

TOOLKIT TOETSEN

Toetscyclus in kaart gebracht



TOOLKIT TOETSEN

Toetscyclus in kaart gebracht



INHOUDSOPGAVE

0	TOETSCYCLUS	7
1a	FORMULEREN VAN LEERDOELEN	9
1b	KIEZEN VAN EEN TOETSVORM	11
2	CONSTRUEREN VAN EEN TOETSMATRIJS	15
3a	CONSTRUEREN VAN EEN KENNISTOETS MET GESLOTEN VRAGEN	17
3b	CONSTRUEREN VAN EEN KENNISTOETS MET OPEN VRAGEN	19
3c	SCREENINGSDOCUMENT KENNISTOETS	21
3d	CONSTRUEREN VAN EEN COMPETENTIETOETS	23
3e	SCREENINGSDOCUMENT COMPETENTIETOETS	25
3f	ONTWIKKELEN VAN RUBRICS	27
3g	OPSTELLEN VAN NORMERING EN CESUUR	31
4a	BETROUWBARE AFNAME EN BEOORDELING	33
4b	HOUDEN VAN EEN KALIBREERSESSIE	35
5	UITVOEREN VAN EEN TOETSANALYSE	37
6	GEVEN VAN FEEDBACK	39
7	REFLECTIE OP TOETSCYCLUS	41
	TOETSVORMEN	43
	BRONNEN	45

0 TOETSCYCLUS

WAT IS DE TOETSCYCLUS?

De toetscyclus is het proces dat doorlopen wordt bij het ontwikkelen van een toets. De toetscyclus bestaat uit 7 fasen (zie figuur 1). In de verschillende fasen worden belangrijke onderdelen behandeld, die bijdragen aan de totstandkoming van een valide, betrouwbare en transparante toets. Alle fasen komen in deze toolkit aan bod. Op deze kaart staat een totaaloverzicht.



HOE IS DE TOETSCYCLUS OPGEBOUWD?

Figuur 1. De zeven fasen van de toetscyclus¹.

<p>1. basisontwerp</p>	<p>Toetsen hangen samen met de leerdoelen die zijn opgesteld en het gegeven onderwijs, maar ook met toetsbeleid van de academie of opleiding en het totale toetsprogramma dat een student gedurende zijn opleiding doorloopt. Deze samenhang wordt constructieve alignment genoemd. In deze fase worden keuzes gemaakt die de basis vormen voor het ontwerp van de toets. De geformuleerde leerdoelen worden onder de loep genomen en leiden tot de keuze voor een passende toetsvorm. Taxonomieën en niveau-indelingen kunnen hierbij helpen.</p>	<p>Kaart 1a gaat dieper in op het formuleren van leerdoelen; kaart 1b op het kiezen van een toetsvorm.</p>
<p>2. construeren toetsmatrijs</p>	<p>Een toetsmatrijs is een tabel waarin wordt aangegeven hoe onderdelen van de toets zijn verdeeld over leerdoelen en op welke (cognitieve) niveaus er wordt getoetst. De zwaarte van een leerdoel bepaalt de weging in de toets. De toetsmatrijs is een soort blauwdruk voor de toets en helpt bij de ontwikkeling ervan. De toetsmatrijs blijft in principe meerdere jaren hetzelfde, waardoor de verdeling van vragen over leerdoelen voor toets, herkansing en toetsen in volgende collegejaren qua inhoud en niveau gelijkwaardig blijven.</p>	<p>Kaart 2 gaat dieper in op hoe je een toetsmatrijs construeert.</p>
<p>3. construeren toets/normeren</p>	<p>In deze fase worden de toetsonderdelen geformuleerd. Deze inhoud moet de leerdoelen en niveaus afdekken zoals in de toetsmatrijs zijn vastgesteld. De toets wordt geconstrueerd en gescreend: vragen voor een kennis-toets en de opdracht voor een competentietoets. Elke toets kent een beoordelingsmodel. Voor kennis-toetsen is dit een antwoordmodel, voor competentietoetsen een beoordelingsformulier. Voor alle typen toetsen geldt dat er gescoord wordt (hoeveel punten krijgt men per onderdeel en hoeveel punten voor de totale toets), dat er een cesuur is (hoeveel van de maximaal te behalen punten moeten worden behaald voor een voldoende) en een normering (welk cijfer bij welk aantal punten). De normering en de cesuur staan vermeld op de toets of op het beoordelingsformulier en zijn hiermee inzichtelijk voor de student.</p>	<p>Kaarten 3a, 3b, 3d en 3f gaan dieper in op hoe je toetsen en beoordelingsmodellen ontwerpt. Met kaarten 3c en 3e kunnen respectievelijk kennis-toetsen en competentietoetsen worden gescreend. Kaart 3g gaat dieper in op hoe je de normering en cesuur van je toets opstelt.</p>

<p>4. afnemen en beoordelen</p>	<p>Wanneer de toets wordt afgenomen, is het belangrijk dat de condities waaronder deze wordt afgenomen zo betrouwbaar mogelijk zijn. Allerlei zaken kunnen negatief van invloed zijn op het zuiver meten van de prestaties van de student, zoals de plek waar de afname plaatsvindt of de tijd die studenten ervoor krijgen. Betrouwbaarheid blijft vervolgens een belangrijk criterium tijdens het beoordelen. Een goed uitgewerkt beoordelingsmodel kan de betrouwbaarheid van de beoordeling vergroten, maar er kan altijd sprake zijn van beoordelaarseffecten die de beoordeling positief of negatief beïnvloeden. Kalibreersessies kunnen helpen om examinatoren op een meer dezelfde wijze tot een beoordeling te laten komen door in gezamenlijkheid het beoordelingsmodel te bespreken en te expliciteren.</p>	<p>Kaart 4a gaat dieper hoe je zorg draagt voor een betrouwbare afname en beoordeling; kaart 4b op hoe je een kalibreersessie houdt.</p>
<p>5. verwerken en analyseren</p>	<p>Na afname van de toets is het zinvol om de resultaten van de toets tegen het licht te houden. Resultaten zoals het slagsingspercentage, gemiddelde cijfer en spreiding van de cijfers kunnen een eerste signaal afgeven om de toets verder te analyseren. Analyse kan aanleiding geven tot het aanpassen van de normering en cesuur, je onderwijs, je toetsopdracht of je beoordelingsmodel.</p>	<p>Kaart 5 gaat dieper op hoe je een toetsanalyse uitvoert.</p>
<p>6. registreren en communiceren</p>	<p>Na de toetsanalyse worden de definitieve toetsresultaten geregistreerd. Deze worden vervolgens gecommuniceerd naar studenten. Slechts het communiceren van het resultaat is geen informatierijke feedback. In deze fase is het belangrijk om leergerichte feedback mee te geven aan de studenten. Op welke onderwerpen scoorden ze goed en op welke minder? Wat hebben ze nog nodig voor het volgende vak? Leergerichte feedback bestaat uit feed-up, feedback en feedforward.</p>	<p>Kaart 6 gaat dieper in op hoe je leergerichte feedback geeft.</p>
<p>7. evalueren en verbeteren</p>	<p>Bij de laatste fase van de cyclus wordt gereflecteerd op de gehele cyclus. Welke fasen gingen goed, wat kon er beter? De conclusies die hieruit worden getrokken, kunnen invloed hebben op allerlei facetten: de toets, afname en beoordeling, samenwerking, leerdoelen, onderwijsactiviteiten, inzage- en feedbackmomenten etc. Dit om de kwaliteit van het gehele onderwijs te verbeteren wanneer de onderwijseenheid waar de toets bij hoorde opnieuw aanvangt.</p>	<p>Kaart 7 gaat dieper in op hoe je je toetscyclus kunt verbeteren.</p>

EXTRA KAARTEN:

- **Toetsvormen:** een overzicht met definities van verschillende toetsvormen. Deze kaart kan dienen als naslagwerk en hoort niet specifiek bij één bepaalde fase van de toetscyclus.
- **Bronnen:** de referenties die voor de toetskaarten zijn gebruikt.

