

Het toetsproces ontleed

Henk van Berkel, Anneke Bax en Desirée Joosten-ten Brinke

2.1 Inleiding – 8

2.2 Fasen in het toetsproces – 8

- 2.2.1 Stap 1: Waarom beoordelen en toetsen: het formuleren van de bedoeling(en) – 11
- 2.2.2 Stap 2: Het formuleren en concretiseren van de doelen – 12
- 2.2.3 Stap 3: Het kiezen van de toetsvorm – 14
- 2.2.4 Stap 4: Het construeren en vooraf controleren van toetsvragen of toetstaken – 14
- 2.2.5 Stap 5: Het samenstellen van de toets – 17
- 2.2.6 Stap 6: Het afnemen van de toets – 17
- 2.2.7 Stap 7: Het nakijken van de toets, inclusief een kwaliteitscontrole – 18
- 2.2.8 Stappen 8 en 9: Zak/slaaggrens en cijfers – 19
- 2.2.9 Stappen 10 en 11: Beslissen en rapporteren – 19

2.1 Inleiding

- Uit welke fasen en stappen bestaat het toetsproces?
- Wat zijn de doelen van toetsen?
- Wat zijn doelstellingen?
- Hoe kunnen doelstellingen nader worden geconcretiseerd?
- Wat is een specificatietabel?
- Welke toetsvormen zijn er?
- Wat zijn inspiratiebronnen voor het construeren van toetsvragen?

Het afnemen van toetsen impliceert meer dan het construeren van toetsvragen; het is een proces dat verschillende fasen en stappen kent. Elk toetsproces is in essentie te herleiden tot een viertal fasen, waarin steeds een ander aspect op de voorgrond staat (zie ook  fig. 2.1):

Toetsproces = Doelbepaling + Meten + Cijfergeven + Beslissen

Deze ontleding van het toetsproces geldt voor alle toetsen met een summatief doel. De grenzen tussen de fasen zijn niet precies te trekken, omdat bij het concretiseren van een fase rekening wordt gehouden met de andere fase – bewust of onbewust. Maar deze vierdeling is als analytisch denkkader goed bruikbaar. Elke fase is onder te verdelen in stappen die in de volgende paragrafen worden uitgewerkt.

Die verschillende fasen, en stappen, dragen alle bij aan de kwaliteit van het toetsproces. De bepaling van de kwaliteit van een toets is moeilijk (►H. 1 en 3). De kwaliteit laat zich niet in een enkel getal uitdrukken omdat het toetsproces en de toets zelf veel onderdelen beslaan die alle hun steentje kunnen bijdragen aan de kwaliteit van de toets. En die onderdelen laten zich niet, net als appels met peren, optellen. Downings indeling telt twaalf stappen (Downing 2006) en die van de Vereniging Hogescholen zeven. Misschien dat het op het eerste gezicht lijkt dat er geen overeenstemming is tussen de verschillende deskundigen, maar dat is slechts schijn. De verschillen in aantal stappen wordt veroorzaakt doordat de auteurs onderdelen van het toetsproces onderbrengen in een andere stap zonder te kort te doen aan het geheel.

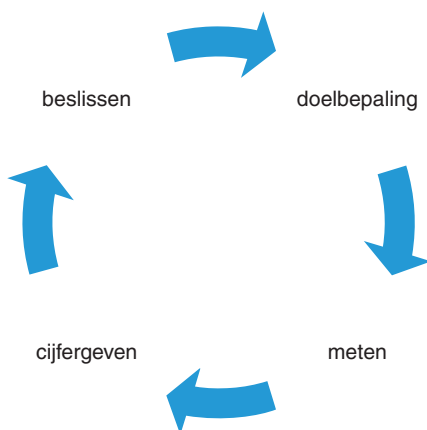
2.2 Fasen in het toetsproces

■ Doelbepaling

In de fase, waarin het doel van de toets wordt bepaald, gaat het om:

- formuleren van de bedoeling(en) van de toets;
- formuleren van de doelstellingen;
- concretiseren van de doelen;
- opstellen specificatietabel.

De doelbepalingfase hoeft niet bij iedere toets te worden doorlopen, omdat de doelstellingen van een toets meestal voor een aantal jaren vastliggen. Bovendien is niet



■ **Figuur 2.1** Fasen in het toetsproces

iedere docent zelf verantwoordelijk voor de doelen. Het is een op opleidingsniveau gedeelde verantwoordelijkheid. In dit hoofdstuk worden het formuleren en concretiseren van de doelstellingen uitgebreid beschreven.

■ Meten

De meetfase omvat alle stappen die te maken hebben met het vormgeven en afnemen van een toets, en wel in een vaste volgorde:

- kiezen van de toetsvorm op basis van de doelen;
- construeren van de toetstaken/toetsvragen;
- vooraf controleren van de toetstaken/toetsvragen;
- samenstellen van de toets;
- afnemen van de toets;
- nakijken van de toets;
- kwaliteitscontrole.

Deze stappen zijn op vrijwel alle toetsvormen toepasbaar. Bij een schriftelijke toets gaat het bijvoorbeeld om het construeren van toetsvragen, terwijl het bij een performance assessment gaat om het construeren van toetstaken. Beide moeten voldoen aan vergelijkbare eisen. De combinatie van toetsvragen/toetstaken vormt de toets. De stappen staan in de volgende paragraaf verder uitgewerkt; aan de kwaliteitscontrole van de toets is een afzonderlijk hoofdstuk gewijd (►H.3).

