zijden van dezelfde medaille (zie figuur 3).

2.3 Bevordering studiesucces

Het staken van de studie en/of het maken van een verkeerde studiekeuze heeft veel impact voor de student in kwestie, die daarmee motivatie, een deel van zijn studiefinanciering en kostbare studietijd verliest. Het studeergedrag van de student - en daarmee het studierendement en de doorstroomsnelheid – wordt het meest beïnvloed door:

1. Toetssystematiek.
2. Kwaliteit van de toetsen.
3. Examenregeling.

Ad 1 Toetssystematiek

De systematiek van toetsen bepaalt sterk het wel of niet slagen voor bepaalde vakken en daarmee de opleiding. Vragen als hoeveel toetsen worden afgenomen, in hoeverre is er compensatie mogelijk/wenselijk, op welke manier kunnen deelttoetsen worden ingezet in het onderwijs, zijn er ingangseisen of uitgangseisen, is er een bindend studieadvies, hoe is de programmering per onderwijsperiode, komen hierin aan bod.

Andere toetssystematieken, die studiesucces bevorderen en daarmee studierendement zijn volgens Cohen-Schotanus (2010):

1. Regelmatig toetsen: de student besteedt dan meer tijd aan de studie. Zie hoofdstuk 3, Functies van toetsing.
2. Niet meer dan één herkansing per toets: hoe meer herkansingen, hoe lager het rendement.
3. Geen snelle herkansing: snel leidt tot uitstelgedrag.
4. Geen concurrentie tussen toetsen: concurrentie tussen studieactiviteiten wordt hierdoor voorkomen, studieactiviteiten beter gespreid en studieresultaten worden gunstig beïnvloed.
5. Wel compensatie tussen toetsen waardoor studenten sneller slagen, ad 3: Examenregeling.
6. Geen absolute cesuur, omdat die geen rekening houdt met de – vaak toevallige- moeilijkheid van de toets. Zie voorbeeld bladzijde 15: de invloed van de cesuur op het rendement.

'Kenmerkend is dat de(ze) meest effectieve rendementsverhogende maatregelen tegenintuïtief en niet duur zijn. Er zijn nogal wat beleidsmakers, die moeite hebben met deze uitkomsten. Dit komt omdat maatregelen moeilijk opgepakt worden als ze tegenintuïtief zijn ook al zijn de maatregelen bewezen effectief. Een voldoende op het ene vak compenseren met een hoger cijfer op een ander tentamen zorgt bijvoorbeeld dat studenten sneller slagen en levert op lange termijn geen kennisachterstand op. Maar docenten hebben die compensatie liever niet en blijven nogal eens hangen in hun intuïtie. Het is daarom voor de kwaliteit van het onderwijs van belang dat de komende jaren ook gewerkt wordt aan de grenzen van het intuïtieve weten'(bron: Rede Tegenintuïtief: Cohen-Schotanus, 2010).

Ad 2 Kwaliteit van de toetsen

Kwaliteit van de toetsen uit zich onder ander in het feit dat het onderwijsprogramma aansluit op het toetsprogramma wat betreft onderwerp en beheersingsniveau.

Ook studenten, die de leerdoelen wel beheersen zakken nogal eens voor een toets. Toetsen variëren in kwaliteit en leiden ten onrechte tot studievertraging bij studenten. De geringe kwaliteit komt onder andere tot uiting in te gedetailleerde tentamenvragen, een onduidelijke stofomschrijving, onnauwkeurige formuleringen of vragen die niet van toepassing zijn op de leerdoelen, onduidelijke beoordelingscriteria en een matige samenhang tussen onderwijsprogramma en toetsprogramma waardoor de student op het verkeerde been gezet wordt. Het ligt niet altijd aan de student wanneer toetsen slecht gemaakt worden! Er valt veel winst in studiesucces te behalen wanneer de kwaliteit van de toetsen goed is. Zie voor verdere uitwerking hoofdstuk 4.

Ad 3 Examenregeling

De examenregeling is van invloed op de toetsresultaten en in het bijzonder de keuze van de opleiding voor een conjunctieve of compensatorische regeling. Bij een conjunctieve regeling zijn studenten verplicht alle onderwijsonderdelen met voldoende af te sluiten. Deze regeling is vrij standaard in het Nederlandse Hoger Onderwijs. Bij een compensatorische regeling kunnen studenten de goede resultaten bij het ene onderdeel gebruiken als compensatie bij een als onvoldoende beoordeeld onderdeel. Rekveld en Starren (1994) geven een overzicht van argumenten voor en tegen beide examenregelingen. Zij concluderen aan de hand van enkele onderzoeken dat een compensatorische regeling leidt tot snellere studievoortgang. Cohen-Schotanus 1995 laat in haar onderzoek zien dat bij invoer van een compensatieregeling de gemiddelde studieduur met ongeveer drie maanden daalde. Haar onderzoek vond plaats bij de faculteit Geneeskunde te Groningen waar een compensatieregeling werd ingevoerd dat studenten één vijf konden compenseren met een zeven van een ander vak. In de grootschalige studie van Van den Berg (2002) wordt aangetoond dat een compensatorische regeling ook bij grote groepen studenten van verschillende opleidingen en universiteiten leidt tot een snellere studievoortgang.

Een reden om geen compensatorische examenregeling door te voeren is het vermeende idee dat het leidt tot minder kwaliteit van de afgestudeerden. Dit lijkt echter niet het geval te zijn volgens Cohen-Schotanus. Zij zocht uit hoe vaak en op welke wijze doctoraalstudenten gebruik zijn gaan maken van de compensatieregel. Er bleek geen standaard moeilijk vak te zijn dat vaak werd gecompenseerd. Er was ook geen standaard combinatie tussen een moeilijk en gemakkelijk vak maar veeleer sprake van compensatie van 'toevallige vijven'. Daarin ziet zij een aanwijzing dat de compensatieregeling niet ten koste gaat van de kwaliteit van de opleiding. (bron: www.studiesucces.ho)

De invloed van de cesuur op het rendement (overgenomen uit Cohen-Scotanus, 2010).

Om u een indruk te geven hoe de cesuur het rendement beïnvloedt, laat ik graag enkele historische gegevens uit Groningen en Maastricht zien. In Groningen hingen (en hangen) we aan een absolute norm: 60% van de vragen goed beantwoord, was de grens voor een voldoende (gecorrigeerd voor raadmogelijkheden). Bij 54 toetsen fluctueerde het percentage gezakte studenten tussen de 20 en 97%.

In Maastricht werd de slaag-zak grens bepaald door naar de groepsprestaties te kijken: Een te grote (negatieve) afwijking van het groepsgemiddelde leverde een onvoldoende op. Bij een dergelijk relatieve cesuur fluctueert het vereiste kennisniveau dat nodig is om voor de toets te slagen.

Het resultaat van deze twee methoden was dat in Groningen gemiddeld per toets 54% van de studenten zakte en in Maastricht 17%. Geen wonder dus dat de Groningers ruim een jaar langer over hun studie deden. De meesten van ons zullen denken dat studenten meer leren bij een strenge cesuur. Als dat zo is, dan moeten de Groningse studenten onderweg heel wat meer kennis hebben opgedaan dan hun collega's uit Maastricht. Maar dat is niet zo, want op landelijke voortgangstoetsen vinden we geen verschillen. Kortom, geen kenniswinst bij een absolute cesuur, maar wel grote doorstroomproblemen. Voor internationale toetsexperts is de door ons intuïtief zo geliefde absolute cesuur absoluut uit den boze, omdat geen rekening gehouden wordt met de –vaak toevallige- moeilijkheid van de toets. Een tentamen bestaat uit een set toetsvragen die worden getrokken uit een groot aantal mogelijke vragen. De ene keer worden meer moeilijk vragen gesteld over de leerstof dan de andere keer. De absolute cesuur houdt hier geen rekening mee en is daarom arbitrair en niet goed te verdedigen. Er zijn andere, beter verdedigbare cesuren die meestal gebruik maken van panels. Om dergelijke kosten te beperken adviseer ik voor lokale toetsen de cesuur van Wijnen, Hofstee en Cohen. De cesuur van mijn promotor Wijnen houdt rekening met de kwaliteit van de toets, de cesuur van mijn promotor Hofstee is de meest intelligente en de cesuur van Cohen de meest pragmatische. november 2010.

3. Functies van toetsing

Het besef dat toetsen en beoordelen sturend is voor zowel het leerproces en de ontwikkeling van de studenten als voor de optimalisatie van de kwaliteit van het onderwijs is een belangrijke ontwikkeling vanuit de jaren negentig (zie bijvoorbeeld Boud, 2000; Longhurst & Norton, 1997; Sluijsmans, 2008; Van der Vleuten & Schuwirth, 2005). Slagen voor de totale verzameling toetsen van een opleiding moet betekenen dat gediplomeerden terecht bevoegd zijn om de aan het getuigschrift verbonden titels te dragen en de rechten uit te oefenen die aan die titels zijn verbonden: het zogenaamde civiel effect van de toetsing: de feedback functie naar het formele registratiesysteem. Vanwege de grote waarde die toetsen maatschappelijk gezien hebben, vervullen ze in een onderwijsprogramma een belangrijke rol met een uitstralend effect naar het studeergedrag van studenten. Voor docenten zijn toetsen van groot belang om te beoordelen of het programma ‘gewerkt’ heeft en of de leerdoelen voor het getoetste onderdeel realistisch en duidelijk genoeg geformuleerd zijn. Voor studenten heeft het slagen voor een toets de hoogste prioriteit. Naast het feedback geven aan het formele registratiesysteem heeft toetsing een ontwikkelingsrichte kant: de feedback op het leren van de student. Als zodanig zijn er twee functies van toetsing te onderscheiden:

1. *Feedback geven op het leren, de zogenaamde formatieve toetsing.*
Deze feedback is ter ondersteuning van de inhoudelijke ontwikkeling van de student, gericht op het opstellen en bijstellen van persoonlijke leer- of ontwikkelplannen, op het verbeteren van prestaties en versterken van interne motivatie. **Kernvraag is: wat kan de student al en wat nog niet?**
2. *Feedback geven naar het formele registratiesysteem, de zogenaamde summatieve toetsing.*
Deze feedback is ter ondersteuning van formele beslissingen die de start, voortgang en afsluiting van het leren bepalen. Leerprestaties die met ‘voldoende’ zijn beoordeeld, worden vastgelegd in certificaten en getuigschriften. **Kernvraag is: Kan de student het ja of neen?**

De begrippen formatieve en summatieve toetsing verwijzen naar de doelstelling van de toetsing en niet naar het instrumentarium, dat bij toetsing gebruikt wordt. Zo heeft toetsing in een competentiegerichte leeromgeving naast de summatieve functie een belangrijke ondersteunende formatieve functie bij het leren. Dit betekent dat studenten tijdens het leerproces fouten mogen maken en daarop niet in formele zin afgerekend worden. Formatieve toetsing is meer dan het voorbereiden van studenten middels oefentoetsen op de summatieve toetsing. Formatieve toetsing is de continue feedback, die de student krijgt tijdens het leerproces. Wel is het wenselijk dat de formatieve en de summatieve toetsing in elkaars verlengde liggen en op elkaar aansluiten (Shavelson e.a., 2008).

De laatste jaren is er veel onderzoek gedaan naar de werking van feedback. In het grootschalige literatuuronderzoek van Hattie en Timperley³ concluderen zij dat feedback effectief kan zijn om het

³ The Power of Feedback, John Hattie and Helen Timperley Review of Educational Research March 2007, Vol. 77, No. 1, pp. 81–112

leerproces te stimuleren mits er kwalitatief goede vorm en inhoud aan gegeven is. Effectieve feedback is alleen mogelijk als er

- duidelijke, gecommuniceerde doelen en succescriteria in de vorm van concrete gewenste resultaten van het leerproces, zijn geformuleerd;
- specifieke informatie aanwezig is over de voortgang van het leerproces;
- vervolgstappen voor het verdere leerproces gemaakt worden.

Er worden drie dimensies van feedback onderscheiden:

1. Richtinggevende dimensie: Feed up. Waar ga ik naar toe? Wat is het doel, wat is het resultaat waar we naar streven?
2. Oordelende dimensie: Feed back. Hoe doe ik het tot nu toe? Welke voortgang heb ik gemaakt naar dat resultaat?
3. Activerende dimensie: Feed forward. Wat kan ik verder doen? Welke activiteiten moet ik ondernemen om dichterbij het resultaat te komen?

Ad 1 Feed up: Waar ga ik naar toe?

Een kritisch aspect van feedback is de informatie die de student krijgt over leerdoelen, die leiden tot resultaten die beoordeeld worden. Dit wordt feed up genoemd: de informatie omtrent de vraag, waar ga ik naar toe?

Goed omschreven leerdoelen kunnen het leerproces positief beïnvloeden. De student kan zich meer inspannen en beter samenwerken met anderen omdat er een duidelijke richting is aangegeven.

Aan doelen zijn twee belangrijke dimensies verbonden:

1. uitdaging;
2. betrokkenheid.