1a FORMULEREN VAN LEERDOELEN

WAT ZIJN LEERDOELEN?

Leerdoelen zijn een beschrijving van welk waarneembaar gedrag studenten na afloop van een vak of onderwijseenheid moeten laten zien. Dit gewenste gedrag komt zoveel mogelijk overeen met de prestaties en producten waar studenten in de beroepspraktijk hun geld mee gaan verdienen. Studenten krijgen geen les om de aangereikte lesstof in hun eigen woorden weer aan de docent te kunnen vertellen. Zij moeten zich bepaalde kennis zo eigen maken, dat zij er in de beroepspraktijk iets mee kunnen doen of maken.

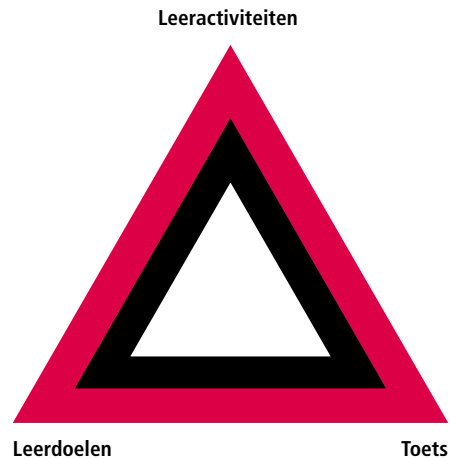
Uit de leerdoelen van een Hbo-opleiding is op te maken wat een beginnend beroepsbeoefenaar moet kunnen. Anders gezegd: wanneer hij 'beroepsbekwaam' is. De beroepsbekwaamheid is verwoord in landelijk vastgestelde eindkwalificaties of (beroeps)competenties. Leerdoelen zijn daarom rechtstreeks af te leiden van deze eindkwalificaties of competenties.

WAAROM ZIJN LEERDOELEN BELANGRIJK?

Goed geformuleerde leerdoelen motiveren studenten en dagen hen uit. Leerdoelen zorgen voor sturing en focus: het is haalbaar en betekenisvol en duidelijk wat er van studenten verwacht wordt. Voor docenten geven leerdoelen richting aan het ontwikkelen van:

- toetsen: waaruit blijkt dat de leerdoelen zijn behaald?
- onderwijs: welke werkvormen en activiteiten helpen om het gewenste gedrag te leren?

Leerdoelen, toets en leeractiviteiten vormen een driehoeksverhouding: op basis van de doelen en het bijbehorende toetsresultaat ontwerp je passende werkvormen en activiteiten, die studenten aanzetten tot gericht leren. Dit wordt 'constructive alignment' genoemd². Om alignment te krijgen is het essentieel dat de leerdoelen worden geformuleerd in termen van waarneembaar gedrag en de beoogde uitkomst van dit gedrag.



Figuur 1. Constructive alignment².

HOE FORMULEER JE LEERDOELEN?

Leerdoelen worden geformuleerd in termen van waarneembaar gedrag: gedrag dat in een bepaalde (beroeps)context leidt tot een gewenst resultaat. In deze formulering staan drie componenten centraal³:

1. Waarneembaar gedrag: wat moet de student laten zien, doen of maken (één centraal werkwoord)?
2. Beroepscontext of relevante inhoud: om welke context of inhoud gaat het?
3. Beoogd resultaat van dit gedrag: waartoe moet de student dit laten zien, doen of maken?

Bij de beantwoording van de vragen focus je eerst op de belangrijkste taken of opdrachten die een beginnend beroepsbeoefenaar (in een bepaalde fase van zijn opleiding) kan krijgen. Van een propedeusestudent verwacht je andere prestaties, dan van een vierdejaars student.

Het leerdoel ziet er met deze componenten als volgt uit:

'De student ... (één centraal werkwoord) ... (inhoud), zodanig dat ... (beoogde resultaat)'.

Bijvoorbeeld:

- De student maakt op basis van een probleemanalyse een onderzoekopzet, zodanig dat het onderzoek uitgevoerd kan worden.
- De student reflecteert op de eigen (Belbin)rol en bijdrage in het team, zodanig dat hij tot sterke punten en verbeteracties komt.
- De student stelt een financiële business case op met diverse scenario's, zodanig dat er in verschillende situaties een onderbouwd ondernemingsbesluit genomen kan worden.

Het beoogde resultaat is 'the next step' van het gewenste gedrag. Afhankelijk van de formulering van het leerdoel wordt dit beoogde resultaat getoetst. In het tweede bovenstaande voorbeeld reflecteert de student en levert hij hierbij zelf een paar sterke punten en verbeteracties aan. In het derde voorbeeld stelt hij een business case op, waardoor een andere persoon een besluit kan nemen.

Het centrale werkwoord in het leerdoel geeft een duidelijke indicatie voor de toetsvorm **(zie kaart 1b)**. Het advies is om per vak of onderwijseenheid maximaal vier leerdoelen te formuleren op een hoger cognitief niveau. In de toetsmatrijs kun je vervolgens aangeven op welk(e) beheersingsniveau(s) getoetst moet worden **(zie kaart 2)**.

In het bovenstaande voorbeeld stelt de student een financiële business case op. Het bijbehorende cognitieve niveau is 'creëren'. Als het relevant is om dit leerdoel bijvoorbeeld ook op de cognitieve niveaus 'begrijpen' en 'analyseren' te toetsen, geef je dit aan in de toetsmatrijs. Je kunt hiervoor verschillende niveau-indelingen of taxonomieën (zoals van Bloom en/of Miller) gebruiken. Je ziet dan dat je voor hetzelfde leerdoel op deze verschillende niveaus een aantal vragen of opdrachten moet formuleren.

1b KIEZEN VAN EEN TOETSVORM

WAT IS EEN TOETSVORM?

Een toetsvorm is de wijze waarop de leerdoelen worden getoetst. Er zijn veel toetsvormen te onderscheiden. Grofweg onderscheiden we kennistoetsen en competentietoetsen.

Om leerdoelen op de cognitieve niveaus te beoordelen, worden kennistoetsen ingezet. De kennis wordt getoetst met open of gesloten vragen en/of door het gebruik van casussen. Open en gesloten vragen hebben beide voor- en nadelen; ga bij de keuze uit van je leerdoelen en bepaal met welk type vragen je deze het beste kunt toetsen. De verschijningsvorm van kennistoetsen kan schriftelijk (analoog of digitaal), mondeling of een combinatie hiervan zijn.

Om leerdoelen te beoordelen die gericht zijn op waarneembaar gedrag t.a.v. beroepsbekwaamheid, worden competentietoetsen ingezet. Er zijn veel verschillende soorten competentietoetsen, waarvan veelgebruikte/bekende een beroepsproduct, performance-assessment, practicum, project, stage, scriptie of portfolio zijn. Een toets heeft, naast een bepaalde vorm, ook een bepaalde functie. Er wordt onderscheid gemaakt tussen summatieve en formatieve toetsen; toetsen van het leren en toetsen óm te leren. Een summatieve toets heeft als doel informatie te

verschaffen op grond waarvan beslissingen kunnen worden genomen (kwalificerend). Het is een toets waar de student een (cijfermatige) beoordeling voor krijgt die wordt vastgelegd. Een formatieve toets is bedoeld om de ontwikkeling van de student te volgen en zo nodig bij te sturen (informerend). Een student krijgt geen formele beoordeling. Dit is bijvoorbeeld een oefentoets of (tussentijdse) feedback op een product.