■ Cijfergeven

De derde fase betreft het waarderingsaspect:

- bepalen van de zak/slaaggrens;
- toekennen van cijfers aan de scores.

De fase van het cijfergeven is wellicht voor studenten de belangrijkste, in ieder geval de spannendste. Maar ook voor docenten is deze fase belangrijk. Het bepalen van de

zak/slaaggrens heeft grote gevolgen voor de studenten. In deze fase moeten docenten op basis van een score bepalen welk cijfer een student krijgt. Het geven van cijfers lijkt simpel, maar vergt grote aandacht van docenten. ► Hoofdstuk 7 gaat hier uitvoerig op in.

■ Beslissen

In de laatste fase worden beslissingen genomen op basis van de in de voorafgaande fase toegekende cijfers:

- nemen en verantwoorden van beslissingen op grond van de toegekende cijfers;
- rapporteren van de toetsuitslag.

In de meest eenvoudige vorm bestaat de toetsuitslag uit niet meer dan het cijfer plus de consequentie voor de studenten, meestal gezakt of (voorwaardelijk) geslaagd, zonder verdere toelichting. Een cijfer alleen geeft echter geen informatie over wat er al wel of nog niet wordt beheerst. Het toevoegen van deze extra informatie maakt een toetsuitslag ook geschikt als feedbackinstrument voor zowel studenten als docenten. Bijvoorbeeld: het aangeven van het hoofdstuk waarop de vraag betrekking heeft of het achterhalen van de antwoorden op extreem moeilijke vragen. Rapportage aan de opleidings- c.q. examencommissie over de cijfers per student en over de kwaliteit van de toets (met behulp van een toetsanalyse na afname) is nodig. In ► H. 3 staat vermeld welke gegevens van belang zijn voor de kwaliteitscontrole en hoe deze worden verkregen.

Hierna staat een overzicht van de fasen in het toetsproces met bij elke fase de stappen die in de volgende paragrafen verder worden uitgelegd.

Fasen en stappen in het toetsproces

Doelbepaling

1. waarom beoordelen en toetsen: het formuleren van de bedoeling(en)
2. het formuleren en concretiseren van de doelen

Metten

3. het kiezen van de toetsvorm
4. het construeren en vooraf controleren van toetsvragen of toetstaken
5. het samenstellen van de toets
6. het afnemen van de toets
7. het nakijken van de toets, inclusief een kwaliteitscontrole

Cijfergeven

8. het bepalen van de zak/slaaggrens
9. het toekennen van cijfers aan de scores

Beslissen

10. het kunnen verantwoorden van de beslissing
11. het rapporteren van de toetsuitslag voor nabespreking, evaluatie en bijstelling

■ **Tabel 2.1** Tijdvak, doel en functie van toetsen

tijdvak	doel	functie
vóór de start	selectie	toelating
tijdens	leermiddel	sturing
tijdens	voortgangscontrol	bijsturing
aan het eind	kwalificatie	niveaubepaling

2.2.1 Stap 1: Waarom beoordelen en toetsen: het formuleren van de bedoeling(en)

Algemeen gesteld: er zijn toetsen met een summatief doel, die leiden tot beslissingen over studenten – bijvoorbeeld zakken of slagen, toelaten of afwijzen – en er zijn toetsen met een formatief doel – bijvoorbeeld om aan te geven over welk onderwerp de student nog te weinig weet. Dit hoofdstuk gaat over toetsen met een summatieve functie, ►H. 8 gaat over toetsen met een formatieve functie.

Er zijn globaal gesproken drie tijdvakken waarin uiteenlopende beslissingen over studenten worden genomen (■tab. 2.1).

■ Toetsen als selectie-instrument

Wanneer, voorafgaand aan het doorlopen van het onderwijsprogramma, moet worden nagegaan of een student geschikt is voor de opleiding, heeft de toets (ingangs-)selectie tot doel. Deze functie van toetsen valt buiten het kader van dit boek.

■ Toetsen als leermiddel en voortgangscontrol

Toetsen kunnen tijdens de onderwijsperiode worden gebruikt om studenten en docenten inzicht te geven in de mate waarin studenten zijn gevorderd in de bestudering van het studieonderdeel of in de mate van verkregen competentie. Wanneer toetsen op een zorgvuldige wijze in het onderwijs zijn geprogrammeerd, vormen ze een middel om studenten te activeren en hun leerproces te sturen. Toetsen die bij wijze van voortgangscontrol zijn afgenomen, beogen het leren te sturen en hebben een formatieve functie. Het belang van dit soort toetsen is de *feedback*functie. De studenten kunnen aan de hand van de toetsuitslag en de feedback bepalen waar hun hiaten in de leerstof zitten en welke vaardigheden zij onvoldoende beheersen. Hiertoe is het uiteraard wel nodig dat de feedback voldoende specifiek en gedetailleerd is, zie ►H. 8.