Black and Wiliam (1998) concludeerden dat 'het geven van uitdagende opdrachten en uitgebreide feedback leidt tot grotere betrokkenheid van studenten en hogere prestaties.'

Zo'n betrokkenheid ontstaat niet vanzelf. De student als actieve participant dient betrokken te worden bij het formuleren van leerdoelen en het maken van resultaatafspraken.

Ad 2 Feed back: Hoe doe ik het tot nu toe?

De feed back dimensie is de informatie, die de student krijgt over de gemaakte progressie: hoe doe ik het tot nu toe? Feed back is informatie over de vergelijking tussen de prestatie –of het gedrag- van de student en het te bereiken doel. De informatie moet **gekoppeld zijn aan een standaard (het leerdoel) of eerder gedrag. De feed back is het meest effectief als deze direct aansluitend op het leergedrag of leerproces**

gegeven wordt. Maar al te vaak wordt de feed back gereduceerd tot het geven van een beoordeling in de vorm van een cijfer. Deze vorm van feed back geeft maar weinig informatie ten aanzien van de gemaakte vorderingen van de student.

Ad 3 Feed forward: Wat kan ik verder doen?

De kracht van feedback wordt vele malen sterker als de student informatie krijgt over te ondernemen acties, die de student moet uitvoeren om dichterbij het leerdoel te komen of nieuwe doelen te stellen. Voor de student is het van belang te weten hoe verder te gaan. Dit wordt feed forward of constructieve feedback genoemd.

Feedback (feed up, feed back, feed forward) kan zich richten op vier niveaus:

1. Taakniveau – hoe goed wordt de taak begrepen en uitgevoerd? Wat doe ik?
2. Procesniveau – welk proces is nodig om de taak te begrijpen en uit te voeren? Hoe doe ik het?
3. Zelfregulerend niveau – hoe worden de activiteiten gemonitord en (zelf) gereguleerd? Hoe stuur ik mijzelf?
4. Persoonlijk niveau – wat betekenen persoonlijke evaluaties voor mij? Wat doet het met mij?

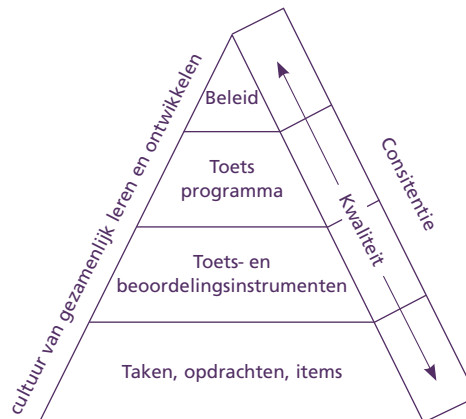
Hattie en Timperley concluderen dat feedback zich moet richten op het niveau van de taak, het proces en de zelfregulatie. Feedback op persoonlijk niveau in leerprocessen is minder effectief om leren te bevorderen en dient vermeden te worden.

De constatering van Hattie en Timperley hebben grote implicaties voor de manier van toetsen en beoordelen. Te vaak wordt toetsing eenzijdig gebruikt als ‘assessment of learning’ ten behoeve van het formele registratiesysteem en externe legitimering. Het is nu net de formatieve feedback, ‘assessment on learning’, die van belang is voor het leren van de student. Uitgaande van de inzichten over de kracht van feedback tijdens het leren dient een balans gezocht te worden tussen enerzijds ‘bestaande en nieuwe waarden’ van toetsing en anderzijds ontwerpeisen, die aan toetsen en beoordelen gesteld worden. Zo dient er meer samenhang tussen summatieve toetsing en formatieve toetsing gezocht te worden waardoor onderwijstijd en beoordelingstijd sterker geïntegreerd worden en leerrendement verhoogd wordt.

4. Kwaliteit van toetsing en examinering

In hoofdstuk twee is beschreven welke ontwerpeisen ten grondslag liggen aan toetsen en beoordelen. Om kwaliteit van toetsing en examinering te borgen zodat opleidingen kwalitatief hoogwaardige getuigschriften verstrekken, positief geaccrediteerd worden en studenten succesvol zijn, is het van belang dat toetsing en examinering tot stand komen vanuit *centrale(organisatie)sturing*. De examencommissie, die belast wordt met het uitvoeren van de kwaliteitsborging van toetsing en examinering, vormt hierin een vitaal element. Vanuit centrale organisatiesturing werken medewerkers samen aan toetsontwerp, toetsontwikkeling, toetscorrectie, toetsanalyse en kwaliteitsborging volgens een gemeenschappelijk toetsbeleidskader, dat organisatiebreed is gecommuniceerd en geaccepteerd. Studenten en werkveldpartners hebben expliciet een rol. Een en ander vraagt kennis en kunde ten aanzien van toetsen en beoordelen.

Om het bewustzijn te stimuleren en de bekwaamheid te bevorderen ten aanzien van toetsen en beoordelen is het noodzakelijk om te beschrijven over welke kennis en kunde het dan gaat. De (aangepaste) kwaliteitspiramide voor eigentijds toetsen en beoordelen van Joosten-ten Brinke & Sluijsmans 2010, biedt de mogelijkheid om op een gestructureerde manier een kennisbasis te beschrijven voor toetsen en beoordelen en van daaruit te bekijken welke kwaliteitsverhogende interventies noodzakelijk zijn om kwaliteit van toetsing en examinering te borgen.



Figuur 4 De kwaliteitspiramide voor toetsen en beoordelen (Joosten-ten Brinke & Sluijsmans, 2010, aanpassing, Jaspers, 2011)

Toelichting

De kennis en kunde ten aanzien van toetsen en beoordelen kan verdeeld worden over vier niveaus waarop de kwaliteit van toetsing beoordeeld en geborgd kan worden. Op het hoogste niveau gaat het om de kwaliteit van het beleid ten aanzien van toetsing en examinering. Wat is de visie van een opleiding ten aanzien van toetsen en beoordelen? Wat zijn de uitgangspunten en kaders? Hoe wordt de kwaliteit van de toetsing geborgd? Welke rol hebben docenten, studenten en werkveldpartners in de toetsing? In hoeverre draagt toetsing bij aan duurzame leerresultaten en het vormen van een blijvende lerende houding bij studenten? Op het tweede niveau gaat het om de kwaliteit van toetsprogramma's. Daarbij kan men denken aan een hele opleiding, of een afgebakend geheel van onderwijs of procedures en protocollen voor het erkennen van verworven competenties. Is het gehele programma in balans voor het meten van de competenties van de student? Het derde niveau betreft de kwaliteit van afzonderlijke toets- en beoordelingsinstrumenten. Op het laagste niveau gaat het om de kwaliteit van taken, opdrachten en/of items. Hieronder wordt elk onderdeel van een toets verstaan waarop een student een reactie kan geven. De niveaus staan niet op zichzelf. Kwaliteit van toetsing wordt geborgd door een sterke samenhang van de vier niveaus (aangegeven door de zijdelingse pijl). Is de kwaliteit op het laagste niveau niet goed, dan heeft dit negatieve invloed op de kwaliteit van de bovenliggende niveaus. Andersom, idem dito. Als onduidelijk is wat de visie op toetsing is en deze niet consistent doorvertaald is naar de lagere niveaus, heeft dit zijn uitwerking naar beneden op de kwaliteit van toetsprogramma's, toetsen en de onderliggende opdrachten, taken en items. Consistentie tussen de niveaus leidt tot kwaliteitsverhoging (Joosten-ten Brinke 2011). Daarnaast wordt kwaliteit van toetsing cruciaal beïnvloed door de cultuur van de organisatie waarin ruimte moet zijn voor het leren en ontwikkelen op individueel/team en organisatieniveau op het gebied van toetsing en examinering. De verschillende niveaus in de kwaliteitspiramide en de cultuur van gezamenlijk leren en ontwikkelen, die randvoorwaardelijk is, worden hieronder toegelicht.

4.1 Beleid

Op het hoogste niveau (aanpassing kwaliteitspiramide Jaspers, 2011) wordt gesproken over beleid ten aanzien van toetsing en examinering. Het formuleren van toetsbeleid is een circulair proces waarbij op systematische en planmatige wijze wordt nagedacht over de toetsing die een opleiding nodig heeft of wenselijk acht om de competenties van de student te beoordelen. In een toetsbeleidsplan worden de uitgangspunten en kaders met betrekking tot ontwerpen, ontwikkelen, afname, rapportage en evaluatie van toetsing beschreven en vastgesteld. Wat wil een opleiding met toetsing bereiken, op welke manier wordt de gekozen onderwijsvisie het beste uitgewerkt? Hoe wordt ervoor gezorgd dat nieuwe, minder ervaren docenten op hetzelfde niveau toetsen als de ervaren en toetskundige docenten, hoe vindt professionalisering van de examencommissie plaats? Wat zijn de gevolgen van de toetsing voor de onderwijsorganisatie? Hoe wordt kwaliteit van toetsing geborgd en wat zijn de financiële kaders?

Het toetsbeleid heeft betrekking op drie beleidsniveaus:

1. Strategisch beleidsniveau: waarom en wat?
2. Tactisch beleidsniveau: hoe?
3. Operationeel beleidsniveau: wie, wat, wanneer en waar?

Idealiter wordt begonnen met het formuleren van strategisch beleid omdat dit beleid kaderstellend is voor de andere twee niveaus. Wanneer op strategisch niveau bijvoorbeeld gekozen wordt voor digitalisering van de toetsen moet dit vertaald worden op het tactische en het operationele niveau. Op tactisch niveau wordt aangegeven hoe de samenhang is met andere toetsvormen, hoe consistentie in het curriculum bevorderd kan worden en hoe een itembank opgebouwd wordt. Op operationeel niveau wordt aangegeven wie de toets maakt, wie de items schrijft, wanneer dat moet gebeuren en wat dat betekent voor de onderwijsorganisatie ten aanzien van beheer en onderhoud.

Beleidsontwikkeling kent geen beginpunt en eindpunt. Urgente problemen op het ene beleidsniveau moeten immers aangepakt worden zonder dat er eerst een uitgebreide analyse op het andere niveau wordt uitgevoerd. Voor alle niveaus geldt dat er structureel geëvalueerd en gereflecteerd wordt op het geformuleerde toetsbeleid opdat het beleid getoetst en geoptimaliseerd kan worden.

Het bovenstaande kan geïnterpreteerd worden alsof de opleiding geheel vrij is in haar formulering van toetsbeleid. Niets is echter minder waar; opleidingen worden in toenemende mate geconfronteerd met externe factoren, die van invloed zijn op de keuze uit methoden, mensen en middelen zoals

- Wettelijke kaders vanuit de regels van de Wet Hoger onderwijs en Wetenschappelijk Onderzoek (WHW).
- Eisen vanuit accreditatie.
- Generieke kwaliteitsdoelstellingen ten aanzien van de opleiding: de Dublin descriptorën.
- Landelijke afspraken tussen opleidingen met betrekking tot nieuwe opleidingsprofielen.
- Vernieuwende activiteiten en inzichten op het gebied van leven lang leren, vraagstukken, elders verworven competenties.
-

Daarnaast zijn er interne factoren te benoemen zoals kwaliteitseisen vanuit de hogeschool, de visie op leren, de gekozen organisatie van toetsing, financiële ruimte en personele middelen.

4.2 Kwaliteit van het toetsprogramma

Op een tweede niveau wordt gesproken over de kwaliteit van het toetsprogramma. Een toetsprogramma is het geheel van toetsen van bijvoorbeeld een volledige bachelor- of masteropleiding. Ook een EVC procedure kan gezien worden als een programma dat moet voldoen aan de kwaliteitscriteria voor toetsprogramma's (Joosten-ten Brinke, 2008).

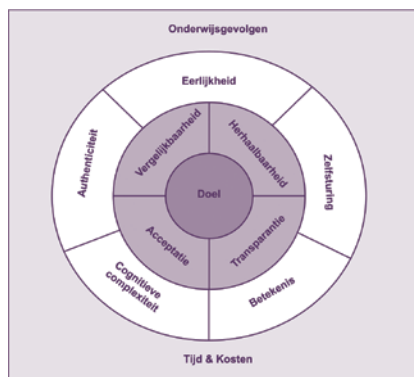
De kwaliteit van een toetsprogramma kan o.a. geborgd⁴ worden door:

1. Competentie Assessment Programma (CAP);
2. Methodemix.

4.2.1 Competentie Assessment Programma (CAP)

In het onderwijs groeit het besef dat leren en toetsen op elkaar moet aansluiten omdat de kwaliteit van toetsing voor een groot deel de kwaliteit van het onderwijs representeert (zie o.a. Schuwirth & Van der Vleuten, 2006). In die discussie is men het erover eens dat toetsen meer is dan het afsluiten van een cursus of module. De samenhang tussen toetsen in een opleiding (zowel formatief als summatief) en de keuzes voor toetsvormen moeten in balans zijn. Deze gedachte heeft ertoe geleid dat er niet meer alleen gekeken wordt naar de psychometrische definitie van kwaliteit in termen van betrouwbaarheid en validiteit maar dat er gesproken wordt over edumetrische kwaliteit, waarin ook de consequenties van toetsing op het leren mee beoordeeld worden.