De keuze voor een toetsvorm volgt logischerwijs uit de leerdoelen die zijn gesteld voor de onderwijseenheid (zie kaart 1a). Het waarneembare gedrag dat in de leerdoelen verwoord is met het centrale werkwoord, geeft een indicatie voor de toetsvorm waarmee de leerdoelen het beste kunnen worden gemeten. Belangrijk is dat de vorm van toetsen een directe meting is van de prestatie en recht doet aan het beoogde resultaat (validiteit). Het beoordelen van presentatievaardigheden, een salesgesprek, het uitvoeren van een behandeling of verzorgen van een les, toets je bij voorkeur met een observatie in de vorm van een performance-assessment, stage, video in portfolio of practicum; reflecterend vermogen toets je met een criteriumgericht interview of verslag.

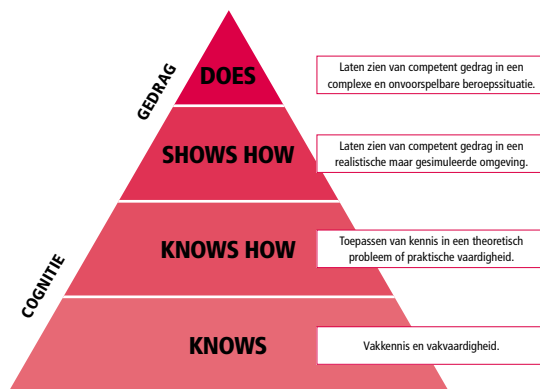
Tabel 1. Uitgewerkte voorbeelden van het kiezen van een toetsvorm.

Bepalen van de toetsvorm	Wat is het leerdoel?	Welk (cognitief) gedrag beoogt het leerdoel en op welk niveau?	Kies de toetsvorm waarmee direct meetbaar is of het leerdoel wordt behaald
Uitgewerkt voorbeeld 1	De student benoemt kenmerken van verschillende ziektebeelden	Hebben van vakkennis: weten en herinneren/ reproduceren	Kennistoets: schriftelijke (digitaal of analoog), mondelinge, open en/of gesloten vragen
Uitgewerkt voorbeeld 2	De student voert o.b.v. een medische casus een behandelplan uit.	Behandelplan uitvoeren: laten zien van competent gedrag in een realistische situatie	Competentietoets: Performance-assessment, stage

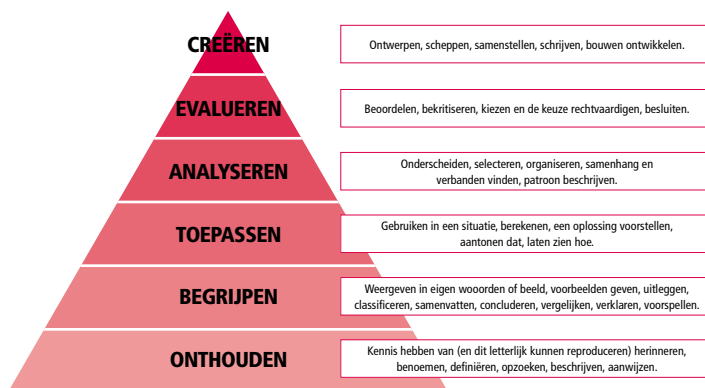
Een niveau-indeling of taxonomie helpt bij het kiezen van een toetsvorm, maar ook bij het formuleren van een vraag of opdracht, omdat deze een indicatie geeft van de moeilijkheidsgraad (niveau). Veel gebruikte taxonomieën zijn de competentiepiramide van Miller (figuur 1) en de taxonomie van Bloom (figuur 2).

Beide taxonomieën zijn van een andere orde en laten zich daardoor niet sluitend één op één vergelijken. Daar waar Bloom enkel ingaat op cognitieve niveaus, maakt Miller binnen competenties onderscheid in cognitie en gedrag. De hoogste cognitieve niveaus van Bloom kunnen

echter ook worden geïnterpreteerd als handelingen die je in een gesimuleerde/gestructureerde beroepssituatie (shows how van Miller) dan wel echte/onvoorspelbare beroepssituatie (does van Miller) kunt uitvoeren. Hierdoor passen vergelijkbare toetsvormen bij de verschillende lagen van beide piramides en zijn sommige toetsvormen in beide beroepssituaties mogelijk. Omdat het in deze kaart gaat om het kiezen van een toetsvorm, zijn beide taxonomieën in tabel 2 naast elkaar gezet. Dit is een vrije/mogelijke interpretatie van hoe de modellen oorspronkelijk zijn bedoeld, die kan helpen om het niveau te duiden en een toetsvorm te kiezen.



Figuur 1. Competentiepiramide van Miller⁴.



Figuur 2. Taxonomie van Bloom⁵.

Tabel 2. Van beheersingsniveau leerdoel naar passende toetsvorm⁶.

Beheersingsniveau leerdoel		Omschrijving	Gewenst (cognitief) gedrag/ voorbeeld toetsvraag/opdracht	Passende toetsvorm (zoals...)
Miller	Bloom			
Knows	Onthouden	Kennis hebben van (en dit letterlijk kunnen reproduceren) door te herinneren, benoemen, definiëren, beschrijven en aan te wijzen.	<ul style="list-style-type: none"> • Geef de definitie van... • Benoem de drie kenmerken van... • Geef volgorde aan van/waar... 	Kennistoets met open en gesloten vragen, voortgangstoets.
	Begrijpen	Het weergeven in eigen woorden, samenvatten, uitleggen, voorbeelden geven, classificeren.	<ul style="list-style-type: none"> • Licht concept/uitspraak/... toe • Leg onderliggend mechanisme uit. • Benoem de 2 verschillen tussen... • Beschrijf de relatie tussen... 	Kennistoets met open en gesloten vragen evt. a.d.h.v. casussen, voortgangstoets.
Knows how	Toepassen	Het gebruiken/toepassen van bestaande kennis in een nieuwe situatie.	<ul style="list-style-type: none"> • Bereken/bewijs in situatie/geval... • Geef de meest waarschijnlijke hypothese/diagnose/uitkomst... • Benoem oorzaken/gevolgen van ingreep... • Geef voor- en nadelen van methode... • Geef een voorbeeld van/evidentie voor... 	Kennistoets met open en gesloten vragen evt. a.d.h.v. casussen, paper/essay, presentatie.
	Analyseren	Het onderscheiden, selecteren, analyseren, organiseren, samenhang en verbanden vinden tussen kennis.	<ul style="list-style-type: none"> • Interpreteer/verklaar... • Geef een onderbouwde beschouwing van... • Teken/schets de... 	Kennistoets met open en gesloten vragen a.d.h.v. casussen, paper/essay, verantwoording.
Shows how	Evalueren	Toetsen en beoordelen van een situatie/product/uitspraak/besluit; onderbouwen van oordeel/conclusies, verantwoorden en reflecteren als professional, weliswaar gesimuleerd.	<ul style="list-style-type: none"> • Vat de conclusie samen en bekritiseer deze. • Geef drie argumenten/redenen voor/tegen... • Beoordeel de uitspraak met behulp/gebruikmaking van... • Bespreek de (ir-)relevante info voor standpunt/ beslissing/voornemen... 	Kennistoets met open en gesloten vragen a.d.h.v. casussen, paper/essay, reflectie, self- en peerassessment, criteriumgericht interview.
	Creëren	Kennis gebruiken om nieuwe ideeën, producten of zienswijzen tot stand brengen als professional, weliswaar gesimuleerd.	<ul style="list-style-type: none"> • Maak een onderzoeksvoorstel/ontwerp/ volgens... • Schrijf/presenteer een advies over/voor... 	Werkstuk/verslag/rapport, project, pitch, practicum, beroepsproduct, onderzoeksvoorstel, performance assessment.
Does		Competent, professioneel gedrag tonen in de beroepspraktijk.	<ul style="list-style-type: none"> • Voer een onderzoek uit waarin... • Voer de dagelijkse handelingen uit die... • Verzamel bewijslast die aantoont dat... 	Werkstuk/verslag/rapport, project, pitch, practicum, performance assessment, stage beoordeling, observatie in dagelijkse praktijk, 360° feedback, scriptie, portfolio.