■ Toetsen als kwalificatie-instrument

De meeste opleidingen hanteren toetsen aan het eind van een onderwijsperiode om na te gaan of studenten het studieonderdeel voldoende beheersen. Deze functie van toetsen wordt aangeduid met de *summatieve* functie. Een studietoets levert doorgaans een score op die aangeeft in welke mate studenten ‘iets’ beheersen. Met andere

woorden, de score weerspiegelt het *niveau* van hun kennis en vaardigheden. De hoogte van het niveau bepaalt de uitkomst van beslissingen: slagen of zakken, voldoende of onvoldoende, toelaten of weigeren enzovoort. Een voldoende score impliceert dat de betrokken studenten voldoen aan de eisen. In dat geval verkrijgen studenten daarvan een officieel bewijs en zijn zij gekwalificeerd. Een summatieve toets is een proeve van bekwaamheid en sluit als het ware iets af.

2.2.2 Stap 2: Het formuleren en concretiseren van de doelen

Het formuleren van doelstellingen is belangrijk om te voorkomen dat onderwijs een kwestie van willekeur gaat worden. In dit boek is het begrip ‘doelstellingen’ – of kortweg ‘doelen’ – gekoppeld aan het gedeelte van de opleiding, het studieonderdeel, dat wordt getoetst. De definitie van doelstellingen aan de hand waarvan docenten een toets maken, luidt:

- » Omschreven kwaliteiten op het gebied van kennis, inzicht en vaardigheden en in voorkomende gevallen (beroeps)houding, waarover degene die het studieonderdeel voltooit, dient te beschikken.

Doelstellingen vormen het startpunt om een toets te construeren.

■ Categorisering en classificatie van doelen

Doelstellingen geven aan wat studenten van het onderwijs moeten gaan opsteken. Een voorbeeld is feitenkennis die vrijwel uitsluitend een beroep doet op het reproduceren van kennis. Maar van studenten mag ook worden verwacht dat ze deze feitenkennis in bepaalde situaties kunnen toepassen of er anderszins mee kunnen omgaan. Deze doelen, in de praktijk aangeduid met reproducerende kennis, inzicht en toepassen, vallen onder de noemer *cognitieve* doelstellingen. Daarnaast zijn opleidingen ook gericht op het bijbrengen van vaardigheden en beroepshoudingen.

De classificatie van Bloom (1972) is nog steeds de bekendste en meest gehanteerde. Hij verdeelt doelstellingen in drie (hoofd)categorieën:

1. Cognitieve doelstellingen: alles wat te maken heeft met verstandelijk functioneren.
2. Affectieve doelstellingen: doelen die te maken hebben met het ontvangen en reageren op signalen.
3. Psychomotorische doelstellingen: doelen gericht op perceptie en het handelen.

Iedere categorie is vervolgens verder onderverdeeld aan de hand van een nadere classificatie, ook wel aangeduid met *taxonomie*. Voor cognitieve doelstellingen hanteert Bloom de volgende classificatie:

- feitenkennis;
- begrip;
- toepassen;

- analyseren;
- synthetiseren;
- evalueren.

Andere veelgebruikte taxonomieën zijn die van Anderson et al. (2001; dit is een herziening van Bloom en Romiszowski (1999)).

De rangordening is niet willekeurig. Het doorlopen van de verschillende klassen vergt een steeds hogere intellectuele activiteit. Een dergelijke onderverdeling is daarom een belangrijk hulpmiddel bij het inhoud geven aan de toets, zodat de vragen in de toets een beroep doen op een breed scala aan intellectuele vaardigheden, en niet bijvoorbeeld uitsluitend op het reproduceren van kennis. In [tab. 11.2](#) in [H. 11](#) staan per niveau voorbeelden van werkwoorden die kunnen worden gehanteerd bij het stellen van vragen van een bepaald niveau.

■ De specificatietabel

Om ervoor te zorgen dat een toets een goede spreiding van vragen heeft over onderwerpen en gewenst niveau, is het hanteren van een specificatietabel, ook wel aangeduid met toetsmatrijs of blauwdruk, nuttig. De simpelste vorm is een lijst met onderwerpen die in de toets aan de orde moeten komen. Deze onderwerpen zijn doorgaans uit de doelen af te leiden. Een inzichtelijke vertaling van de doelen naar onderwerpen maakt duidelijk of alle doelen inderdaad in de toets aan de orde komen. Het is een vereiste de blauwdruk te maken voordat de constructie van vragen begint en de geconstrueerde vragen vervolgens in te passen in de blauwdruk. Dan pas kan de toets representatief en evenwichtig zijn.

Een toets moet representatief zijn ten opzichte van de doelen van het studieonderdeel. Immers met behulp van de toetsuitslag worden belangrijke uitspraken gedaan over het studieonderdeel. Niet alleen ten aanzien van zakken en slagen, maar ook over waar precies de hiaten van studenten zitten, zodat zij zich gericht kunnen bijspijkeren. De toets moet dus voldoende representatief zijn om er valide gevolgtrekkingen uit te kunnen maken. Een term die synoniem is aan representativiteit is *inhoudsvaliditeit*.