Om de kwaliteit van het toetsprogramma van een opleiding te screenen middels zelfevaluatie, hebben Baartman, Prins, Kirschner en Van der Vleuten (2007) twaalf kwaliteitscriteria benoemd. De grondgedachte achter de kwaliteitscriteria is dat er een combinatie gemaakt wordt van klassieke ideeën over kwaliteit van toetsing en nieuwe opvattingen over wat goede assessments zijn in competentiegericht onderwijs. Hierdoor worden kennis over en ervaring met het werken met klassieke beoordelingsmethoden gecombineerd met kennis en ervaring met het werken met nieuwe beoordelingsmethoden van assessments.



Figuur 5 Wiel Baartman

Toelichting figuur 5

De kwaliteitscriteria zijn in een wiel in een aantal lagen opgedeeld. In de kern van het wiel staat het kwaliteitscriterium geschiktheid voor onderwijsdoelen. Dit is het basiscriterium voor toetsing, dat bepaalt dat de gebruikte toetsvormen moeten passen bij het onderwijsprogramma en bijvoorbeeld in het geval van competentiegericht onderwijs gericht moeten zijn op competenties. In de middelste ring staan de kwaliteitscriteria, vergelijkbaarheid, herhaalbaarheid, acceptatie en transparantie. Deze criteria vormen de basisprincipes voor het toetsprogramma, die vaak voorwaardelijk zijn voor de criteria in de buitenste ring. Zo moet het toetsprogramma eerst doorzichtig en begrijpelijk zijn voor de onderwijsorganisatie (transparantie), voordat

⁴ Er zijn veel borgingsmechanismen mogelijk (zie ook hoofdstuk 5); in deze publicatie worden twee mechanismen uitgebreid besproken.

gekeken wordt naar de betekenis ervan voor studenten. De buitenste ring bestaat uit de kwaliteitscriteria eerlijkheid, zelfsturing, betekenis, cognitieve complexiteit en authenticiteit. Deze kwaliteitscriteria komen vaak voort uit nieuwe inzichten over wat goede manieren van leren en beoordelen zijn. Buiten het wiel zien we tenslotte de ruimte waarin de beoordeling zich afspeelt.

Hier zijn de kwaliteitscriteria Tijd, Kosten en Onderwijsgevolgen te vinden (hoogste niveau kwaliteitspiramide), die beiden betrekking hebben op de samenhang tussen de toetsen en de rest van het onderwijsprogramma. De kwaliteitscriteria geven een concrete uitwerking aan het begrip validiteit. Het begrip betrouwbaarheid heeft een nieuwe invulling gekregen in met name de kwaliteitscriteria vergelijkbaarheid en herhaalbaarheid van beslissingen. Niet de cijfermatige aantoonbare betrouwbaarheid bepaalt de kwaliteit van de toetsing maar de door de betrokkenen als redelijk ervaren beoordelingsresultaten zijn doorslaggevend (Cluitmans & Klarus, 2005).

Voor een uitgebreidere beschrijving van de twaalf kwaliteitscriteria en indicatoren Competentie Assessment Programma (CAP), zie bijlage 2.

4.2.2 Methodemix

Het hanteren van een methodemix is een (andere) manier van toetsprogrammering om sterktes en zwaktes van de toetsen in betrouwbaarheid en validiteit op te vangen. 'Als zodanig bestaat er geen 'beste methode' van toetsen omdat de diverse methoden de zwakste schakel niet allemaal op dezelfde plaats hebben' (Straetmans, 2001). De methodemix heeft bovendien als voordeel dat er meer informatie wordt verzameld. Veel informatie verzamelen is erg belangrijk om een uitspraak te doen over het (competentie) niveau van de student.

Een indeling in drie soorten van toetsvormen is volgens Straetmans (2001) -hands-on, simulatie en hands-off.

1. Hands-on staat voor toetsvormen waarbij de student in een realistische of zelfs reële werksituatie eventueel met gebruik van echte apparatuur, gereedschappen of instrumenten, kenmerkende taken uitvoert. Het praktisch deel van het rijexamen is een goed voorbeeld van een hands-on toets of het geven van een proefles in een EVC-procedure voor zij-instromers in het lerarenberoep.
2. Bij hands-off toetsvormen worden taken voorgelegd waaruit moet blijken of studenten de cognitieve component van een competentie beheersen. Dit kan op papier maar er wordt steeds vaker gebruik gemaakt van multimedia om taaksituaties zo realistisch mogelijk te presenteren.

3. Onder simulatie worden al die vormen verstaan waarbij de student de bekwaamheid demonstreert onder nagebootste werkomstandigheden met gebruikmaking van al dan niet realistische apparatuur. Dit soort vormen wordt in de Angelsaksische literatuur aangeduid met de term *work sample test*. De uit te voeren taken zijn te beschouwen als een steekproef uit de taken die de kandidaat in de reële werksituatie zou tegenkomen.

	Hands-on	simulatie	Hands-off
Accuraatheid toetsscore	-	0	+
Generaliseerbaarheid toetsscore	-	0	+
Overdraagbaarheid toetsscore (transfer)	+	0	-

Figuur 6 Sterktes en zwaktes van hands-on, simulatie en hands-off (overgenomen uit Straetmans, 2001).

Er is een tendens om terug te keren naar de vertrouwde gestandaardiseerde kennistoetsen. Dit komt zeker door de polarisatie in de discussie over het ‘oude’ en het ‘nieuwe’ leren. Daarnaast bezitten docenten niet altijd de deskundigheid om op andere manieren te toetsen. Zo worden klassieke kennistoetsen vaak gezien als de enige manier om goed te toetsen omdat de objectiviteit van de beoordeling hoog is. Deze vooronderstelling klopt niet. De objectiviteit van de beoordeling is inderdaad hoger dan bij andere manieren van toetsen; de validiteit (meet de toets wat hij beoogt te meten?) en de betrouwbaarheid (is het resultaat op een ander moment van toetsing hetzelfde?) zijn bij klassieke kennistoetsen echter laag. Deskundigheidsbevordering op het gebied van toetsen is van cruciaal belang. Zo leert men vooronderstellingen te toetsen op basis van nieuwe theoretische inzichten en daarmee intuïtie aan evidentie.

Een voorbeeld van een onjuiste vooronderstelling over openboek toetsen

Veel docenten denken dat openboek toetsen te gemakkelijk zijn voor studenten terwijl uit onderzoek blijkt dat studenten er op de lange termijn juist meer van leren. Heijne-Penninga (2010) vergeleek ‘openboek toetsen en geslotenboek toetsen in een groot onderzoek onder Groningse studenten Geneeskunde. Ze ontdekte dat openboek toetsen de studenten uiteindelijk meer grip geven op de stof omdat ze hen dwingen om zelf verbanden te leggen. Als iemand later iets opdiept uit zijn geheugen komen al die verbanden naar boven en niet alleen dat ene losse feitje. Openboek toetsen maken kennisverwerving en -gebruik dus efficiënter. Het effect van openboek toetsen geldt wel alleen voor de lange termijn. Heijne-Penninga laat zien dat studenten in de eerste drie jaar van hun studie juist beter presteren met ‘stampen’ voor gesloten boektoetsen. Dat komt omdat ze dan het grote plaatje nog niet zien. Pas later valt alles op zijn plek.

Openboek toetsen zetten studenten juist aan tot ‘diep leren’. Ze krijgen inzicht in de stof door samenvattingen te maken en andere literatuur te raadplegen. Een andere uitkomst van het onderzoek van Heijne-Penninga is dat openboek toetsen even moeilijk zijn als gesloten boek toetsen. Ook is de kwaliteit vergelijkbaar.

4.3 Kwaliteit van toets- en beoordelingsinstrumenten

Op het derde niveau wordt door Joosten-ten Brinke (2011) gesproken over de kwaliteit van de toets- en beoordelingsinstrumenten. Met toets- en beoordelingsinstrumenten worden alle mogelijke toetsvormen bedoeld. Beoordelingsmodellen bij schriftelijke toetsen bestaan uit het juiste antwoord inclusief een verwijzing naar de kennisbasis en de leerstof. Deze beoordelingsmodellen worden meestal gelijk bij de constructie van de toets ontwikkeld. Bij het beoordelen van portfolio's is dit in de praktijk veel minder het geval.

Elke toets moet voldoen aan de criteria van bruikbaarheid, validiteit en betrouwbaarheid. Dit betekent dat toetsen efficiënt en eerlijk zijn, dat ze meten wat ze beogen te meten en dat er overeenstemming is tussen de docenten over de beoordeling. Om te voldoen aan deze criteria moet bij toetsconstructie de volgende vragen beantwoord worden: Wat wordt er getoetst? Bij wie wordt de toets afgenomen? Wie zijn de beoordelaars (docenten, studenten en/of werkveldpartners)? Hoeveel toetstijd is er beschikbaar? Hoe wordt er getoetst? Hoe wordt er beoordeeld?

De validiteit van toetsen kan onder andere verhoogd worden door te werken met een toetsmatrijs. In zo'n toetsmatrijs wordt aan de hand van twee variabelen (de inhoud en het beheersingsniveau) de basis vastgelegd voor toetsing en de te ontwikkelen vragen. Hiermee kunnen schrijfopdrachten gegeven worden aan itemconstructeurs en is het mogelijk om meerdere jaren achter elkaar inhoudelijk vergelijkbare toetsen te ontwikkelen. De betrouwbaarheid van toetsen kan zeker verhoogd worden door gezamenlijke toetsconstructie. Het gezamenlijk ontwikkelen en bespreken van beoordelingscriteria bevordert de consistentie van beoordelingen. Het ontwikkelen van toetsen is niet alleen het schrijven van toetsopdrachten of items, maar betekent ook dat antwoordmodellen en correctievoorschriften vooraf beschikbaar zijn waarbij de beoordelaars het vooraf met elkaar eens zijn over het juiste antwoord en de bijbehorende scoringsregels.

Verschillende toetsvormen in een competentiegericht leeromgeving worden in hoofdstuk 6 besproken.

4.4 Kwaliteit van taken, opdrachten of items

Tot slot geeft Joosten-ten Brinke (2011) aan dat het laagste niveau van de kwaliteitspiramide de kwaliteit van de taken, de opdrachten en de items betreft. Elke taak, opdracht of vraag die in een toets gebruikt wordt, of dit nu een meerkeuzevraag, een open vraag, of een opdracht tot het schrijven van een werkstuk is, moet voldoen aan criteria van relevantie, objectiviteit, efficiëntie en geschikte moeilijkheidsgraad. De relevantie van een taak, opdracht of item heeft betrekking op de mate waarin deze toetst wat het beoogd te toetsen met andere woorden op de mate waarin de taak, de opdracht of het item een directe relatie heeft met het leerdoel. Voor toets taken in competentiegericht onderwijs betekent dit dat toets taken opgebouwd zijn rond realistische, betekenisvolle situaties. Een taak, opdracht of item is objectief indien de score onafhankelijk is van degene die de beoordeling uitvoert. Dit impliceert dat antwoordmodellen zo min mogelijk ruimte voor subjectieve interpretatie openlaten en dat deskundigen het eens zijn over het juiste antwoord of over de criteria waaraan een juist antwoord moet voldoen. Het criterium van efficiëntie

betekent dat, wanneer men bij het ontwikkelen van een taak, opdracht of item kan kiezen uit verschillende soorten taken, opdrachten of items, die in alle opzichten gelijkwaardig zijn, de taak, de opdracht of het item dat het minste tijd kost in termen van afname en correctie, de voorkeur heeft. Tot slot moet de moeilijkheidsgraad van een taak, opdracht of item aanvaardbaar zijn, gegeven het niveau van de studenten. Digitale toetsing vraagt op dit niveau van de kwaliteitspiramide specifieke aanvullende kennis en kunde.

4.5 Digitale toetsing

Er zijn vele redenen op velerlei gebied te geven waarom een opleiding overstapt op digitalisering van de toetsen. Allereerst zijn er winstpunten te noemen op het gebied van de zogenoemde ‘grote’ kwaliteit: een analyse uit de computer geeft oneindig meer informatie dan een handmatig verwerkte scoring, zowel voor de student als voor de opleiding. Een eenmaal opgebouwde itembank maakt de opleiding minder kwetsbaar ten aanzien van toetsbeheer. Het eigen kapitaal wordt gebruikt en is toegankelijk voor iedere belanghebbende. Een beheerorganisatie van gedigitaliseerde toetsen is op den duur minder kostbaar dan ‘ouderwetse’ papieren toetsen vooral qua docententijd!.

Een itembank vraagt om een structuur: dit dwingt de opleiding om de onderwijsvisie concreet te vertalen naar dimensies binnen toetsen en items. Een speciale vorm van toetsen, de voortgangstoets, is zelfs uitsluitend te verwerken binnen een geautomatiseerd systeem. Wanneer er eenmaal een digitale itembank bestaat voor papieren toetsen is het overschakelen naar een elektronische (tijd- en plaatsafhankelijke) leeromgeving, met daarin toetsen, gemakkelijker en vanzelfsprekender.