2 CONSTRUEREN VAN EEN TOETSMATRIJS

WAT IS EEN TOETSMATRIJS?

Een toetsmatrijs is een tabel waarin wordt aangegeven hoe de leerdoelen moeten worden verdeeld over de toetsen en op welk(e) niveau(s) elk leerdoel wordt getoetst. Een toetsmatrijs is daardoor een blauwdruk van de toets: het geeft aan voor elk leerdoel op welk cognitief of gedragsniveau vragen of opdrachten moeten worden bedacht. De toetsmatrijs borgt de validiteit van de toets.

In de toetsmatrijs in tabel 1 zie je een verdeling van de leerdoelen op de verschillende (cognitieve) niveaus en wordt zichtbaar wat de weging is per leerdoel en niveau.

Tabel 1. Voorbeeld van een eenvoudige toetsmatrijs.

Leerdoel	Niveau 1	Niveau 2	Niveau 3	Totaal
Leerdoel X		10%	25%	35%
Leerdoel Y		20%		20%
Leerdoel Z		20%	25%	45%
Totaal		50%	50%	100%

HOE STEL JE EEN TOETSMATRIJS OP?

1. Formuleren van leerdoelen

Een belangrijke randvoorwaarde voor het opstellen van een toetsmatrijs zijn goed geformuleerde leerdoelen (zie kaart 1a). Het advies is om per vak of onderwijsseenheid maximaal vier leerdoelen te formuleren op een hoger cognitief niveau. In de toetsmatrijs kun je vervolgens aangeven op welk(e) beheersingsniveau(s) de leerdoelen getoetst moeten worden.

2. Bepalen van de niveaus waarop het leerdoel getoetst moet worden

Het centrale werkwoord in het leerdoel geeft een duidelijke indicatie voor de toetsvorm (zie kaart 1b). Stel, het leerdoel betreft het opstellen van een financiële businesscase. Het bijbehorende cognitieve niveau is 'creëren'. Als het relevant is om dit leerdoel bijvoorbeeld ook op de cognitieve niveaus 'begrijpen' en 'analyseren' te toetsen, geef je dit aan in de toetsmatrijs. Je kunt hiervoor verschillende niveau-indelingen of taxonomieën (zoals van Bloom en/of Miller) gebruiken. Je ziet dan dat je voor hetzelfde leerdoel op deze verschillende niveaus een aantal vragen of opdrachten moet formuleren.

3. Aanbrengen van een weging van de leerdoelen

Hoe belangrijker een bepaald leerdoel is, hoe meer gewicht dit leerdoel krijgt in de toets en dus hoe groter het aandeel vragen of opdrachten behorend bij dit leerdoel. Het aandeel blijkt uit de moeilijkheidsgraad van de vragen en/of het aantal vragen/opdrachten voor een leerdoel. Krijgen studenten bepaalde informatie of moeten ze zelf meerdere denkstappen maken om tot een antwoord of oplossing te komen? De weging van de leerdoelen maak je heel bewust en beredeneerd. Met een percentage kun je aanduiden hoeveel gewicht het leerdoel krijgt in de toets en op welk(e) niveau(s) je dit leerdoel beoogt te meten in de toets.

Uiteindelijk moet de toetsmatrijs een weerspiegeling zijn van de bedoeling die je als docent hebt om te kunnen beoordelen in hoeverre de leerdoelen zijn behaald. In deze fase bepaal je het absoluut aantal vragen/opdrachten in de toets. Om cognitieve overload bij de student te voorkomen bij het maken van de toets, is het belangrijk de toets niet te lang te maken⁷.

4. Gebruik van een format toetsmatrijs

In tabel 2 zijn de niveau-indelingen die worden gebruikt om competentieniveaus te duiden (knows-knows how, shows how-does) van Miller⁴ en de cognitieve niveaus van Bloom⁵ gecombineerd.

Je stelt de toetsmatrijs als volgt op:

- In de eerste kolom neem je de eindkwalificatie(s) op waartoe het leerdoel behoort.
 - In de tweede kolom neem je de leerdoelen op van de onderwijseenheid.
 - In de opvolgende kolommen neem je de niveau-indeling die wordt gehanteerd op (in dit format zijn de niveau-indeling van Miller en Bloom met elkaar gecombineerd)
 - In de laatste kolom 'weging' geef je aan in welke mate (percentage) het leerdoel terugkomt in de toets en meetelt in het eindcijfer.
- Vervolgens specificeer je (in percentages) de weging van het leerdoel in (verschillende) niveau(s). Dan wordt zichtbaar welk type vragen of opdrachten op welk niveau dienen terug te komen en in de toets.
 - Eventueel duid je het precieze aantal vragen of opdrachten tussen haakjes.
 - Ter controle tel je de percentages van elk niveau (verticaal/kolom) en van elk leerdoel (horizontaal/rij) op. Vervolgens tel je de totalen zowel horizontaal als verticaal op; deze moeten beide uitkomen op 100%. Ook de som van het aantal vragen of opdrachten tussen haakjes (per niveau of per leerdoel) moet overeenkomen met het totaal aantal vragen van de toets.

Tabel 2. *Format toetsmatrijs.*

Beheersingsniveau/type vragen of opdrachten							Weging
Eindkwalificatie	Leerdoelen	Knows		Knows how		Shows how	Does
		Onthouden	Begrijpen	Toepassen	Analyseren	Evalueren	Creëren
2.2	Leerdoel 1	10% (3)					10% (3)
3	Leerdoel 2		25% (4)				25% (4)
4	Leerdoel 3		15% (3)	15% (2)			30% (5)
4	Leerdoel 4		5% (3)	20% (3)			25% (6)
7.1	Leerdoel 5				10% (2)		10% (2)
	Totaal	10% (3)	45% (10)	35% (5)	10% (2)		100% (20)

3a CONSTRUEREN VAN EEN KENNIS- TOETS MET GESLOTEN VRAGEN

WAT IS EEN KENNISTOETS MET GESLOTEN VRAGEN?

Bij het basisontwerp heb je de keuze gemaakt voor een kennistoets met (onder andere) gesloten vragen. Een kennistoets met gesloten vragen bestaat uit vragen waarbij de student uit een beperkt aantal mogelijke antwoorden een goed antwoord of goede antwoorden moet kiezen.

De meest voorkomende gesloten vraagvormen zijn:

1. Meerkeuzevragen: een vraag, gevolgd door meestal 3 of 4 antwoordmogelijkheden
2. Ja/nee of juist/onjuist-vragen: is de bewering goed of fout?

Met name bij digitale toetsing kan gebruik worden gemaakt van vraagvormen waarbij de student een bewerking moet uitvoeren met de verschillende antwoordmogelijkheden⁸:

- ordenen ten opzichte van elkaar (rangschikken)
- aangeven wat bij elkaar hoort (matchen)
- markeren van deelgebieden in afbeeldingen (hotspotvraag)
- gecombineerde vragen die gebaseerd zijn op een casus (samengestelde vragen).