Een specificatietabel dient er verder voor te zorgen dat in de uiteindelijke toets, naast een spreiding van de vragen over belangrijke onderwerpen, ook op alle beoogde (cognitieve) niveaus wordt getoetst. Dit hangt nauw samen met de doelen. Daarom is het nodig dat docenten per onderwerp aangeven op welk niveau (of welke niveaus) zij vragen over het betreffende onderwerp gaan stellen.

Het voorbeeld in [tab. 2.2](#) gaat uit van de situatie dat de toets tien onderwerpen bevraagt op drie cognitieve niveaus: kennisreproductie, begrip en toepassing. De gehele toets bestaat uit vijftig vragen.

■ **Tabel 2.2** Voorbeeld van een specificatietabel

aantal vragen per onderwerp	kennisreproductie	inzicht	toepassen
overzicht van onderwijssystemen	2	1	2
plaats van onderwijs in de maatschappij	0	3	2
onderwijsvormen in relatie tot: motivatie, leereffect, inzet, betrokkenheid, eigen verantwoordelijkheid, zelfstandigheid	3	3	4
rol van motivatie	2	2	3
geheugen	2	0	1
scholing en diploma's	2	0	2
duaal leren	1	2	2
attitudes	1	1	1
evalueren	4	0	4

2.2.3 Stap 3: Het kiezen van de toetsvorm

De keuze voor een toetsvorm dient aan te sluiten bij de doelstellingen. Met behulp van een toets moet de beoogde leeropbrengst kunnen worden gemeten. Zo is het bijvoorbeeld niet geloofwaardig om met een meerkeuzevraag te toetsen of een student een onderzoek kan presenteren. Daar zijn andere toetsvormen voor, zoals een performance assessment (►H. 15).

De schriftelijke toetsvormen met open of gesloten vragen komen (nog steeds) het meest voor. Onder open toetsvormen vallen essay-, korte-antwoord- en aanvultoeetsen (►H. 12), maar ook mondeling toetsen (►H. 14). Gesloten toetsvormen zijn: ja/nee-, juist/onjuist-, meerkeuze- en matchingtoetsen (Dousma et al. 1997, zie ook ►H. 11). Bij open toetsvragen moeten studenten zelf een antwoord formuleren ('vrije antwoorden'), terwijl zij bij gesloten toetsvragen een van de alternatieven moeten aankruisen ('geprecodeerd'). Deze toetsen kunnen worden gekoppeld aan casusmateriaal. Bij medische opleidingen is een 'casus' dan een patiënt of een beschrijving ervan, en de toetsen vinden plaats met gestructureerde open vragen (►H. 13).

Voor een uitvoerige beschrijving van de verschillende toetsvormen met hun voor- en nadelen, wordt verwezen naar de betreffende hoofdstukken in dit boek.

2.2.4 Stap 4: Het construeren en vooraf controleren van toetsvragen of toetstaken

■ Het construeren van toetsvragen: inspiratiebronnen

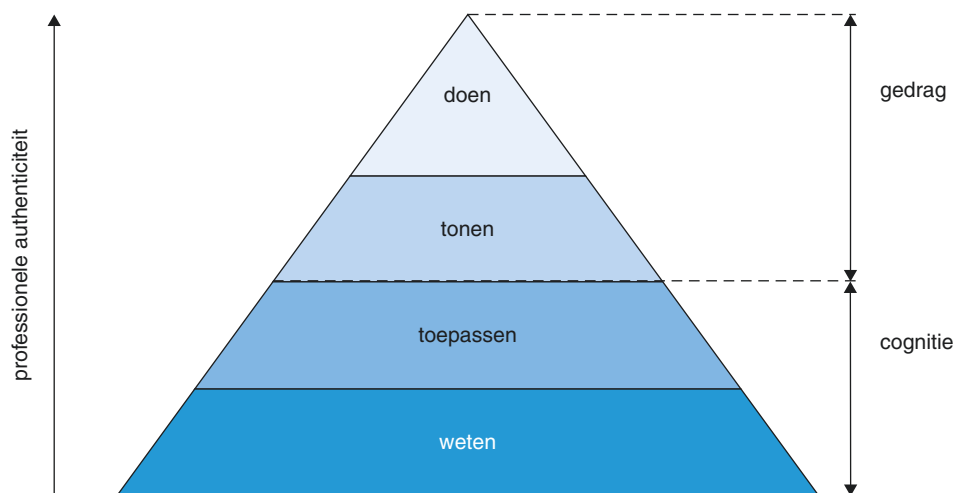
Op basis van de gekozen toetsvorm moeten vragen/toetstaken worden geconstrueerd. Omdat veel docenten huiverig zijn om reeds gebruikte toetsvragen in de toets op te nemen, is het construeren van vragen een steeds terugkerende grote klus. Het