Maar het bezitten van een geautomatiseerd toetssysteem levert ook op praktisch gebied kwaliteit op: de winst in de snelheid van toetsaanmaak, het verwerken tot scores en de opname van de scores in het studentvolgsysteem.

Het gebruik van toetsen met gesloten vragen vraagt in ieder geval om een geautomatiseerde omgeving. Handmatig aanmaken en verwerken van dergelijke toetsen is zoiets als de kaartenbak gebruiken in plaats van een database, of het telraam gebruiken in plaats van de rekenmachine. De voordelen zijn legio. De nadelen zijn meteen de schaduwzijden van de voordelen: het is in de startfase kostbaar en inspannend. Onderhoud blijft geld kosten en aandacht vragen en de organisatie moet er rijp voor zijn. Anders is het een breekijzer in de diverse samenwerkingsverbanden binnen de organisatie. De intercollegiale relatie kan onder druk komen te staan.

In de wereld van geautomatiseerde papieren toetsen is zelfsturing moeilijk te realiseren vanwege de complexe logistiek van papieren toetsaanmaak en toetsverwerking. Beeldschermtoetsing realiseert – naast een snellere aanbieding en verwerking – veel nieuwe mogelijkheden zoals multimediale toepassingen, flexibilisering in tijd, plaats en niveau (adaptief toetsen). ‘Zelfsturing door de student’ als organisatieprincipe wordt eenvoudiger wanneer deze ondersteund wordt vanuit een elektronische leeromgeving. Leren en toetsen worden op deze wijze steeds meer met elkaar geïntegreerde processen. Bij de overgang naar het geautomatiseerd verwerken van papieren toetsen kan een checklist gebruikt worden. Zie bijlage 3, Checklist Geautomatiseerd toetsen.

4.6 Cultuur van gezamenlijk leren en ontwikkelen

In een goede toetscultuur is het vanzelfsprekend dat de betrokken partijen, instelling, docenten en studenten zich maximaal inspannen voor een goed toetsresultaat. Concreet dienen betrokken partijen bij de evaluatie van bijvoorbeeld toetsresultaten zich automatisch de volgende vragen te stellen:

1. Was de inhoud van de toets een juiste representatie van de gecommuniceerde leerdoelen?
2. Zijn de toetsinhoud, de gecommuniceerde leerdoelen en het uitgevoerde onderwijsprogramma consistent?
3. Hebben alle studenten goed gewerkt voor deze toets?

maar ook:

- Hoe kan ik mij professionaliseren in nieuwe vormen van toetsen?
- Hoe ziet het toetsprogramma eruit (nieuwsgierigheid)?
- Hoe kunnen we voortdurend verbeteren en vernieuwen?
- Wat kunnen we van elkaar leren?
- Hoe kunnen we professionaliseren?

De antwoorden op deze vragen geven docenten, opleidingsteams en opleidingsorganisatie belangrijke informatie over de kwaliteit van de toetsen, het toetsprogramma en het toetsbeleid waardoor het toets- en onderwijsprogramma verbeterd kan worden.

Kanttekening

Consistentie tussen de visie op leren en toetsen is niet exclusief voorbehouden aan het toetsprogramma. In de hogeschoolorganisatie dient zich een parallel leerproces te voltrekken.

Als zodanig moet er eenheid in architectuur zijn tussen de visie op leren en toetsen van studenten en de visie op leren en toetsen van docenten. Hierdoor ontstaat er een congruent leerklimaat waarin docenten en studenten van en met elkaar leren en zich gezamenlijk kunnen ontwikkelen tot een lerende, elkaar toetsende gemeenschap.

5. Kwaliteitssysteem toetsing en examinering

In hoofdstuk vier is beschreven dat de kwaliteit van toetsing geborgd wordt op vier verschillende, met elkaar samenhangende niveaus in een cultuur waarin het gezamenlijk leren en ontwikkelen centraal staat. Om de gewenste kwaliteit op elke niveau afzonderlijk en in samenhang met elkaar te (blijven) realiseren is een kwaliteitssysteem nodig. Een definitie die zeer bruikbaar is om de verschillende aspecten van een kwaliteitssysteem uit te leggen is⁵:

Een georganiseerde zorgvuldigheid van samenwerken.

Om systematisch kwaliteit te kunnen leveren moet in de organisatie nagedacht worden hoe de gewenste kwaliteit bereikt kan worden uitgaande van de huidige kwaliteit. Er dienen werkafspraken, procesvolgordes en procedureafspraken te worden vastgesteld, medewerkers met de juiste deskundigheid te worden ingezet en juiste randvoorwaarden aanwezig te zijn zoals een werkend roostersysteem, een goed studievolsysteem, voldoende lokalen en ondersteuning. De organisatie moet met andere woorden zijn zaken zó *organiseren* dat het leveren van de gewenste kwaliteit mogelijk wordt gemaakt. En omdat kwaliteit een bewegend doel is ‘steeds toenemende verwachtingen’ moet de organisatie zijn zaken ook nog eens flexibel organiseren. Iedereen moet dus *zorgvuldig* omgaan en handelen naar die voorwaarden om kwaliteit te leveren.

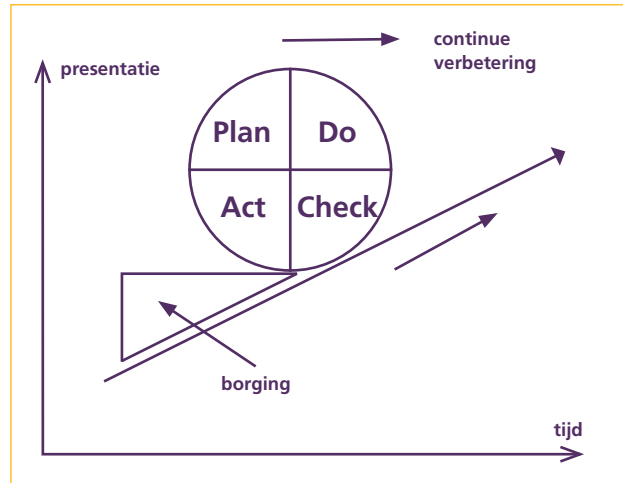
Deze zorgvuldigheid moet niet leiden tot het schoonvegen van eigen straatjes, of het ‘over de muur gooien’ van zaken. Het is namelijk de bedoeling om als *organisatie* kwaliteit te leveren en daar *gezamenlijk aan mee te werken*. Of, zoals Philips dat eens treffend in een personeelsadvertentie verwoordde: ‘Kwaliteit maak je niet alleen, maar wel zelf’.

5.1 PDCA-cyclus

Nu geeft bovenstaande definitie wel aan welke aspecten relevant zijn in een kwaliteitssysteem maar zij maakt niet duidelijk hoe dat handen en voeten gegeven kan worden. Zonder op een specifieke invulling in te gaan, is het mogelijk om invulling te geven aan een kwaliteitssysteem. W. Edwards Deming heeft daarvoor een cirkel gebruikt⁶ (zie figuur 7).

⁵ voor meer definities zie van der Bij, 2004

⁶ Zie van der Bij, 2004



Figuur 7 Deming-cirkel

De Deming-cirkel laat zien, dat er in een organisatie vier soorten activiteiten nodig zijn:

- Plan: Bedenk van tevoren welke producten of diensten geleverd gaan worden, en op welke wijze deze moeten worden gemaakt. Deze activiteit is een belangrijk deel van dat wat hierboven is beschreven bij het *organiseren*.
- Do: Voer uit wat er in Plan bedacht is (*zorgvuldig samenwerken*).
- Check: Verifieer geregeld of dat wat er in Plan bedacht is ook daadwerkelijk in Do heeft plaatsgevonden. Als dat niet het geval is, onderzoek dan waar dat door komt (bijvoorbeeld: door te ambitieuze bewerkingstijden, te weinig of verkeerde ondersteuning, meer klanten en op andere tijden dan verwacht, veranderde klantenwensen). Tezamen met de volgende activiteit is deze activiteit het gedeelte van het *organiseren* dat niet in Plan zit (het *flexibel* zijn).
- Act: Doe iets met de constatering uit Check; het is immers niet zinvol om alleen maar te constateren dat iets niet is gegaan zoals gedacht, zonder daarbij na te gaan of dat in de toekomst vermeden kan worden. Dat betekent dat er wijzigingen in het *organiseren* aangebracht moeten worden.

Niet voor niets is de Deming-cirkel een cirkel: na de besluitvorming van Act, zal er opnieuw ge-Plan-d moeten worden (lees: aanpassingen bedacht worden), waarna dit wordt uitgevoerd, de haalbaarheid gecheckt en acties daarop ondernomen. Beleid, afspraken en procedures worden vastgelegd en op basis van de Check bijgesteld zodat iedereen in de organisatie zich aan de afspraken kan houden en op aangesproken worden.

5.2 In elkaar grijpende PDCA-cycli

De kwaliteitspiramide laat zien dat het realiseren van kwaliteit van toetsen en beoordelen een complexe aangelegenheid is. Het betreft veel verschillende processen, die binnen verschillende teams of afdelingen plaats vinden waarbij veel verschillende mensen van binnen en buiten de organisatie, betrokken zijn. Er is geen sprake van één PDCA-cyclus maar van vele in elkaar grijpende cycli op allerlei niveaus: strategisch, tactisch, operationeel en binnen allerlei functionele gebieden. Een voorbeeld van een PDCA-cyclus op operationeel niveau:

Ontwikkelen en afnemen toets

Plan: Opstellen ontwerpcriteria toets en bepalen werkwijze (onderdeel toetsbeleid)

Do: Ontwikkelen en afnemen toets (op basis van gestandaardiseerde werkwijze)

Check: Toetsanalyse en tevredenheidmeting

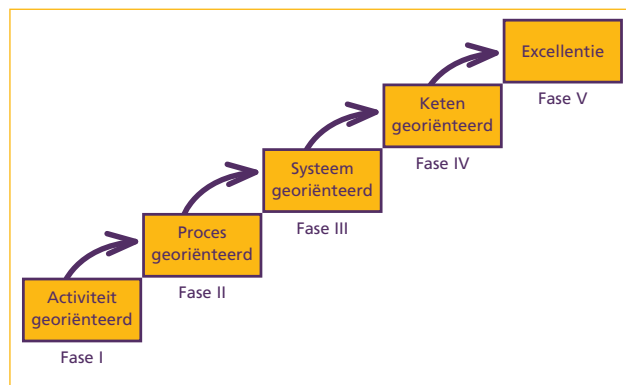
Act: Bijstellen ontwerpcriteria/werkwijze, bijstellen toets, bijscholen examinatoren, examencommissies etc..

Bij deze cyclus zijn examinatoren betrokken en medewerkers vanuit een ondersteunend team. Zij zorgen samen dat er kwaliteit geleverd kan worden ieder vanuit de eigen expertise.

De kwaliteit van de toetsing blijkt uit de analyse van de toetsresultaten maar ook uit de kwaliteit van alle processen waarmee dit resultaat tot stand komt. Elk proces, hoe klein ook, is daarmee object van een PDCA-cyclus.

5.3 Voortdurend verbeteren en vernieuwen

Voortdurend verbeteren en vernieuwen is een continu proces met de PDCA-cyclus als motor voor verandering. Door het steeds explicieter maken van de PDCA-cycli, de toenemende betrokkenheid van medewerkers, studenten en externe belanghebbenden én het steeds meer in samenhang uitvoeren van de verschillende PDCA-cycli ontwikkelt kwaliteitszorg zich naar een steeds hoger niveau. Dat wordt weergegeven in nevenstaand groeimodel⁷.



Figuur 8 Groeimodel organisatie-ontwikkeling

⁷ Ontleend aan INK-model

In het model genoemde generieke kenmerken zijn specifiek ingevuld voor het onderwerp ‘toetsing en examinering’.