Daarnaast biedt het toevoegen van video en ander multimediamateriaal aanvullende mogelijkheden ten opzichte van analoge toetsen op papier.

HOE CONSTRUEER JE EEN GESLOTEN VRAAG?

Onderdelen

Een gesloten toetsvraag waarbij gekozen moet worden tussen verschillende antwoordmogelijkheden, kent de volgende onderdelen^{8, 9}:

- Gegeven/stam: dit bevat alle informatie die de student nodig heeft om de vraag te kunnen beantwoorden. Deze informatie staat niet ter discussie. Het kan ook een casus zijn. Dit onderdeel is niet noodzakelijk, maar helpt vaak wel om de vraag een context mee te geven.
- Vraag/stelling: datgene wat aan de student gevraagd wordt, of waarvan de student

een uitspraak moet doen over de correctheid.

- Alternatieven: de antwoorden waaruit gekozen kan worden. Onderverdeeld (uiteraard niet vermeld op de toets) in:
 - o De sleutel: juiste antwoord(en)
 - o De afleider(s): foute antwoord(en)

Let op:

- Een stelling is vaak moeilijker te formuleren dan een vraag, omdat het beweerde in alle denkbare gevallen juist of onjuist dient te zijn.
- Wanneer je besluit dat er meerdere antwoorden juist zijn, heeft dit consequenties voor de beoordeling. Als een student slechts één van de gevraagde antwoorden correct heeft, hoeveel punten krijgt hij dan?
- Bij een toets met gesloten vragen, kan bij de normering rekening worden gehouden met de gokkans (**zie kaart 3g**).

Stappenplan⁸

1. Bestudeer het niveau en de leerdoelen uit de toetsmatrijs.
2. Selecteer een vraagtype dat aansluit bij de inhoud zoals in de toetsmatrijs, bij het niveau van de taxonomie die hieraan ten grondslag ligt en de wijze van toetsen (analoog of digitaal).
3. Formuleer de stam.
4. Formuleer de vraag/stelling zo, dat deze ook zonder de antwoordalternatieven te beantwoorden is. Zorg dat uit de vraag blijkt of er één of meerdere antwoorden juist zijn.
5. Formuleer de sleutel.
6. Formuleer de afleiders; kies voor een aantal waarbij alle afleiders inhoudelijk verdedigbaar zijn.
7. Bekijk de vragen na minimaal een dag opnieuw en beantwoord de vragen zonder de alternatieven te bekijken. Pas indien nodig aan.
8. Laat een collega de vragen controleren of de toets blind maken (zonder sleutel).

Houd tijdens het construeren van de vragen het screeningsdocument kennistoets (**zie kaart 3c**) ernaast; dat helpt om je vragen en daarmee je toets valide, betrouwbaar en transparant te maken.

VARIANTEN EN VOORBEELDEN

Ja/nee variant

(voorbeeld op niveau begrijpen)

Stam: Meester Tom is in groep 4 bezig met optelsommen. Hij kent veel verschillende manieren om deze sommen op te lossen en legt deze aan de leerlingen uit.

Stelling: In dit voorbeeld heeft meester Tom de rol van coach.

Ja / Nee

Rangschikvraag in digitaal toetsysteem

(voorbeeld op niveau onthouden)

Leraar lichamelijke opvoeding Ben Janssen onderzoekt of sporten invloed heeft op depressieve gevoelens bij jongeren.

Als hij zijn rapportage maakt na afloop van het onderzoek, in welke volgorde zal hij dan zijn variabelen presenteren in zijn hypothesen? Selecteer de opties hieronder, sleep ze naar het lege vak en plaats ze in de correcte volgorde.

Onafhankelijke variabele
Covariaat
Afhankelijke variabele

--

Vragende variant

(voorbeeld op niveau toepassen)

Een studietoets bestaat uit 70 vragen. De vragen hebben een gemiddelde intercorrelatie van 0,14. Welke waarde benadert het best coëfficiënt alfa?

- a. 0,50
- b. 0,75
- c. 1,00

Casusvraag

(voorbeeld op niveau evalueren)

Je bent huisarts. Je maakt een visite bij mevrouw van Doorn (65 jaar oud). Ze heeft plotseling hevige pijn in haar buik gekregen. Ze vertelt je verder dat [...]. Je besluit lichamelijk onderzoek te doen en vindt [...]. Je vraagt je af of je haar naar het ziekenhuis moet sturen of dat je voorlopig beter thuis kunt aanzien. Wat is de beste actie op dit moment?

- a. Je stuurt haar naar de SEH
- b. Je verwijst haar naar de polikliniek van de internist
- c. Je verwijst haar naar de polikliniek van de gynaecoloog
- d. Je wacht even af en ziet het beloop van de klachten aan.

3b CONSTRUEREN VAN EEN KENNISTOETS MET OPEN VRAGEN

WAT IS EEN KENNISTOETS MET OPEN VRAGEN?

Bij het basisontwerp heb je de keuze gemaakt voor een kennistoets met (onder andere) open vragen. Bij een toets met open vragen produceert en formuleert de student zelf een antwoord.

De verschillende soorten open vragen worden met name bepaald door de lengte van het antwoord. Er zijn korte en lange open antwoordvragen, waarbij het antwoord kan variëren van één ingevuld woord/cijfer tot een langer antwoord of zelfs een opstel/betoog van meerdere pagina's (als onderdeel van een kennistoets). Er is geen harde grens tussen wat kort en lang is te noemen. In- en aanvulvragen, waarbij de student een antwoord invult halverwege of aan het eind van een zin/alea, zijn ook vormen van open vragen^{8,9}.

Alle typen open vragen lenen zich voor digitale afname. Zowel de in- en aanvulvragen als de korte antwoordvragen kunnen automatisch gecorrigeerd worden. De lange antwoordvragen dienen handmatig gecorrigeerd te worden. Toch kan ook dan voor de digitale variant van de lange antwoordvraag worden gekozen, vanwege de borging van de kwaliteit (analyse mogelijkheden), de archivering van de toetsen en de leesbaarheid van de antwoorden.

HOE CONSTRUEER JE EEN OPEN VRAAG?

Onderdelen

Een open vraag bestaat uit de volgende onderdelen⁸:

- Informatie (eventueel): een situatieomschrijving of achtergrondinformatie die de student nodig heeft om de vraag te kunnen beantwoorden. Dit is vaak nodig bij een vraag die een hoger niveau dan puur feitenkennis toetst.
- Vraag/opdracht: datgene wat de student moet beantwoorden/opschrijven, met eventuele instructies, bv. over de lengte van het antwoord.
- Voorgestructureerde antwoordruimte (eventueel): deze kan direct onder de vraag/opdracht

worden geplaatst en geeft enigszins richting aan de omvang van het antwoord.

- Antwoordmodel: eenduidige en volledige omschrijving van het modelantwoord (uiteraard niet vermeld op de toets).