Tabel 2.3 Inspiratiebronnen voor het construeren van vragen

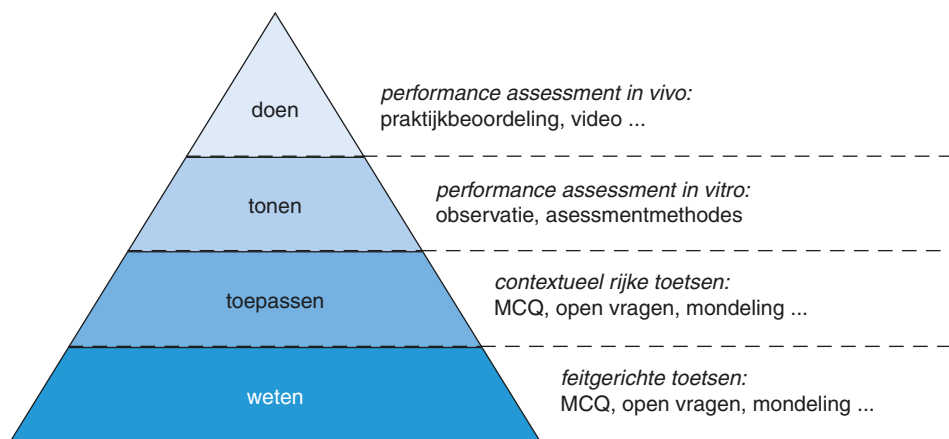
Doelstellingen	Toetsen dienen na te gaan in hoeverre de student de doelen beheerst en moeten dus een inhoudelijke relatie hebben met de doelen. Nu zijn deze vaak globaal van aard. Nodig zijn daarom concretisering in de vorm van nadere beschrijvingen van wat studenten moeten kennen en kunnen nadat zij het studieonderdeel hebben doorlopen. Een toetsvraag is in feite een ultieme operationalisatie van een doelstelling
leerstof	De leerstof dient een afspiegeling te zijn van de doelen; daarmee is de leerstof een goede inspiratiebron
vervolgopleiding	De inhoud van eventuele vervolgopleidingen kan ideeën opleveren voor het maken van toetsvragen. Het gaat om niet meer dan ideeën. Het kan niet de bedoeling zijn dat in de vooropleiding datgene wordt getoetst wat in de vervolgopleiding aan de orde komt. Primair uitgangspunt vormen de doelen van het 'eigen' onderwijs
werkveld	De meeste opleidingen zullen zich richten op een bepaalde beroepsuitoefening. De werkzaamheden in het toekomstige beroep kunnen ideeën opleveren voor het maken van vragen, mits men uiteraard die facetten uit het beroep neemt die in het te toetsen studieonderdeel worden behandeld. Ook hier moet men voorzichtig zijn. In het bijzonder moet men letten op het beheersingsniveau waarop de toets zich richt
analyse van vaardigheid	Er zijn onderwijssituaties waarin, naast een theoriemodule, practica of vaardigheidstrainingen zijn geprogrammeerd die inhoudelijke overeenkomst hebben met de theoriemodule. Bijvoorbeeld: binnen een verpleegkundige opleiding bestaat naast de theoriemodule 'Vervoer in het lichaam' de training 'Bloeddrukmeten'. Vragen over de kennis die studenten nodig hebben om bloeddruk te kunnen meten, zijn in de kennistoets op hun plaats
antwoorden op gebruikte vragen	De uiteenlopende antwoorden van studenten op open toetsvragen geven inspiratie voor nieuwe vragen of voor het formuleren van afleiders (dit zijn de niet-correcte alternatieven). Het komt regelmatig voor dat een studieonderdeel wordt getoetst met een combinatie toets, bestaande uit een gesloten en een open deel. Bij het construeren van de meerkeuzevragen is het zoeken naar aantrekkelijke afleiders vaak een probleem (► H. 11). Inspiratie voor het formuleren van een afleider kan men dan halen uit de (foute) antwoorden op open toetsvragen van een vergelijkbare toets
studenten	De docent kan discussies tussen studenten tijdens het onderwijs beluisteren waarin studenten de nadruk leggen bij de bestudering van de leerstof. Kennelijk zijn dat aanstekelijke en voor studenten relevante onderwerpen geweest. Indien docenten hiervan korte aantekeningen maken, kunnen de notities later worden omgezet in toetsvragen. De mogelijke tegenwerping 'maar omdat ze er zo veel over gediscussieerd hebben, weten ze allemaal het antwoord!'; is niet ter zake doende. Er is toch niets mis mee wanneer studenten bijna allemaal de vraag correct beantwoorden indien de vraag betrekking heeft op een relevant onderwerp? Kennelijk is het onderwijs op dit punt goed aangeslagen
tussentijdse gedachten	Gedurende de periode dat het onderwijs wordt voorbereid en gegeven, komen dikwijls gedachten op bij de docent over mogelijke toetsvragen. Het is nuttig deze gedachten meteen op te schrijven zodat ze later kunnen worden omgezet in toetsvragen

is denkbaar dat de inspiratie hiervoor na verloop van tijd opraakt. Toch bestaan er meestal meer bronnen dan verwacht waaruit docenten kunnen kiezen (► tab. 2.3).

De beste methode voor het construeren van toetsvragen is uit te gaan van de doelstellingen en de beheersingsniveaus zoals ze vastgelegd zijn in de specificatietabel en die



■ **Figuur 2.2** De assessmentpiramide (Miller 1990)



■ **Figuur 2.3** Assessmentpiramide en toetsvormen

vervolgens te concretiseren in de vorm van toetsvragen. Op die wijze geconstrueerde toetsvragen dragen de beste garantie in zich dat de toets representatief is voor de doelen, en dus inhoudsvalide is.