Fase 1 Activiteitgeoriënteerd	Deze fase kenmerkt zich door een grote autonomie voor de medewerkers. Er is geen toetsbeleid geformuleerd. Iedereen streeft ernaar om zijn taak zo goed mogelijk uit te voeren en individuele deskundigheid wordt hoog gewaardeerd. Er is veel ruimte voor eigen initiatief en creativiteit en er is weinig bemoeienis van bovenaf. De nadruk ligt op het oplossen van problemen op basis van klachten en incidenten. Er wordt af en toe geëvalueerd.
Fase 2 Procesgeoriënteerd	Er is een toetsprogramma dat door iedereen gehanteerd wordt en de toetsinstrumenten en –opdrachten worden ontwikkeld op basis van kwaliteitscriteria. Het ontwikkelen en afnemen van toetsen wordt beheerst doordat alle processtappen zijn geïdentificeerd en alle taken en verantwoordelijkheden (strategisch, tactisch, operationeel niveau) zijn belegd. Er zijn uniforme afspraken over de manier waarop toetsen worden gemaakt, afgenomen en beoordeeld. Professionalisering van betrokken medewerkers vindt planmatig plaats. Sturing vindt plaats op basis van indicatoren. Op basis van metingen en kennis van de processen worden verbeteringen aangebracht. Alle belangrijke aspecten worden systematisch geëvalueerd.
Fase 3 Systeemgeoriënteerd	Er wordt op alle niveaus gewerkt aan verbetering van de toetsing als geheel. Er is consistent toetsbeleid dat vertaald is naar een adequaat toetsprogramma en toetsorganisatie. Beleid en uitvoering worden in onderlinge samenhang verbeterd. Toetsresultaten worden stelselmatig en in samenhang met elkaar geanalyseerd. Betrokkenen, zoals leden van de examencommissie en examinatoren, professionaliseren zichzelf en elkaar continu. Op alle niveaus wordt gewerkt aan de verbetering van de toetsing. Sturing vindt vooraf, tijdens en achteraf plaats met behulp van indicatoren.
Fase 4 Ketengeoriënteerd	De toetsorganisatie is intern op orde. De manier van werken wordt afgestemd met klanten en (werkveld) partners. Kennis en capaciteiten van alle betrokkenen en partners worden ingebracht en maximaal benut. Met partners wordt structureel aan innovatie gewerkt. De toetsorganisatie is flexibel en anticiperend op de vraag van studenten en werkveldpartners. Er vindt benchmarking plaats.
Fase 5 Excellentie	De organisatie staat met de kwaliteit van toetsing aan de top en is een voorbeeld voor andere, soortgelijke organisaties. Het proces van continu verbeteren is in de organisatiestructuur en -cultuur verankerd. De organisatie beheerst de kunst om tijdig bakens te verzetten vanuit een lange-termijn-perspectief zoals gedefinieerd in het beleid over toetsing en examinering.

Het model beschouwt de inrichting van een kwaliteitssysteem als een ontwikkeling naar een steeds hoger niveau. Bovenstaand model kan een hulpmiddel zijn voor een opleiding en in het bijzonder de examencommissie om te bepalen in welke mate de kwaliteit van toetsing en examinering is geborgd⁸.

⁸ In het INK-model wordt de positiebepaling op negen deelgebieden uitgevoerd. Dat zijn de vijf organisatiegebieden leiderschap, strategie en beleid en management van medewerkers, middelen, processen en de vier resultaatgebieden waaronder klanten en medewerkers

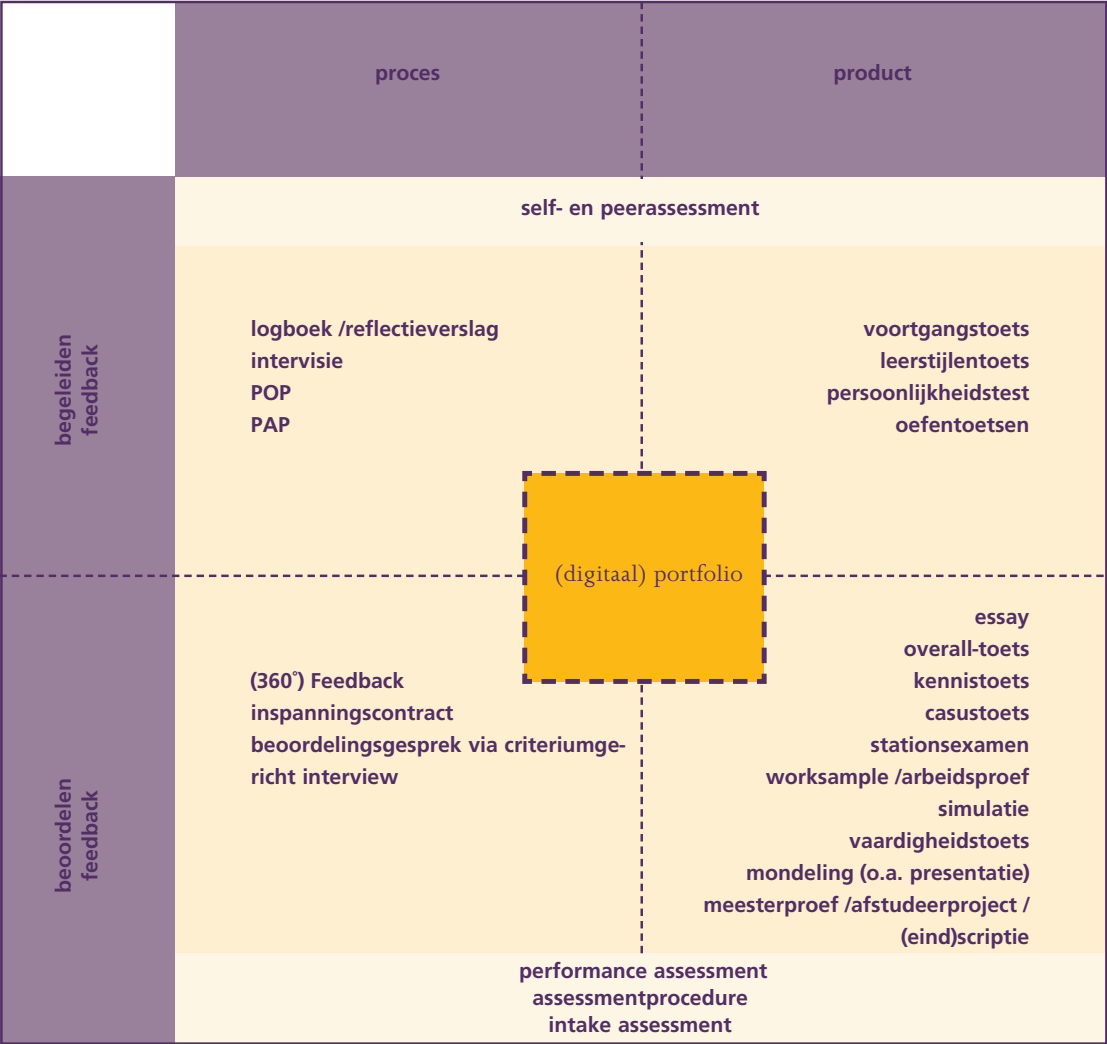
6. Toetsvormen in een competentiegericht leeromgeving

6.1 Indeling in toetsvormen: een werkmodel

Om in de grote hoeveelheid van toetsvormen in een competentiegericht leeromgeving lijn te brengen, is door Jaspers en Heijmen-Versteegen (2004) een werkmodel gedefinieerd (figuur 9). In dit werkmodel is uitgegaan van functies van toetsen - begeleiden en beoordelen - de rol van feedback - en de focus van toetsing namelijk proces en product. Door de functie en focus op twee assen te positioneren ontstaan vier kwadranten, waarin toetsvormen geplaatst worden. In het hart van het model is portfolio opgenomen omdat in die toetsvorm alle onderdelen samenkomen. Uitgaande van de methodemix wordt een keuze(s) op organisatieniveau gemaakt uit elk van de vier kwadranten.

Self-, co- en peerassessment zijn in het model in figuur 3.1 opgenomen als ingang bij het *begeleiden* van het leerproces van de student zowel voor proces als product gerichte toetsvormen (zie § 3.2.1 en § 3.2.2). De bijvoeglijke naamwoorden bij *assessment* die hier gebruikt worden, duiden de methode aan waarmee de daaronder aangegeven toetsvormen worden ingezet. *Selfassessment* is gericht op reflectie van het eigen leerproces, hierbij heeft de student zicht op de beoordelingscriteria bij de taken. Zelfbeoordeling prikkelt studenten om systematisch naar het eigen functioneren te kijken en dit te verbeteren. Het is subjectief van aard en er is geen onderlinge vergelijking tussen studenten mogelijk vanwege diversiteit aan beoordelingsstandaarden. Bij *peerassessment* beoordelen studenten elkaar via beoordelingscriteria. Het is gericht op reflectie, daarnaast op samenwerkend leren. *Performance assessment*, *assessmentprocedure* en *intake assessment* zijn in het model in figuur 9 opgenomen als ingang bij het *beoordelen* van het leerproces van de student zowel voor proces als product gerichte toetsvormen (zie § 3.2.3 en § 3.2.4). 'Een *performance assessment* is een geïntegreerd beoordelingsmoment door de opleiding van het handelen en gedrag van de student in een voor het beroep kritische situatie. Deze methode van toetsing wordt gebruikt om zicht te krijgen op de beheersing van competenties, die voor een beroep kenmerkend of van doorslaggevende betekenis zijn. De beoordeling vindt plaats door een gestandaardiseerde observatie door gekwalificeerde assessoren, die getraind zijn in het observeren, registreren en beoordelen van assessments. Op deze manier wordt de betrouwbaarheid zoveel mogelijk gewaarborgd (Van Brakel & Heijmen-Versteegen, 2003, p.33).' Een *assessmentprocedure* - ook wel *assessment center methode* (ACM) - heeft een betrouwbare voorspellende waarde op prestaties in het werk en opleiding en leerpotentieel. Het maakt in principe gebruik van een gevarieerd aantal instrumenten zoals weergegeven in de onderste helft van het model in figuur 3.1. Een *assessmentprocedure* is kostbaar om te ontwikkelen en te onderhouden. Het vereist specifieke training en doorontwikkeling van assessoren, hun bekwaamheid is essentieel voor de kwaliteit van de *assessmentprocedure*. Het *intake-assessment* richt zich op het in kaart brengen van Eerder, Erkenning of Elders Verworven Competenties en potentieelbeoordeling om een advies en oordeel te geven over (zij)

instroommogelijkheden in een bepaalde opleiding en nog te ontwikkelen competenties. Het vormt de basis voor de beslissing over het opleidingstraject van een student en geeft dus input voor het persoonlijk ontwikkelingsplan (POP).



Figuur 9 Werkmodel competentiegericht leren (Jaspers & Heijmen-Versteegen, 2004)

6.2 Beschrijving van toetsvormen

In deze paragraaf worden de toetsvormen zoals ingedeeld in figuur 3.1 nader uitgewerkt per segment uit de figuur. Voor de beschrijving van de toetsvormen en de verschillende functies zijn diverse bronnen gebruikt, deze zijn opgenomen in de literatuurlijst van deze publicatie.

In deze paragraaf zijn de toetsvormen opgenomen die primair een begeleidend karakter hebben, dat gericht is op het leerproces. Achtereenvolgend worden besproken: Feedback, logboek/reflectieverslag, interview, persoonlijk ontwikkelplan en persoonlijk activiteitenplan.

Toetsvorm	Feedback
<i>Doel</i>	De begeleider geeft op een zodanige manier feed up, feed back, feed forward, dat de student geholpen wordt zelf oplossingen en antwoorden te vinden op vooraf vastgestelde leerdoelen of om nieuw doelen te stellen. De student leert tevens het eigen handelen te evalueren, hierop te reflecteren, de ervaring te verbinden aan andere ervaringen en theoretische concepten.
<i>Kenmerken</i>	Feedback geven is een vorm van het bevorderen van leren en het sturen van <i>ontwikkeling</i> . De begeleider helpt de student bij het formuleren van de volgende vragen: Waar ga ik naar toe, hoe doe ik het tot nu toe en wat kan ik doen om dichter bij het leerdoel te komen of om nieuwe doelen te stellen?
<i>Functie</i>	Formatief/summatief
<i>Betrokkenen bij het ontwerp</i>	Begeleider en student
<i>Betrokkenen bij de beoordeling</i>	Begeleider en student
<i>Feedback op de toets</i>	Informatie over eigen leerproces en niveau in relatie tot de geformuleerde eindkwalificaties/competenties.

Toetsvorm	Logboek /reflectieverslag
<i>Doel</i>	De eigen vooruitgang erkennen en patronen in het eigen gedrag te identificeren. Dit alles ter bevordering van de eigen competentieontwikkeling.
<i>Kenmerken</i>	• Beschrijving van ervaringen en reflectie daarop • Beschrijving van gevoelens • Gestructureerde beschrijving • Feedback is essentieel • De validiteit en betrouwbaarheid van het resultaat worden sterk beïnvloed door factoren als introspectief vermogen, motivatie, eerlijkheid en belangen die op het spel staan
<i>Functie</i>	Formatief. Kan summatief worden ingezet als het via het portfolio ter beoordeling wordt aangeboden door de student.
<i>Betrokkenen bij het ontwerp</i>	De student en eventueel de studieloopbaanbegeleider van de opleiding
<i>Betrokkenen bij de beoordeling</i>	De student, medestudenten, de studieloopbaanbegeleider en eventueel het werkveld, steeds op initiatief van de student
<i>Feedback op de toets</i>	De evaluatie van lezers en de zelfbeoordeling van studenten worden regelmatig besproken. Feedback kan zowel schriftelijk als mondeling met als doel de bevordering van de competentieontwikkeling van de student.