Stappenplan^{8,9}

1. Bestudeer het niveau en de leerdoelen uit de toetsmatrijs.
2. Selecteer een vraagvorm die aansluit bij de inhoud zoals in de toetsmatrijs en bij het niveau van de taxonomie die hieraan ten grondslag ligt.
3. Schrijf per leerdoel de kernbegrippen op; wat moeten de studenten kennen en kunnen?
4. Formuleer op basis van het genoteerde bij stap 3 eerst de modelantwoorden op de vragen die je wilt ontwikkelen, dus nog niet de vragen zelf!
5. Formuleer vervolgens vragen die precies leiden tot de modelantwoorden van stap 4.
6. Vul de vragen aan met informatie en/of instructies.
7. Splits eventueel de informatie en vraag.
8. Benadruk de noodzakelijke aspecten in het modelantwoord en formuleer indien nodig antwoordrestricties.
9. Geef de puntenverdeling (scoring) aan; ken punten toe aan de noodzakelijke aspecten. **(zie kaart 3g).**
10. Maak een instructie voor de beoordelaar.
11. Bekijk de vragen na minimaal een dag opnieuw en beantwoord deze zonder de modelantwoorden. Pas de vragen indien nodig aan.
12. Laat een collega de vragen controleren of de toets blind maken (zonder antwoordmodel).

Houd tijdens het construeren van de vragen het screeningsdocument kennistoets **(zie kaart 3c)** ernaast; dat helpt om je vragen en daarmee je toets valide, betrouwbaar en transparant te maken.

Antwoordmodel

Het doel van het antwoordmodel is om ervoor te zorgen dat één beoordelaar vergelijkbaar beoordeelt en/of dat verschillende beoordelaars vergelijkbaar beoordelen⁸.

Neem het volgende in acht bij het antwoordmodel:

- Neem naast het juiste antwoord, ook alternatieve juiste formuleringen en voorbeelden van foute antwoorden op.
- Omschrijf hoe er omgegaan wordt met 'half goede' antwoorden
- Geef aan hoe mogelijke afwijkingen worden beoordeeld (bv. meer antwoorden dan gevraagd, (door-)rekenfouten of taalfouten)
- In het geval van meerdere punten voor een vraag: omschrijf scorepunten per denkstap en geef aan hoe de punten over de onderdelen van het antwoord zijn verdeeld.

VARIANTEN EN VOORBEEDEN

Invulvraag

(voorbeeld op niveau onthouden)

De moord op ... vormde de directe aanleiding voor de Eerste Wereldoorlog.

Korte antwoordvraag

(voorbeeld op niveau begrijpen)

Informatie: Floor (8 jaar) is verdrietig, omdat haar hond dood is gegaan. Juf Sanne kan zich goed verplaatsen in Floor, omdat haar opa onlangs is overleden. Ze laat aan Floor merken dat ze zich voor kan stellen hoe zij zich voelt, door er met haar over te praten.

Vraag: Welk vermogen toont Sanne in dit voorbeeld?

Lange antwoordvraag

(voorbeeld op niveau onthouden)

Informatie: Het KNMI bepaalt elk uur de temperatuur op een groot aantal meetpunten.

- Vragen:
1. Geef één reden waarom dat elk uur gebeurt.
 2. Geef één reden waarom dat op een groot aantal meetpunten gebeurt.

Betoogvraag in digitaal toetsstelsysteem met videofragment

(voorbeeld op niveau toepassen)

Informatie: <Videofragment van een deel van een consult tussen een psycholoog en cliënt>

Vraag: Bekijk het videofragment en geef aan, ondersteund met argumenten, welke behandelaanpak jouw voorkeur heeft.

Opstelvraag

(voorbeeld op niveau onthouden)

Informatie: Betrouwbaarheid en validiteit zijn kenmerken van toetskwaliteit binnen het hbo.

Vraag: Beschrijf hoe een toetsmatrijs bijdraagt aan het borgen van de betrouwbaarheid en validiteit van een toets.

3c SCREENINGSDOCUMENT KENNISTOETS

VALIDITEIT mate waarin de toets de beoogde leerdoelen meet	O/V	TOELICHTING waarom is het (on)voldoende?
De toetsvorm sluit aan op de leerdoelen en de onderwijsactiviteiten.		
Het beheersingsniveau van de toetsvragen (onthouden, begrijpen, toepassen e.d.) sluit aan op de leerdoelen (zoals aangegeven in de toetsmatrijs).		
Het aantal vragen per leerdoel stemt overeen met de zwaarte van het leerdoel (zoals aangegeven in de toetsmatrijs).		
De inhoud van de toetsvragen sluit aan op de leerdoelen.		
Er wordt een aantal punten toegekend dat past bij de omvang van de vraag en de weging van het bijbehorende leerdoel in de toetsmatrijs.		
BETROUWBAARHEID mate waarin de toets onder dezelfde omstandigheden leidt tot hetzelfde eindoordeel	O/V	TOELICHTING waarom is het (on)voldoende?
ALGEMEEN		
De inleiding van de vraag is duidelijk te onderscheiden van de toetsvraag.		
Alle vragen in de toets zijn onafhankelijk van elkaar.		
De toets maakt onderscheid tussen studenten die goed en minder goed presteren.		
Het aantal vragen is in overeenstemming met de beschikbare tijd voor het tentamen.		
Bij elke vraag wordt voldoende en relevante informatie gegeven; wel de benodigde, maar geen overbodige informatie.		
De informatie wordt geclusterd aangeboden, bijvoorbeeld niet de vraag op de ene pagina en de antwoordopties op de andere pagina.		
Normatieve begrippen als 'moeten', 'dienen', 'belangrijk', 'zinloos', 'wenselijk' zijn zoveel mogelijk vermeden.		
Semi-kwantitatieve begrippen als 'soms', 'zelden', 'veel', 'groot', 'langzaam', 'kort', 'kenmerkend', 'typisch', 'karakteristiek', 'specifiek' zijn zoveel mogelijk vermeden.		
Absolute woorden als 'altijd', 'nooit', 'alleen', 'even vaak', 'identiek' zijn zoveel mogelijk vermeden.		
Vage begrippen zoals 'kunnen', 'spelen een rol' zijn zoveel mogelijk vermeden.		
Vragen zijn geen strikvrAGEN.		
Negatieve woorden zijn door middel van <u>onderstrepen</u> , lettergebruik of vet gemarkeerd.		
OPEN VRAGEN		
Er is duidelijk wat voor soort antwoord wordt verwacht (voorbeeldantwoord, aangeven soort informatie).		
Er wordt niet naar de mening van de student gevraagd.		
De vraag is enkelvoudig gesteld (één vraag of probleemstelling tegelijk).		
De vragen zijn efficiënt (kort, eenvoudig, begrijpelijk, duidelijk).		
MEERKEUZE VRAGEN		
Minimaal 40 4-keuze vragen, 60- 3 keuzevragen of 80 2-keuzevragen.		
Er is rekening gehouden met de gokkans: aantal items is voldoende om de betrouwbaarheid te borgen.		