Vooraf in opleidingen waar vaardigheden belangrijk zijn, is de zogenoemde piramide van Miller (1990) handig bij het formuleren van toetsvragen (■ fig. 2.2). In de piramide symboliseren de lagen de complexiteit van competenties. Weten is noodzakelijk om te kunnen toepassen. Toepassen is weer noodzakelijk om iets te kunnen tonen. Om een verrichting in praktijk daadwerkelijk uit te kunnen uitvoeren, moet je als student de verrichting eerst hebben getoond.

In de tweede piramide staan de toetsvormen toegevoegd aan de verschillende niveaus (■ fig. 2.3). Het onderste niveau, weten, is te meten met relatief eenvoudige toetsen als meerkeuze- en open toetsvormen. De tweede laag, toepassen, vereist wat meer complexe toetsen. Niet zozeer wat de toetsvorm betreft – die kunnen gelijk zijn aan die voor

het toetsen van het laagste niveau. Maar wel dienen de toetsen ‘rijk’ te zijn. Dat wil zeggen, de vragen dienen plaatjes van praktische situaties te bevatten waarover studenten vragen beantwoorden. Bij computerafname kunnen dat ook filmpjes zijn. De hoogste twee niveaus, waar gedragingen moeten worden getoetst, zijn niet schriftelijk te toetsen (►H. 18). Docenten moeten daarin studenten observeren terwijl ze, bijvoorbeeld in de onderwijssituatie, een proef uitvoeren of een gesprek voeren. In de hoogste laag, het doen, gaat het om te toetsen of studenten in een ‘echte’ praktijksituatie, een werkplek (►H. 19), in staat zijn een vaardigheid uit te voeren.

Wanneer de docent eenmaal voldoende inspiratie heeft voor het construeren van toetsvragen, kan worden gestart met de constructie. De verschillende hoofdstukken in dit boek bevatten vele tips over hoe dit in zijn werk gaat. Voor specifieke constructieregels bij het formuleren van open en gesloten vragen wordt verwezen naar de ►H. 11 en 12.

■ Het vooraf controleren van de toetsvragen of toetstaken

Wanneer de toetsvragen zijn geconstrueerd, moet, voorafgaand aan de toetsafname, een indicatie worden verkregen over de kwaliteit van elke vraag door de vragen te controleren. Alle toetsvragen moeten voldoen aan inhoudelijke en vormtechnische eisen.

Over een aantal eigenschappen van een toetsvraag, zoals relevantie, eenduidigheid, objectiviteit en specificiteit, is vooraf een oordeel te geven. Een dergelijke controle kan bijvoorbeeld gebeuren door de vragen aan collega's voor te leggen en te bespreken. Over aspecten als de moeilijkheid en het onderscheidingsvermogen van een vraag is het lastig om vooraf een oordeel te geven – dit moet na afname worden gecontroleerd – zie onder stap 7 (en ►H. 3).

2.2.5 Stap 5: Het samenstellen van de toets

Een toets is meer dan een verzameling vragen. De toets als geheel zal inhoudsvalide moeten zijn. Dat wil zeggen: de toets moet meten wat men wil meten. Om dit te bereiken is het in ieder geval noodzakelijk dat de doelstellingen evenwichtig in de toets aan de orde komen. Wanneer bijvoorbeeld ‘het verkrijgen van inzicht in ...’ een belangrijk doel is, dan is een toets die slechts bestaat uit vragen die een beroep doen op kennisreproductie beslist niet representatief. Het belangrijkste hulpmiddel om een representatieve toets te verkrijgen is het vooraf opstellen van een specificatietabel (stap 2). Met behulp van een specificatietabel kunnen docenten de toetsvragen verdelen over de belangrijkste onderwerpen van het studieonderdeel en over het gewenste niveau (kennisreproductie, toepassing, inzicht/begrip).

Tevens moet bij het samenstellen van de toets ook worden gedacht aan de studentinstructie, beoordelingscriteria en een antwoord- en scoremodel (bij open vragen).

2.2.6 Stap 6: Het afnemen van de toets

De condities waaronder een docent een toets afneemt, zoals de zaalruimte of temperatuur, kunnen een belangrijke invloed hebben op het toetsresultaat. De studenten moeten in elk geval voldoende tijd krijgen om de toets af te leggen – in Nederland is het niet gebruikelijk een toets het karakter van een *speed test* te geven. Verder moet het de

2 studenten vooraf bekend en duidelijk zijn wat zij precies kunnen verwachten. Wat zijn de onderwerpen van de vragen? Welke vragen zijn het belangrijkste en krijgen een hoger gewicht bij de beoordeling? Wat is de vraagvorm? Wanneer komt de uitslag? Ten behoeve van de transparantie is het goed om studenten dit soort informatie op voorhand te geven, bijvoorbeeld in de studiegids of in de vorm van een toetsinstructie. Aan het eind van dit hoofdstuk staat een uitgebreide toetsinstructie als voorbeeld.