Toetsvorm	Intervisie
<i>Doel</i>	Verbetering van het handelingsrepertoire en probleemoplossend vermogen van de student.
<i>Kenmerken</i>	• Feedback geven door medestudenten aan een door de student zelf ingebrachte cases • De student combineert de verschillende invalshoeken, deskundigheden en ervaringen van medestudenten om tot nieuwe oplossingen en inzichten te komen • Draagt bij aan ontwikkeling van de reflectieve vaardigheden van groepsleden • Communicatieve vaardigheden vormen het fundament van leerzame en geslaagde intervisiebijeenkomsten • Vaste of wisselende groep van maximaal zes personen • Begeleider die het leerproces en het communicatieproces van de studenten in de groep faciliteert. • Vrijwilligheid van deelname • Heldere methodiek en doelen • Vertrouwelijkheid • Elke groep stelt zelf de regels vast • Het gaat om ondersteuning en uitdaging: Er worden geen oordelen uitgesproken
<i>Functie</i>	Formatief.
<i>Betrokkenen bij het ontwerp</i>	Groep van studenten en een begeleider. De begeleider is of een opleidingsdocent of een student met een meer gevorderd competentieniveau.
<i>Betrokkenen bij de beoordeling</i>	Groep van studenten en begeleider
<i>Feedback op de toets</i>	Mondeling en te verwerken in het reflectieverslag of portfolio.

Toetsvorm	Persoonlijk Ontwikkel Plan of Persoonlijk Opleidings Plan (POP)
<i>Doel</i>	Het prikkelen van studenten om zichzelf en hun pakket aan vragen, eisen en leerwensen in kaart te brengen en daarop te reflecteren. Het vastleggen van individuele afspraken over leerarrangementen. Deze staan in dienst van de te verwerven competenties.
<i>Kenmerken</i>	Geeft aan het waarom, wat, hoe, waar, wanneer, met wie en welke opbrengst Geeft aan op welke manier geëvalueerd /getoetst wordt Geeft aan welke leerarrangementen gevolgd worden Het POP is geen doel op zich, leren is het doel De neiging er een papieren document van te maken is bijzonder groot, past volledig in de structuur, die opleidingen gewend zijn
<i>Functie</i>	Formatief; summatief als het voorstel van de student moet worden goedgekeurd door de opleiding (bijvoorbeeld bij minoren).
<i>Betrokkenen bij het ontwerp</i>	Student en studieloopbaanbegeleider
<i>Betrokkenen bij de beoordeling</i>	Studieloopbaanbegeleider en indien goedkeuring nodig is: Examencommissie van de betreffende opleiding.
<i>Feedback op de toets</i>	Mondeling en individueel

Toetsvorm	Persoonlijk Activiteiten Plan (PAP)
<i>Doel</i>	Het plannen en uitwerken van de leerarrangementen zoals vastgelegd in het POP.
<i>Kenmerken</i>	Plannen ontwikkelactiviteiten naar: tijd; plaats; studieactiviteiten.
<i>Functie</i>	Formatief
<i>Betrokkenen bij het ontwerp</i>	Student en studieloopbaanbegeleider
<i>Betrokkenen bij de beoordeling</i>	Student en studieloopbaanbegeleider
<i>Feedback op de toets</i>	Mondeling en individueel

Toetsvormen voor begeleiding gericht op het product

In deze paragraaf zijn de toetsvormen opgenomen die primair een begeleidend karakter hebben, dat gericht is op een product. Achtereenvolgend worden besproken: Voortgangstoets, leerstijlentoets en persoonlijkheidstest. Met oefentoeetsen wordt in het model bedoeld dat de toetsvormen genoemd in de beoordelende kwadranten in een oefensituatie kunnen worden ingezet door de student.

Toetsvorm	Voortgangstoets
<i>Doel</i>	Metten van de mate waarin de student de cognitieve doelstellingen van de totale opleiding beheerst. Geeft de student en de opleiding inzicht in het ontwikkelingsniveau van de student op het gebied van kennis. Het biedt de student en ook de opleiding informatie over mogelijke bijsturing van het leerproces.
<i>Kenmerken</i>	• Toets op eindniveau van de opleiding • Periodieke toets • Geschikt voor grote groepen studenten in redelijk homogene opleidingen • Curriculum-onafhankelijke toets • Toetsgericht studeren is niet mogelijk • Gericht op functionele, beklijvende kennis • Hoge betrouwbaarheid is realiseerbaar • Geen herkansingen nodig • Talige toets • Centrale organisatie vereist voor ontwikkeling van de toets, mogelijk buiten de grenzen van een individuele opleiding
<i>Functie</i>	Oorspronkelijk formatief. Er zijn veel voorbeelden van opleidingen die de voortgangstoets summatief inzetten, dit gaat in tegen het eigenlijke doel van de toets.
<i>Betrokkenen bij het ontwerp</i>	Multidisciplinair team van docenten/toetsontwikkelaars en werkveld, eventueel ook studenten
<i>Betrokkenen bij de beoordeling</i>	Bij voorkeur de toetscommissie van de opleiding, die de coördinatie van de voortgangstoets op zich neemt
<i>Feedback op de toets</i>	Percentage beheersing mogelijk op subonderwerpen van de toets, vaak in vergelijking met andere studenten in dezelfde groep.

Toetsvorm	Leerstijltoets
<i>Doel</i>	Student krijgt inzicht in de eigen manier van leren. Dit inzicht kan gebruikt worden bij de vormgeving van de studieloopbaan (relatie met POP en PAP, zie hiervoor).
<i>Kenmerken</i>	- Betrouwbaarheid van bestaande tests is aanvaardbaar - Plaats en tijdonafhankelijk af te nemen - Talige toets
<i>Functie</i>	Formatief
<i>Betrokkenen bij het ontwerp</i>	Gestandaardiseerde reeds beproefde toets, bijvoorbeeld Inventaris Leerstijlen gebaseerd op J. Vermunt
<i>Betrokkenen bij de beoordeling</i>	De student en eventueel de tutor of studieloopbaanbegeleider
<i>Feedback op de toets</i>	Directe schriftelijke feedback op afzonderlijke aspecten van de leerstijl van de student.

Toetsvorm	Persoonlijkheidstest
<i>Doel</i>	Inzicht geven in de persoonlijkheid van de student en via portfolio relatie leggen met het beroep waarvoor wordt opgeleid.
<i>Kenmerken</i>	Enige voorspellende validiteit voor sommige personeigenschappen in specifieke situaties Gemakkelijk en goedkoop te onderhouden Risico van bedreiging privacy Talige test
<i>Functie</i>	Formatief, mag nimmer summatief ingezet worden vanuit ethisch oogpunt.
<i>Betrokkenen bij het ontwerp</i>	Bestaande en beproefde persoonlijkheidstests. Zie onder andere boek: Welke kleur heeft jouw parachute? door R.N. Bolles (2003), waar meerdere tests in zijn beschreven.
<i>Betrokkenen bij de beoordeling</i>	De student en eventueel de studieloopbaanbegeleider of coach
<i>Feedback op de toets</i>	Inzicht in persoonskenmerken en drijfveren

Toetsvormen voor beoordeling gericht op het leerproces

In deze paragraaf zijn de toetsvormen opgenomen die primair een begeleidend karakter hebben, dat gericht is op het leerproces. Achtereenvolgend worden besproken: 360° feedback, inspanningscontract en beoordelingsgesprek via criteriumgericht interview.

Toetsvorm	360° feedback
<i>Doel</i>	Het beoordelen van competenties die op de werkplek, in afstudeeropdrachten en projectonderwijs centraal staan. Te denken valt aan sociale en management competenties of aan mondeling communiceren.
<i>Kenmerken</i>	• Directe manier om informatie over gedragscompetenties in te winnen • Inzet van meerdere beoordelaars resulteert in hoge betrouwbaarheid en validiteit • Tijdrend voor veel betrokkenen • Afhankelijk van goede bedoelingen en discipline van betrokken actoren • Niet bruikbaar voor selectiedoeleinden • Het doet enkel uitspraken over door de student 'getoonde' competenties in de huidige context • Schriftelijke vragenlijst die ingevuld wordt door de student, personen op de werkplek, de stagebegeleider op de werkplek en eventueel ook de stagebegeleider van de onderwijsinstelling
<i>Functie</i>	Summatief, mits is voldaan aan eisen van betrouwbaarheid en validiteit. Deze toets kan ook formatief worden ingezet.
<i>Betrokkenen bij het ontwerp</i>	Coach en /of studieloopbaanbegeleider
<i>Betrokkenen bij de beoordeling</i>	Iedereen die de student (in die situatie) 'kent'
<i>Feedback op de toets</i>	Vanuit meerdere invalshoeken feedback op handelen en /of competenties van de student.

Toetsvorm	Inspanningscontract / presentieverplichting
<i>Doel</i>	Het onderstrepen van het wezenlijke belang van deelname om succesvol gedrag te vertonen. Het leert de student dat verantwoordelijkheid naar de eigen leergroep, de docenten en het opleiding met alle kostbare voorzieningen hiermee concreet vorm heeft.
<i>Kenmerken</i>	• In het contract is geformuleerd welke inspanningen een student moet leveren om een voldoende beoordeling te krijgen • In het contract is opgenomen wat de gevolgen zijn, wanneer de student aan de bepaalde verplichtingen niet voldoet • Altijd in samenhang met een andere toetsvorm
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	De opleiding gaat een wederzijdse verplichting (contract) met de student aan.
<i>Betrokkenen bij de beoordeling</i>	Opleiding en student
<i>Feedback op de toets</i>	Voldaan of niet voldaan

Toetsvorm	Beoordelingsgesprek via criteriumgericht interview
<i>Doel</i>	Het beoordelen van competenties via in het verleden concreet waarneembaar gedrag en/of expliciteren van in het verleden getoond gedrag voor evaluatie van en de reflectie op dat gedrag via een gerichte vraagmethodiek (STARR: Situatie – taak – actie – resultaat – reflectie)
<i>Kenmerken</i>	• Redelijk valide mits gestructureerd en gebaseerd op beroepanalyse • Criteria zijn vooraf aan de student bekend gemaakt • Kwaliteit van interview is sterk afhankelijk van de kwaliteiten van de interviewer
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Opleiding en eventueel student(en)
<i>Betrokkenen bij de beoordeling</i>	Getrainde interviewers die de STARR-methodiek kunnen toepassen.
<i>Feedback op de toets</i>	Vanuit meerdere criteria feedback op handelen en/of competenties van de student.

Toetsvormen voor beoordeling gericht op het product

In deze paragraaf zijn de toetsvormen opgenomen die primair een beoordelend karakter hebben, dat gericht is op het product. Achtereenvolgend worden besproken: Essay, overall-toets, kennistoets, casustoets, stationsexamen, worksample/arbeidsproef, simulatie, vaardigheidstoets, mondeling (onder andere presentatie) en meesterproef/afstudeerproject / (eind)scriptie.

Toetsvorm	Essay
<i>Doel</i>	Toetsen van kennis, begrip, inzicht en toepassing van kennis.
<i>Kenmerken</i>	• Reproductie van feiten • Leggen van combinaties of toepassingen van kennis • Antwoorden bestaan uit opsommingen, redeneringen, betogen en argumentaties • Antwoordsleutel vereist • Betrouwbaarheid is lastig punt: Er zijn vaak meerdere – ondanks antwoordsleutel - interpretaties mogelijk door de beoordelaars • Snelle constructie tegenover lange correctietijd • Talige toets • Breed toepassingsgebied • Student moet zelf antwoorden formuleren
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Multidisciplinair team van docenten/toetsontwikkelaars, waar mogelijk onder coördinatie van een toetscommissie
<i>Betrokkenen bij de beoordeling</i>	Inhoudskundige docenten/toetsontwikkelaars van de opleiding, die deel uitmaken van het multidisciplinair team
<i>Feedback op de toets</i>	De toetsuitslag geeft de student inzicht in sterke en zwakke punten door differentiatie in aparte scores.

Toetsvorm	Overall-toets
<i>Doel</i>	De overall-toets (OALT) richt zich op vijf competenties. De toets meet in welke mate de student: 1. Een nieuw probleem of aspect van een probleemsituatie kan <i>definiëren</i> door eerder bestudeerde relevante begrippen, modellen of theorieën. 2. Een nieuw probleem kan <i>analyseren</i>. 3. De resultaten van de analyse kan <i>synthetiseren</i>, 4. Mogelijke oplossingen of te nemen beslissingen kan <i>beargumenteren</i>, 5. Oplossingen of beslissingen kan <i>evalueren</i>.
<i>Kenmerken</i>	De OALT is ontwikkeld in een probleemgestuurd curriculum. De OALT meet probleem-oplossende vaardigheden en heeft tien karakteristieken (Segers, 2002, p.141): 1. Elke toetsvraag is gebaseerd op, en refereert aan probleemsituaties zoals beschreven in artikelen; de artikelen waarin de probleemsituaties zijn beschreven, zijn verschillend van aard. 2. De toets is gebaseerd op een set van artikelen. 3. De artikelen beschrijven een probleemsituatie in haar totaliteit. 4. De probleemsituaties zijn besproken tijdens de onderwijsgroep bijeenkomsten. 5. Probleemsituaties worden vanuit verschillende disciplines benaderd. 6. De probleemsituaties worden tijdens een zelfstudieperiode bestudeerd. 7. De toetsvragen richten zich op kernaspecten van de probleemsituaties. 8. Zowel gesloten- als de open-vraagvorm wordt gebruikt waarbij beide vraagvormen zich richten op een specifiek beheersingsniveau. 9. De toets heeft een open-boekkarakter. 10. De toets wordt door een multidisciplinair team geconstrueerd.
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Multidisciplinair team van docenten/ toetsontwikkelaars
<i>Betrokkenen bij de beoordeling</i>	Multidisciplinair team van docenten/ toetsontwikkelaars; ook mogelijkheden voor co- en peerassessment
<i>Feedback op de toets</i>	Resultaat. Studenten geven aan behoefte te hebben aan discussie over de mogelijkheden en beperktheden van het geleerde bij varianten op de probleemtaak.