De toetsvragen zijn te beantwoorden zonder de antwoordmogelijkheden te zien.		
Er is slechts één antwoordmogelijkheid de juiste.		
Het juiste antwoord staat willekeurig op diverse posities.		
Alle antwoordmogelijkheden (afleiders en sleutel) zijn vergelijkbaar qua lengte, woordgebruik en formulering.		
Voor de student die de stof niet beheerst zijn alle antwoordmogelijkheden even aannemelijk.		
Antwoordmogelijkheden sluiten grammaticaal en inhoudelijk aan op de vraag.		
Antwoordmogelijkheden kennen geen overlap en sluiten elkaar uit.		
Antwoordmogelijkheden bevatten geen herhaling uit de vraag.		
Dubbele ontkenning tussen de vraag en één of meer antwoordmogelijkheden is vermeden.		
ANTWOORDMODEL		
Er is een uitgewerkt antwoordmodel met beoordelaarsinstructie, uitgewerkte antwoorden, scoringsvoorschrift, normering en cesuur.		
De beoordelaarsinstructie bevat richtlijnen voor het opmaken van de beoordeling (scoren-normeren-waarderen/becijferen).		
De beoordelaarsinstructie beschrijft hoe met afwijkingen wordt omgegaan, zoals 'half goede' antwoorden, te veel antwoorden, reken- of taalfouten.		
Het antwoordmodel bevat het juiste antwoord (bij meerkeuze eenduidig).		
Het antwoordmodel bevat uitwerkingen van de antwoorden zoals:		
- zo veel mogelijk voorbeelden van goede antwoorden		
- voorbeelden van veel voorkomende antwoorden		
- voorbeelden van (foute) antwoorden die geen punten krijgen		
- kernwoorden die in elk geval in het antwoord moeten voorkomen.		
Het antwoordmodel bevat informatie over de puntentoekening per vraag (<i>bij voorkeur 1 scorepunt per denkstap</i>).		
Er zijn geen doorrekenfouten mogelijk tussen verschillende vragen.		
CESUUR – NORMEREN – WAARDEREN		
De cesuur doet recht aan het behalen van de leerdoelen (<i>niet automatisch op 55% gesteld?</i>).		
Er is een normeringstabel waarin het aantal behaalde scorepunten worden omgezet naar een cijfer.		
TRANSPARANTIE mate waarin de informatie over de toets voor alle betrokkenen helder is	O/V	TOELICHTING waarom is het (on)voldoende?
Toetsvorm en type vragen zijn vooraf bekend.		
Toegestane hulpmiddelen zijn vooraf bekend.		
De beschikbare tijd is bekend.		
Aantal vragen is weergegeven.		
Aantal pagina's van de toets en het einde van de toets zijn weergegeven.		
De toetsvragen zijn eenduidig en taalkundig correct geformuleerd.		
Puntenverdeling, normering en cesuur is duidelijk.		
De lay-out van de toets is overzichtelijk.		
De pagina's zijn genummerd.		
Aangegeven is waar de student zijn uitwerkingen op moet schrijven.		

3d CONSTRUEREN VAN EEN COMPETENTIETOETS

WAT IS EEN COMPETENTIETOETS?

Een competentietoets bestaat uit een (beroeps) opdracht die resulteert in een (beroeps)product of prestatie. Dit product of deze prestatie weerspiegelt de kerntaken van een beroepsbeoefenaar⁸.

Competentietoetsen beoordelen het gewenste professionele gedrag (competenties) afgeleid van de beroepsbekwaamheid waartoe wordt opgeleid.

Competentietoetsen kennen een (beroeps) opdracht en een beoordelingsformulier met beoordelingscriteria waarmee het product of de prestatie van de student beoordeeld wordt.

Een beoordelingsformulier bestaat uit:

- Beoordelingscriteria; aan welke eisen dient de prestatie o.b.v. de leerdoelen te voldoen?
- Een beoordelingschaal; op welke schaal wordt een prestatie gewaardeerd?
- Normering en cesuur (**zie kaart 3g**): richtlijn over hoe de beoordeling tot een eindoordeel (cijfer) leidt.
- Beoordelaarsinstructie: richtlijnen over het gebruik en invullen van het beoordelingsformulier.

Er zijn veel verschillende competentietoetsen, waarvan de meest gebruikte/ bekende een beroepsproduct, performance-assessment, practicum, project, stage, scriptie of portfolio zijn.

HOE STEL JE EEN COMPETENTIETOETS OP?

Hieronder volgt een stappenplan voor het opstellen van een competentietoets. Houd tijdens het opstellen het screeningsdocument competentietoets (**zie kaart 3e**) ernaast; dat helpt om je opdracht en criteria en daarmee je toets valide, betrouwbaar en transparant te maken.

1. Focus op de prestatie

Goede leerdoelen representeren om welk gedrag het gaat en welke prestatie daarvoor geleverd dient te worden. Over deze prestatie wil je met behulp van de toets uitspraak doen. Om tot een valide uitspraak te komen, dient de prestatie verder

geconcretiseerd te worden in beoordelingscriteria. De opdracht aan de student moet vervolgens leiden tot deze prestatie.

2. Formuleer de beoordelingscriteria

Voordat je de toets construeert, start je met het ontwikkelen van de beoordelingscriteria. Hierdoor krijg je scherp wat je wilt gaan beoordelen en welke opdracht hieraan ten grondslag dient te liggen. Formuleer de beoordelingscriteria in samenwerking met betrokken collega's, zodat overeenstemming en duidelijkheid t.a.v. de beoordelingscriteria wordt bereikt. Beoordelingscriteria geven weer aan welke eisen het resultaat van de prestatie moet voldoen. Ofwel: welk 'bewijs' geleverd moet worden zodanig dat de prestatie als voldoende kan worden beoordeeld¹⁰.

De leerdoelen vormen wederom het uitgangspunt voor het formuleren van de beoordelingscriteria en worden op het beoordelingsformulier vermeld. Bepaal per leerdoel aan welke criteria (eisen) de student moet voldoen om uitspraak te kunnen doen over de beheersing van het leerdoel; wat moet de student minstens laten zien voor een voldoende prestatie? Er zullen meerdere beoordelingsaspecten te onderscheiden zijn. Bij een presentatie wordt bijvoorbeeld niet alleen gekeken naar de mondelinge overtuigingskracht, maar ook naar structuur en de visuele ondersteuning van de presentatie. Het is aan te raden om per criterium/leerdoel (het nummer van) de bijbehorende eindkwalificaties op te nemen op het beoordelingsformulier.

De zwaarte van de leerdoelen kan worden weerspiegeld met een wegingsfactor of meerdere eisen/criteria die aan de prestatie worden gesteld. Een goed beoordelingscriterium is geformuleerd in termen van concreet waarneembaar gedrag. Hoe concreter de beoordelingscriteria zijn geformuleerd, hoe meer ze zich lenen voor feed-up aan de student (**zie kaart 6**). Ze sturen het leergedrag van de student.

Holistisch en analytisch beoordelen

De mate waarin de beoordelingscriteria worden uitgewerkt is ook afhankelijk van hoe de prestatie het best beoordeeld kan worden; volstaat een meer globale (holistische) beoordeling van de prestatie of is het wenselijk veel specifiek (analytisch) de prestatie te beoordelen? Bij analytisch beoordelen worden veel beoordelingscriteria onderscheiden om de prestatie te beoordelen en wordt veelal per criterium uitspraak gedaan over de beheersing ervan. Bij holistisch beoordelen wordt over een beperkt aantal beoordelingscriteria samen een uitspraak gedaan.

3. Formuleer de opdracht (de toets)

Op basis van de beoordelingscriteria formuleer je de opdracht; dit is de toets. De opdracht aan de student moet leiden tot de gewenste prestatie in lijn met de leerdoelen. De opdracht is dusdanig geformuleerd dat de student weet welke prestatie van hem verwacht wordt om zijn competenties te bewijzen. Het is een concrete en heldere instructie aan de student en geeft weer onder welke condities de opdracht plaatsvindt. Ook blijkt uit de opdracht welke hulpmiddelen en bronnen de student mag inzetten en eventueel hoeveel tijd de student aan de opdracht mag besteden.

4. Stel de beoordelingsschaal op

Een beoordelingsschaal geeft de niveaus van beheersing weer van de prestatie. A.d.h.v. de beoordelingsschaal wordt de geleverde prestatie gewaardeerd. De keuze voor de reikwijdte (2-, 3, 10-puntsschaal) van de beoordelingsschaal is afhankelijk van hoe specifiek je de prestatie wilt waarderen. Voor te stellen is dat naarmate de reikwijdte van de beoordelingsschaal groter wordt, het onderscheid tussen niveau 6 of 7 lastiger te definiëren is. Wanneer er een verdere specificatie gewenst is van de waarde van de prestatie op elk niveau van de beoordelingsschaal, ontwikkel je hiervoor rubrics (**zie kaart 3f**). Een rubric is een vorm van een analytische beoordelingsschaal.