2.2.7 Stap 7: Het nakijken van de toets, inclusief een kwaliteitscontrole

■ Het nakijken van de toets

Het nakijken van de toets hoeft bij gesloten toetsen weinig tijd te kosten als dit per computer gebeurt. Met behulp van de antwoordsleutel is de uitslag in korte tijd bekend. Voordat de uitslag bekend wordt gemaakt, moet duidelijk zijn dat de vragen kwalitatief goed zijn. Het nakijken van de vragen uit een open toets is veel tijdrovender omdat dit meestal handwerk is (►H. 10 voor het digitaal nakijken van open vragen). Het hanteren van een antwoordmodel bij het nakijken is onontbeerlijk (►H. 12). Voor het beoordelen van de toetstaken bij vaardigheidstoetsen of bij het uitvoeren van opdrachten dienen beoordelingsmodellen beschikbaar te zijn. Rubrics zijn hiervoor ook zeer geschikt (►H. 16).

■ Kwaliteitscontrole

Achteraf kan blijken dat er – ondanks alle zorgvuldigheid – kwalitatief toch iets is aan te merken op bepaalde vragen, toetstaken en/of de toets als geheel. Een groot voordeel van het computertijdperk is dat men op een betrekkelijk eenvoudige en snelle wijze inzicht verkrijgt in een aantal kwaliteitsaspecten van de vragen en van de toets als geheel: de moeilijkheid, het onderscheidingsvermogen van de vraag (beantwoordt de groep studenten die de stof beheerst de vraag correcter dan de groep die de stof minder beheerst?), de (on)aantrekkelijkheid van de afleiders en de betrouwbaarheid van de toets. Dit soort informatie is belangrijk op het moment dat een docent de toetsuitslag interpreteert. Wanneer de toets bijvoorbeeld zeer onbetrouwbaar is geweest (►H. 3 voor de normen hiervoor), is het niet eerlijk tegenover studenten om hun een cijfer te geven. Immers, een zeer onbetrouwbare toets betekent niets anders dan dat men net zo goed met een dobbelsteen de uitslag had kunnen bepalen. Bij het beoordelen van werkstukken kan de kwaliteitscontrole uitwijzen dat bepaalde docenten soepeler beoordelen dan anderen of dat beoordelingscriteria niet helder genoeg geweest zijn.

Ook kan de kwaliteitsanalyse docenten op het spoor zetten van inhoudelijke tekortkomingen van de vraag. Wanneer blijkt dat bijvoorbeeld slechts weinig studenten een bepaalde vraag correct beantwoorden, kan dit zijn veroorzaakt door de slechte voorbereiding van studenten, maar ook door de (toevallig) moeilijk uitgevallen vraag of door een onduidelijke vraagstelling. Als de analyse aantoont dat veel studenten een ander alternatief dan het correcte hebben gekozen, kan – bij nadere bestudering van de vraaginhoud – blijken dat dit alternatief óók correct is.

2.2.8 Stappen 8 en 9: Zak/slaaggrens en cijfers

Het bepalen van de zak/slaaggrens (de cesuur), het toekennen van cijfers en het geven van waarderingen aan de cijfers zijn sterk met elkaar verbonden. Rapporten en cijferlijsten drukken de waardering uit die de docenten hebben voor de leerprestaties van de studenten. Het waarderingscijfer voor prestaties wordt ontleend aan een cijferschaal die, althans theoretisch., meestal loopt van 0 of 1 tot 10. Voor iedere stap op deze schaal bestaat een verbale kwalificatie die in alle gevallen een waarderende betekenis heeft. Zo is een 8 'goed' en een 5 'bijna voldoende'. Dergelijke kwalificaties impliceren dat er ergens grenzen zijn getrokken. Dat is echter niet eenvoudig omdat er geen hanteerbare methode is die in elke situatie toepasbaar is. Daarom resteert docenten niets anders dan uit het beschikbare aanbod die methode te kiezen die het beste is in de gegeven situatie. Daarvoor zijn wel enkele criteria te geven (►H. 7).

2.2.9 Stappen 10 en 11: Beslissen en rapporteren

Beslissingen binnen het onderwijs hebben een zeer bijzondere plaats gekregen. De telkens terugkerende spanning in een examenperiode maakt dit duidelijk. Vreemd is dit overigens niet, want het zijn juist deze beslissingen die een belangrijke betekenis hebben voor de toekomst van studenten. De verlenging van een studieperiode met een jaar of meer, het niet toegankelijk zijn van de beoogde vervolgopleiding, ingedeeld worden bij een andere groep, in de knel raken met de nog beschikbare studiefinanciering, een streep moeten zetten door uitgestippelde toekomstplannen enzovoort, zijn voorbeelden van belangrijke gevolgen van een toetsuitslag. Hoewel beslissingen die worden genomen op basis van hoge of lage cijfers onbetwistbaar zullen zijn, is het verschil tussen de nog net positieve en de nog net negatieve beslissing bijna altijd heel erg klein. Alleen al op die grond moet een zo groot mogelijke zorgvuldigheid met betrekking tot het nemen van beslissingen worden bepleit (Wijnen 1993). Een voorbeeld van een zorgvuldige besluitvorming zijn de antwoorden van studenten (of hun paper) die net niet of net wel voldoende zijn, ten tweede male te beoordelen, nu door een andere dan de eerstverantwoordelijke docent.