Toetsvorm	Kennistoets
<i>Doel</i>	Het vaststellen van het niveau van de student van beroeps- en vakgerichte kennis.
<i>Kenmerken</i>	• Verschillende verschijningsvormen: Toetsen met gesloten vragen of met open vragen of met een combinatie van beide • Kan zowel via papier als via computer worden afgenomen • Talige toets • Bij grote groepen en gesloten vragen is betrouwbaarheid en validiteit goed in kaart te brengen • Bij gesloten vragen kan het nakijken van de toets automatisch geschieden, biedt dan mogelijkheden voor controle van betrouwbaarheid van de toets • Vaak ter afsluiting van een bepaalde studieperiode afgenomen
<i>Functie</i>	Summatief; heel goed ook formatief in te zetten als oefentoets.
<i>Betrokkenen bij het ontwerp</i>	Inhoudsdeskundige docenten/toetsontwikkelaars van de opleiding en eventueel studenten of werkveld
<i>Betrokkenen bij de beoordeling</i>	Docenten/toetsontwikkelaars van de opleiding
<i>Feedback op de toets</i>	Resultaat, meestal een cijfer berekend via de goed en fout beantwoorde vragen. Zo mogelijk uitgesplitst naar kennisgebied binnen de toets. Indien geautomatiseerd ook een vergelijking met de resultaten van de overige studenten in de groep.

Toetsvorm	Casustoets
<i>Doel</i>	Het beoordelen van het niveau van kennis en vaardigheden van een student via een probleem of gevalsbeschrijving, waarin een beroep wordt gedaan op het probleemoplossende vermogen van de student.
<i>Kenmerken</i>	• Probleem of gevalsbeschrijving, ontleend aan de beroepspraktijk • Generieke hbo-competenties worden getoetst • Beroepspecifieke kennis wordt getoetst • Kan onderdeel zijn van een kennistoets of OAT • Stimuleert multidisciplinair en creatief denken bij de student
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Docenten/toetsontwikkelaars via informatie uit het werkveld
<i>Betrokkenen bij de beoordeling</i>	Docenten/toetsontwikkelaars en eventueel werkveld en/of medestudenten
<i>Feedback op de toets</i>	Resultaat, meestal in de vorm van een cijfer. Mogelijk een nabespreking of eindgesprek met toelichting op: Kwaliteit van de analyse van de student, kwaliteit van de probleemoplossing, manier van werken van de student.

Toetsvorm	Stationsexamen
<i>Doel</i>	De student laat zien dat hij in diverse gecontroleerde (simulatie)settings een representatieve reeks vaardigheden beheerst en kan uitvoeren en zomogelijk kort kan antwoorden. De student leert dat vaardigheden een essentieel element vormen van de beroepsvorming. Oefenen daarvan baart kunst. Ook leert de student onder tijdsdruk te werken en binnen een tijdsbestek van 1-2 uur een veelheid aan verschillende vaardigheden uit te voeren. Kennistoetsing (zie hiervoor) kan worden opgenomen in een stationsexamen.
<i>Kenmerken</i>	• Reeks van vaardigheden die de student volgens opdracht uitvoert in een gestelde tijd in een vastgesteld aantal ruimten • Een ruimte heet 'station' • De uitgevoerde vaardigheid wordt direct geobserveerd • Een compleet stationsexamen bestaat uit een circuit van meerdere stations (8-12)
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Docenten / toetsontwikkelaars in overleg met het werkveld
<i>Betrokkenen bij de beoordeling</i>	In elk station wordt de student beoordeeld door een observator. De observator is een getrainde deskundige of een getrainde simulatiepatiënt met een beoordelingsformulier. De beoordelingen worden ingevuld op een scanformulier voor een automatische (of handmatige) verwerking.
<i>Feedback op de toets</i>	Een beoordeling op een 5-puntsschaal: Onvoldoende /matig /voldoende /ruim voldoende / goed.

Toetsvorm	Worksample / arbeidsproef
<i>Doel</i>	Toetsing van competentie en kennis en vaardigheden.
<i>Kenmerken</i>	Taak wordt in een authentiek context uitgevoerd Student is op de hoogte van de beoordelingssituatie Zeer realistische manier van beoordelen In het algemeen is de validiteit relatief hoog Hoge acceptatiegraad voor zowel de te beoordelen persoon als voor de beoordelaars Ontwikkeling is relatief duur Afname kost veel tijd, waardoor het aantal uit te voeren taken beperkt wordt gehouden. Hierdoor kunnen verkeerde conclusies getrokken worden.
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Opleidingsdocenten en werkveld, eventueel student
<i>Betrokkenen bij de beoordeling</i>	Beoordelaar van de opleiding en van de werkplek, eventueel de cliënt of de patiënt
<i>Feedback op de toets</i>	Een eindcijfer, dat de gecombineerde waardering betreft van de volgende aspecten: In welke mate de student kan functioneren als een zelfstandige professional; in hoeverre de student in staat is het eigen product kritisch te beoordelen; de kwaliteit van het product van de student en hoe dit op anderen overkomt.

Toetsvorm	Simulatie
<i>Doel</i>	Metten van één of meerdere competenties in een realistische, maar gestandaardiseerde situatie.
<i>Kenmerken</i>	• Nagebootste reële werkomgeving in rollenspel. • Gericht op klantgerichte vaardigheden. • Gestandaardiseerde situatie. • Benadert de authentieke situatie. • Vraagt veel voorbereiding. • Verschillende vormen van simulaties: De simulator, de gesprekssimulatie, de computersimulatie en de acteurssimulatie (Bergsma, 2003).
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Opleidingsdocenten en werkveld
<i>Betrokkenen bij de beoordeling</i>	Getrainde observatoren die beoordelen via een lijst met beoordelingscriteria
<i>Feedback op de toets</i>	Een score die kan worden opgenomen in het portfolio en voor POP en PAP.

Toetsvorm	Vaardigheidstoets
<i>Doel</i>	De toets controleert of de student over de gewenste vaardigheden beschikt, ofwel kan demonstreren dat bepaalde beroepsvaardigheden correct en adequaat worden uitgevoerd.
<i>Kenmerken</i>	• Training van observatoren vergroot de betrouwbaarheid • Criterialijsten van de beoordeling kunnen meer gedetailleerd of meer globaal van karakter zijn, afhankelijk van het doel van de toetsing • Validiteit over het algemeen voldoende te waarborgen • Duur van de beoordeling moet lang genoeg zijn, dit is een probleem bij de uitvoering • Arbeidsintensief
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Docententeams en eventueel werkveld
<i>Betrokkenen bij de beoordeling</i>	Getrainde assessoren
<i>Feedback op de toets</i>	Een cijfer en een mondelinge of schriftelijke toelichting. De feedback kan gegeven worden door de assessoren/docenten, medestudenten en de student zelf (reflectie).

Toetsvorm	Mondeling (onder andere presentatie)
<i>Doel</i>	Het kunnen presenteren, redeneren en daarover mondeling communiceren en discussiëren over een onderwerp.
<i>Kenmerken</i>	• Lage betrouwbaarheid. • Moeilijk te garanderen validiteit. • Zeer strikte regels rond afname en procedure vereist. • Deze toetsvorm wordt in het algemeen niet aanbevolen wegens grote subjectiviteit. • Vooraf is bekend wat onderwerp van toetsing is: De inhoud, het presenteren als vaardigheid of beide. • Mondeling wordt altijd in samenhang met andere toetsvormen afgenomen
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Docententeams en eventueel studenten
<i>Betrokkenen bij de beoordeling</i>	Assessoren
<i>Feedback op de toets</i>	Veelal mondeling direct na afname.

Toetsvorm	Meesterproef /afstudeerproject /(eind)scriptie
<i>Doel</i>	Toetsing van competent zijn op bachelorniveau. Is de student startbekwaam?
<i>Kenmerken</i>	• Taak wordt in een authentiek context uitgevoerd • Student is op de hoogte van de beoordelingssituatie • Zeer realistische manier van beoordelen • In het algemeen is de validiteit relatief hoog • Student krijgt gelegenheid om het competent zijn zelf aan te bewijzen • Arbeidsintensief • Levert een bijdrage aan de doorontwikkeling van het beroep (innovatief vermogen)
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Opleiding en student.
<i>Betrokkenen bij de beoordeling</i>	Beoordelaar van de opleiding en zo mogelijk van de werkplek.
<i>Feedback op de toets</i>	Een eindbeoordeling, dat de gecombineerde waardering betreft van de volgende aspecten: In welke mate de student kan functioneren als een zelfstandige professional; in hoeverre de student in staat is het eigen product en zichzelf kritisch te beoordelen.

(Digitaal) portfolio

Via het (digitaal) portfolio komt alles met alles in samenhang. Het portfolio is voor de student leidend of dragend voor de opleiding.

Toetsvorm	(digitaal) portfolio
<i>Doel</i>	Een (digitaal) portfolio heeft of kan een aantal doelstellingen hebben: • Beoordelen van de student. • Begeleiden van het leerproces van de student. • Presenteren van de competenties door de student (showmap).
<i>Kenmerken</i>	• Vooruitblik, reflectie en inzichtelijk maken van het leerproces. • Authentiek beeld van de student. • Ruimte voor individuele profilering van de student. • Instrument bij vraaggestuurde curricula: Gaat uit van de leerbehoefte van de student. • Beoordeling via verschillende bronnen die in portfolio samen komen. • Instrument voor begeleiding en beoordeling. • valide; • actueel; • dynamisch; • inhoudsgericht; • leren-leren gericht; • rijk aan bewijzen van competentie; • interactief; • longitudinaal.
<i>Functie</i>	Summatief
<i>Betrokkenen bij het ontwerp</i>	Student, medestudent en opleiding
<i>Betrokkenen bij de beoordeling</i>	Student, medestudent, opleiding en eventueel werkveld
<i>Feedback op de toets</i>	Continue

Bijlage 1 **Format Jaarverslag examencommissie Fontys Hogescholen, september 2010**

Jaarverslag Examencommissies van de opleiding (jaartal invullen)

Met ingang van september 2010 bepaalt de WHW dat de examencommissies, in het kader van de gewenste transparantie, jaarlijks een verslag van haar werkzaamheden moet opstellen ten behoeve van het instellingsbestuur. Het jaarverslag dient per kalenderjaar opgesteld te worden aan de hand van dit format. Richtlijn: maximaal drie pagina's.

Voorwoord

In het voorwoord kan kort vermeld worden waarom dit jaarverslag is opgesteld, wat in dit verslag te lezen is. Ook kunnen eventueel bijzondere gebeurtenissen, ontwikkelingen genoemd worden die onder de aandacht gebracht mogen worden.

Dit jaarverslag is opgesteld op grond van artikel 7.12b lid 4 van de WHW en bevat een verslag van de werkzaamheden van de examencommissie in het jaar 2010.

1. Werkwijze

In dit onderdeel wordt beschreven hoe de examencommissie is samengesteld en hoe haar werkwijze is. Wanneer de examencommissie werkt met subcommissies/kamers kan dat hier opgenomen worden. Denk daarbij ook aan de volgende onderwerpen. Hoe is de vergaderfrequentie? Is er een rol voor externe leden? Welke? Hoe vindt verslaglegging plaats? Hoe vindt besluitvorming plaats (denk bv. aan het opstellen van een nieuwe procedure, maar ook aan hoe besloten wordt over verzoeken van individuele studenten)? Hoe vindt communicatie/informatievoorziening plaats (is er een nieuwsbrief, een portal etc.)?

De examencommissie bestaat uit de volgende leden:

Naam	functie in organisatie	functie in examencommissie	lid sinds	lid tot
.. ..	..	..	..	..

De examencommissie vergadert één keer per

2. Professionalisering

In dit onderdeel wordt beschreven wat de examencommissie doet aan professionalisering van zichzelf, maar ook van de organisatie of van docenten (bv. met betrekking tot toetsing). Geef ook aan op welke terreinen professionalisering heeft plaats gevonden.

3. Resultaten

3a. Beleidsmatige taken

In dit onderdeel wordt opgenomen wat de examencommissie aan beleid heeft ontwikkeld dit jaar, denk aan fraudebeleid, vrijstellingenbeleid, beleid met betrekking tot cum laude, tot het verstrekken van voorzieningen o.a. aan studenten met chronische ziekte/handicap, langstudeerders e.d..

3b. Uitvoerende taken

Hier kan feitelijke, meer statistische informatie opgenomen worden, eventueel ook in een eenvoudig overzicht (geen namen van studenten noemen)

- beoordeling examens/vaststelling geslaagden/niet geslaagden (aantallen: propedeuse/bachelor/master)
- uitbrengen studieadviezen (aantallen, welke soorten adviezen, positief, voorwaardelijk, afwijzend)
- beoordelen meldingen van fraude (aantal, welke sancties)
- beoordelen van individuele verzoeken van studenten, o.a. met betrekking tot:
 - verlenen vrijstelling (aantal, gehonoreerd of afgewezen)
 - voorzieningen voor studenten met chronische ziekte/handicap/topsport (aantal, welke voorzieningen)
 - extra herkansing (aantal, gehonoreerd of afgewezen)
 - verzoek om herbeoordeling/second opinion (aantal, gehonoreerd of afgewezen)
 - hardheidsclausule, inclusief verzoek om alternatief afstudeertraject (aantal, gehonoreerd of afgewezen)
- beroepszaken
 - aantal beroepschriften ingediend bij het College van beroep voor de examens
 - welke onderwerpen (aantallen per onderwerp)
 - schikking (in hoeveel gevallen is het gekomen tot een schikking)
 - oordeel (wat was het oordeel van het College)

3c. Evaluatieve taken

Hier kan informatie opgenomen worden over toetsanalyses en andere activiteiten die hebben plaatsgevonden ten behoeve van het monitoren van de toetskwaliteit en het maken van procedures.

4. Nawoord

Welke plannen zijn er voor het volgend jaar?

Bijlage 2 Competentie Assessment Programma (CAP): twaalf kwaliteitscriteria en indicatoren (Baartman e.a.)

Acceptatie

- De leerlingen kunnen zich vinden in de beoordelingscriteria.
- De leerlingen kunnen zich vinden in de wijze waarop het CAP wordt uitgevoerd.
- De docenten kunnen zich vinden in het doel, de beoordelingscriteria en de procedures van het CAP.
- Het beroepenveld kan zich vinden in de beoordelingscriteria en de procedures van het CAP.
- Leerlingen, docenten en het beroepenveld hebben vertrouwen in de kwaliteit van het CAP.

Authenticiteit

- De beoordelingsopdracht bevat activiteiten die leerlingen in hun toekomstig beroep moeten uitvoeren.
- De werkomstandigheden tijdens de beoordeling lijken op die in het toekomstig beroep.
- De sociale context tijdens de beoordeling (bijvoorbeeld samenwerking) lijkt op die in het toekomstig beroep.
- De beoordelingscriteria lijken op de criteria waarop werknemers in het toekomstig beroep worden afgerekend.

Cognitieve complexiteit

- De beoordelingsopdrachten roepen de denkstappen op die beginnend beroepsbeoefenaren hanteren.
- Bij het uitvoeren van een beoordelingsopdracht moet een leerling uitleggen waarom hij bepaalde keuzes heeft gemaakt.
- De beoordelingscriteria zijn (ook) gericht op de gehanteerde denkstappen.
- De beoordelingsopdrachten vereisen het denkniveau dat beginnend beroepsbeoefenaren nodig hebben.

Vergelijkbaarheid

- De beoordelingsopdrachten zijn voor alle leerlingen vergelijkbaar en eventuele verschillen worden verantwoord.
- De werkomstandigheden zijn voor alle leerlingen vergelijkbaar en met eventuele verschillen wordt rekening gehouden in het oordeel.
- De beoordelingscriteria zijn voor alle leerlingen vergelijkbaar en eventuele verschillen worden verantwoord.
- De beoordelingsprocedure is voor alle leerlingen vergelijkbaar en eventuele verschillen worden verantwoord.

Tijd en kosten

- Vooraf aan de implementatie van een CAP wordt een inschatting gemaakt van de tijd en kosten.
- Bij de keuze voor een mix van beoordelingsbronnen in een CAP wordt rekening gehouden met de tijd en de kosten.
- Jaarlijks wordt geëvalueerd of het CAP efficiënter kan worden uitgevoerd.
- De positieve gevolgen van het CAP wegen op tegen de tijd en kosten die worden geïnvesteerd.

Onderwijsgevolgen

- Het CAP roept bij de leerlingen de gewenste leerprocessen op in de voorbereiding naar een beoordeling.
- De leerlingen worden op een positieve manier beïnvloed door het CAP.
- De docenten worden op een positieve manier beïnvloed door het CAP.
- Het CAP wordt verbeterd als onverwachte negatieve gevolgen worden gevonden.
- Het curriculum wordt aangepast als de resultaten van het CAP dit vereisen.

Eerlijkheid

- Er zijn procedures vastgesteld voor het corrigeren van eventueel gemaakte fouten tijdens de beoordeling.
- Op basis van de belangrijkheid van de (deel)competenties is een weging vastgesteld.
- De beoordelaars zijn niet bevooroordeeld.
- De beoordelingsopdrachten zijn gevarieerd.
- Leerlingen ervaren de beoordeling als eerlijk.

Geschiktheid voor onderwijsleerdoelen

Het CAP dekt het competentieprofiel / de eindtermen.

- In het CAP worden kennis, vaardigheden en attitudes geïntegreerd beoordeeld.
- Het CAP bestaat uit een mix van verschillende beoordelingsvormen.
- Het CAP bevat zowel summatieve als formatieve beoordelingsvormen.
- De gekozen beoordelingsvormen passen bij de leerdoelen van het onderwijs.

Ontwikkeling van zelfsturend leren

De leerlingen beoordelen zichzelf of elkaar.

- De beoordelingsvormen stimuleren het op een goede manier (leren) geven en ontvangen van feedback.
- De beoordelingsvormen stimuleren het (leren) reflecteren op de eigen ontwikkeling.
- De beoordelingsvormen stimuleren het formuleren van eigen leerdoelen, gebaseerd op de eigen ontwikkeling.

Betekenisvolheid

- De leerlingen vinden de feedback van de formatieve beoordelingsmomenten zinvol voor hun leerproces.
- De leerlingen vinden de feedback van de summatieve beoordelingsvormen zinvol voor hun leerproces.
- De leerlingen ervaren de beoordeling als leermoment.
- De leerlingen vinden de beoordelingscriteria betekenisvol met betrekking tot hun toekomstig beroep.
- De docenten en bedrijven vinden de beoordelingscriteria betekenisvol met betrekking tot de eisen die zij stellen aan toekomstige beroepsbeoefenaars.

Herhaalbaarheid van beslissingen

- Voor een summatief eindoordeel wordt elke competentie meerdere keren beoordeeld.
- Voor een summatief eindoordeel wordt het oordeel van meerdere beoordelaars gecombineerd.
- Bij summatieve beoordeling worden beoordelaars met een verschillende achtergrond ingezet (bijvoorbeeld docent/bedrijfsleven/vervolgonderwijs/medeleerling).
- Tussen de verschillende beoordelaars vindt een gelijkwaardig overleg plaats waarin iedereen zijn oordeel onderbouwt.
- De beoordelaars zijn getraind en competent voor de verschillende beoordelingsvormen.
- Een summatief oordeel wordt gebaseerd op beoordelingen in verschillende werksituaties.

Transparantie

- De leerlingen weten of een beoordeling formatief of summatief is bedoeld.
- De leerlingen kennen en begrijpen de beoordelingscriteria.
- De leerlingen weten en begrijpen hoe het CAP wordt uitgevoerd.
- De docenten kennen en begrijpen het doel, de criteria en de procedure van het CAP.
- De praktijkbegeleiders kennen en begrijpen het doel, de criteria en de procedure van het CAP.
- Een externe partij kan op basis van de vastgestelde procedures en de beschrijving van het uitgevoerde CAP een controle uitvoeren.

Bijlage 3 Checklist Geautomatiseerd toetsen

Zestien tips voor het formuleren van gesloten vragen

Relevantie

1. Kies onderwerpen/vraagstukken die passen bij het te toetsen onderwijsresultaat.
2. Probeer de vraag een praktische context te geven.
3. Elimineer een vraag die niet binnen enkele minuten door uw collega's beantwoord wordt.

Onbetwiste kennis

4. De inhoud van de vraag en het juiste antwoord mogen niet discutabel zijn.
5. Vermijd op zichzelf staande getallen en percentages.
6. Vermijd normatieve begrippen, zoals moeten, dienen, belangrijk, zinloos, wenselijk.
7. Vermijd semikwantitatieve begrippen, zoals soms, zelden, veel, groot, langzaam, kort, kenmerkend, typisch, karakteristiek, specifiek.
8. Vermijd vage begrippen, zoals kunnen, spelen een rol.
9. Vermijd absolute formuleringen, zoals altijd, nooit, alleen, even vaak, identiek.
10. Ga na of het gevraagde altijd opgaat, zonder (denkbare) uitzonderingen.

Algemene richtlijnen

11. Vermijd ontkenningen, zoals geen, niet, on-.
12. Maak korte zinnen, vermijd of verklaar 'moeilijke' woorden, vermijd bijzinnen.
13. Vermijd strikvragen: Vragen waar ook de goede student snel intrapt.
14. Zorg dat meer dan gezond verstand nodig is om de vraag goed te beantwoorden.
15. Maak vergelijkingen: 'Ligt dicht bij de 10% dan 40%' verdient de voorkeur boven 'komt vaker voor dan'.
16. Zorg ervoor dat de items die in de toets worden opgenomen, onafhankelijk van elkaar zijn.

7. Referenties

- 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.
- Cluitmans, J.J. (2002). *Aan de slag met competenties: competentiegericht leren in MBO en HBO*. Nuenen.
- Cohen-Schotanus, J. (1995). Studieduur en de kwaliteit van tentamens. *Onderzoek van Onderwijs*, 24, 26-28.
- Cohen-Schotanus, J. (1995). De praktijk van de compensatie. *Onderzoek van Onderwijs*, 24, 60-62.
- Cohen-Schotanus, J. (2010). *Tegenintuïtief*. Rijksuniversiteit Groningen en Universitair Medisch Centrum Groningen.
- Duthler Associates (2011). *Handreiking voor examencommissies*. Den Haag.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- HBO-raad (2008). *Werken aan kwaliteit*. Den Haag.
- HBO-raad (2011). *Geslaagd. Handreiking examencommissie*. Den Haag.
- INK (2006). *Presteren in het onderwijs. Van ambitie naar realisatie*. Zaltbommel.
- Inspectie van het Onderwijs (2009). *Boekhouder of wakend oog*. Den Haag.
- ISO (2009). *Toetsing: een vak apart*. Utrecht: Interstedelijk studenten overleg.
- Jaspers, M. (2003). *Toetsing, een strategische keuze. Kadernotitie Fontys visie en beleid ten aanzien van toetsing*. Eindhoven: Fontys Hogescholen.
- Jaspers, M. & Heijmen-Versteegen, I. (2005). *Toetswijzer competentiegericht begeleiden & beoordelen in het Hoger Onderwijs*. Eindhoven: Fontys Hogescholen.
- Jaspers, M. & Schade, M. (2002). *Toets&Beleid, organisatiegestuurd toetsbeleid en geautomatiseerde toetsing*. Eindhoven: Fontys Hogescholen.

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

Joosten-ten Brinke, D. & Sluijsmans, D. M. A. (2010). *De kwaliteitspiramide voor eigentijds toetsen en beoordelen*. Tilburg/Nijmegen: Fontys Lerarenopleiding Tilburg/Hogeschool van Arnhem en Nijmegen.

Linn, R.L. (1990). Admissions testing: Recommended uses, validity, differential prediction, and coaching. *Applied measurement in Education*. 3 (4), 297-318.

Onderwijsraad

- Examinering in het hoger onderwijs, 2004
- Examinering: draagvlak en toegankelijkheid, 2006
- Advisering over examinering; een stand van zaken, 2008
- Een diploma van waarde, 2010.

Rekvelde, I.J. & Starren J. (1994). Een examenregeling zonder compensatie in het Nederlandse Hoger Onderwijs? Een vergelijking tussen compensatie en conjunctie. *Tijdschrift voor Hoger Onderwijs*, 12, 4, p. 210-219.

Senge, P. (2007). *Lerende Scholen*, Academic Service, Den Haag.

Sluijsmans, D. M. A. (2008). *Betrokken bij beoordelen* (Lectorale rede). Nijmegen: Hogeschool Arnhem en Nijmegen.

Sluijsmans, D. M. A. & Joosten-ten Brinke, D. (2010). *Bewust en bekwaam toetsen. De effecten van ontwikkelingsgerichte zelfbeoordeling door lerarenopleiders op hun toetspercepties, toetskennis en toetskunde*. Projectvoorstel. Nijmegen/Tilburg: HAN/FLOT.

Van Berkel, A. E. & Bax, H. J. M. (2006). *Toetsen in het hoger onderwijs*. Houten/Diegem: Bohn Stafleu Van Loghum.

Van der Bij, H. & Broekhuis, M. & Gieskes, J. F. B. (2004). *Kwaliteitsmanagement in beweging. Van blauwdruk naar contingenties en dynamiek*. Deventer: Kluwer.