■ Het rapporteren van de toetsuitslag

De toetsuitslag kan bestaan uit het behaalde cijfer plus eventueel de vermelding van de consequentie voor de studenten – bijvoorbeeld gezakt of (voorwaardelijk) geslaagd – en het aantal behaalde studiepunten. Een toetsuitslag kan ook functioneren als een feedbackinstrument wanneer studenten hieruit bijvoorbeeld tevens kunnen opmaken hoe zij hebben gescoord op onderdelen van de toets, op de afzonderlijke vragen of op de clustering hiervan tot bepaalde onderwerpen.

Naast een rapportage aan studenten dient de verantwoordelijke docent ook aan de opleiding c.q. de examencommissie te rapporteren. Uiteraard bestaat deze rapportage uit de cijfers per student ten behoeve van de administratie, maar de rapportage moet tevens enkele meest voor de hand liggende resultaten van de toets bevatten die een

indicatie geven voor de kwaliteit van de toets, zoals: het gemiddelde cijfer, de verdeling van de cijfers, het aantal geslaagden en de betrouwbaarheid. Deze rapportage kan dan ook gebruikt worden om de kwaliteit van de toets met collega's te bespreken en op basis hiervan het ontwerp van de toetsing voor een komende periode aan te passen.

Voorbeeld van een toetsinstructie

Toetsinstructie bij de toets *Inleiding in de Onderwijskunde*.

Zo meteen krijg je de toets met de vragen uitgedeeld. Lees, voordat je begint, het onderstaande zorgvuldig. Je hebt 2,5 uur om de toets te maken; tegen het einde van de toets word je drie keer gewaarschuwd dat het inlevertijdstip nadert.

Invulinstructie

- Omcirkel bij de vijftig meerkeuzevragen het alternatief (A, B of C) dat naar jouw idee het beste antwoord bevat op de vraag. Breng vervolgens je keuze over naar het optisch leesbare formulier. Op dat formulier tref je een kolom aan met cijfers. De cijfers corresponderen met de nummers van de meerkeuzevragen. In een van de drie kolommen achter het cijfer zet je je keuze. Indien je 'A' antwoordt, zet dan een streepje tussen de lijntjes in de eerste kolom, met 'B' antwoord je in de tweede kolom en kies je 'C', dan een streepje in de derde kolom.
- Geef in ieder geval een antwoord; fout gokken leidt *niet* tot strafpunten!
- Zet het streepje met een potlood, hardheid HB. Indien je het antwoord wilt veranderen, gum het dan helemaal weg en plaats het op een andere plaats opnieuw. Mocht je je potlood en gum zijn vergeten, ga er dan *NU* een halen bij de surveillant.
- Het tweede gedeelte van de toets bestaat uit twaalf open vragen. Formuleer je antwoord direct in de ruimte onder de vraag. De vragen zijn zo gesteld dat de ruimte voor het antwoord voldoende is. Schrijf *NIET* in de kantlijn of op de achterzijde.

Scoring en uitslag

- Voor ieder correct antwoord op een meerkeuzevraag krijg je één punt. De twaalf open vragen leveren samen zestig punten op, iedere vraag vijf punten. Je kunt dus in totaal 110 punten behalen.
- De cijfertoekenning houdt rekening met hoe de toets is gemaakt door je medestudenten. Maar je kunt ervan op aan dat als je zestig punten hebt behaald, je in ieder geval een voldoende hebt.
- Indien je schrijft buiten de toegestane ruimte, wordt dit door de beoordelaar niet gelezen. Deze tekst maakt dus geen deel uit van het antwoord.
- Foute beweringen in je antwoord leiden tot puntenaftrek, ook als het goede antwoord er wel staat. De minimale score op de vraag is echter nooit negatief.
- In principe hangt over drie weken de uitslag van de toets op het mededelingenbord bij Bureau Onderwijs.

Algemeen

- Het is *NIET* toegestaan tijdens de toets hulpbronnen te raadplegen als een rekenmachine, woordenboek, aantekeningen, telefoon, iPod, iPad, enzovoort.
- Het is *NIET* toegestaan te communiceren met een medestudent, ook niet via het uitwisselen van potloden, papier, gum en/of pen.
- Vermeld allereerst op *IEDERE* bladzijde van de toets én op het optisch leesbare formulier je naam en studentnummer.
- Controleer de paginanummering van je toetsboekje. Indien een bladzijde ontbreekt, steek dan *NU* je hand op en vraag een nieuw toetsboekje.
- De examencommissie maakt je er op attent dat tijdens deze toets het fraudereglement zoals vermeld in het examenreglement van toepassing is.
- Indien je een vraag wilt stellen, steek dan je hand op.
- De eerste dertig minuten mag je de zaal niet verlaten, ook niet voor het bezoeken van het toilet.
- Indien je het toilet wilt bezoeken, steek dan je hand op en vraag toestemming aan de surveillant.
- Als je klaar bent met de toets, lever je de toets in bij de surveillant, legitimeer je je met je collegekaart, en verlaat de zaal.
- Na afloop van de toets word je verzocht een kort evaluatieformulier buiten de toetszaal in te vullen en in te leveren bij de surveillant.

Uit: Van Berkel (1999)

Toetsen in het hoger onderwijs

van Berkel, H.; Bax, A.; Joosten-ten Brinke, D. (Eds.)

2017, XVIII, 354 p. 29 illus. in color., Softcover

ISBN: 978-90-368-1678-6