5. Stel de normering en cesuur op

Of je holistisch of analytisch beoordeelt heeft ook consequenties voor hoe je de normering

en cesuur van de toets opstelt en vastlegt in het beoordelingsformulier. Wanneer je analytisch beoordeelt, ligt de cesuur wellicht hoger dan bij een holistische beoordeling, aangezien je per criterium een oordeel velt en elk criterium van voldoende niveau moet zijn. Bij een holistische beoordeling is de insteek dat je een oordeel velt over een set aan criteria (een onderdeel), dus bijvoorbeeld per beoordelingsaspect zoals in het geval bij presenteren t.a.v. de overtuigingskracht, structuur en visuele ondersteuning. Criteria binnen zo'n set (onderdeel) zijn met elkaar te compenseren, terwijl dat in het geval van een analytische beoordeling niet mogelijk is.

De cesuur van de toets leg je vast door een richtlijn (scoringsvoorschrift) op te stellen welke eindscore de student moet behalen om te slagen voor de toets. Moet de student voor elk criterium minstens een voldoende scoren of voor elk onderdeel? Welke criteria of onderdelen mogen (eventueel) met elkaar gecompenseerd worden? Op basis van de cesuur kan de normering van de toets worden bepaald; welke scores leiden tot welk cijfer? Dit kan a.d.h.v. een formule, tabel of beschrijving inzichtelijk worden gemaakt. (**zie kaart 3g**).

6. Stel een beoordelaarsinstructie op

Tot slot beschrijf je alle relevante zaken in een beoordelaarsinstructie, zodat de examinator weet hoe dit beoordelingsformulier te gebruiken en in te vullen. Naast het scoringsvoorschrift beschrijf je eveneens hoeveel examinatoren betrokken zijn bij de beoordeling, wie welke rol heeft en hoe de beoordeling moet worden opgemaakt/ingevuld. De kern is dat de beoordelaarsinstructie alle benodigde informatie bevat die bijdraagt aan een betrouwbare beoordeling.

3e SCREENINGSDOCUMENT COMPETENTIETOETS

VALIDITEIT mate waarin de toets de beoogde leerdoelen meet	O/V	TOELICHTING waarom is het (on)voldoende?
De toets is gebaseerd op de toetsmatrijs.		
Aantal en niveau van de toetsopdracht(en) komen overeen met de toetsmatrijs.		
Het beheersingsniveau van de toetsopdracht(en) (kennis, inzicht, toepassing, probleemoplossen e.d.) sluit(en) aan op de leerdoelen.		
De inhoud van de toetsopdracht(en) sluit aan op de leerdoelen.		
De beoordelingscriteria zijn afgestemd op de opdracht(en) qua niveau en inhoud.		
De beoordelingscriteria toetsen (representeren) de leerdoelen.		
Het aantal beoordelingscriteria komt overeen met de zwaarte van het leerdoel (conform de toetsmatrijs).		
In het beoordelingsformulier is een koppeling gemaakt tussen de leerdoelen en de beoordelingscriteria.		
De toets meet de beoogde competentie(s).		
BETROUWBAARHEID mate waarin de toets onder dezelfde omstandigheden leidt tot hetzelfde eindoordeel	O/V	TOELICHTING waarom is het (on)voldoende?
OPDRACHT(EN)		
De opdracht maakt het voor de student mogelijk zijn competentie(s) te bewijzen.		
De opdrachten maken een onderscheid tussen studenten die goed en minder goed presteren (specificiteit).		
In het geval van een groepsopdracht is duidelijk wat het aandeel is van iedere individuele student.		
De tijd die de student nodig heeft om de toetsopdracht te maken is in overeenstemming met de beschikbare tijd.		
AFNAME		
De afnamecondities zijn duidelijk omschreven zodat de beoordelingen van kandidaten vergelijkbaar zijn.		
De omstandigheden waaronder de toets wordt uitgevoerd, zoals voorbereidingen, hulpmiddelen, context/setting, inrichting toetsruimte zijn voor alle studenten hetzelfde.		
BEORDELING (BEORDELINGSFORMULIER - CRITERIA - SCHAAL - NORMERING/CESUUR - INSTRUCTIE)		
In het beoordelingsformulier zijn de beoordelingscriteria, een beoordelingschaal, normering en cesuur en instructies voor de beoordelaar opgenomen.		
De beoordelingscriteria zijn specifiek en eenduidig: benoemen eigenschappen van de prestatie.		
De beoordelingscriteria geven weer aan welke norm de prestatie moet voldoen.		
De beoordelingscriteria zijn op basis van consensus/ kalibratie tot stand gekomen.		

Er is gebruik gemaakt van een eenduidige beoordelingsschaal waarmee de prestatie gewaardeerd kan worden.		
Het aantal beoordelingscriteria is voldoende om een uitspraak te kunnen doen over de prestatie van de student.		
De beoordelaarsinstructie bevat richtlijnen voor het opmaken van de beoordeling (scoren-normeren-waarderen/becijferen).		
De beoordelaars zijn bekwaam om de toets af te nemen en de prestatie van de student te beoordelen.		
Er worden meerdere beoordelaars betrokken in de beoordeling van de toets.		
Er wordt een onafhankelijke beoordelaar betrokken bij de beoordeling.		
TRANSPARANTIE mate waarin de informatie over de toets voor alle betrokkenen helder is	O/V	TOELICHTING waarom is het (on)voldoende?
De toetsopdracht, beoordelingscriteria, cesuur & normering, procedure (ontvankelijkheidseisen) & voorwaarden (plaats en tijd van de toetsafname, hulpmiddelen en mate van zelfstandigheid) zijn vooraf bekend voor de student.		
De toetsopdracht is eenduidig en geconcretiseerd naar wat er van de student wordt verwacht.		
De opdracht geeft weer onder welke voorwaarden (plaats, tijd, hulpmiddelen, mate van zelfstandigheid) de opdracht moet worden uitgevoerd.		
De beoordelingscriteria zijn transparant; geven weer wat van de student verwacht wordt (concreet gedrag) en welke eisen er worden gesteld aan het resultaat.		
De puntenverdeling (scorepunten), normering en de cesuur zijn duidelijk.		
In het beoordelingsmodel is ruimte voor feedback opgenomen.		
De toets kent (een) formatieve toetsmoment(en) in voorbereiding op de toets.		

3f ONTWIKKELEN VAN RUBRICS

WAT ZIJN RUBRICS?

Rubrics [ofwel rubrieken] zijn beschrijvingen van bepaald gedrag of kenmerken van een product op verschillende prestatieniveaus. Rubrics vormen een uitgewerkte beoordelingschaal van een beoordelingsformulier om prestaties en producten a.d.h.v. niveaubeschrijvingen te kunnen waarderen. De prestatieniveaus zijn gebaseerd op specifiek uitgewerkte beoordelingscriteria al dan niet in gedragskenmerken of indicatoren. Rubrics zijn geschikt als instrument voor het beoordelen van beroepsproducten en prestaties waaraan een kwaliteitsniveau moet worden toegekend en wanneer je veel belang hecht aan de ontwikkelgerichte functie van een toets.

WAT ZIJN VOOR- EN NADELEN VAN RUBRICS?

Voordelen

Een belangrijk voordeel van rubrics ten opzichte van een 'gewoon' beoordelingsformulier is dat de beoordelingscriteria op verschillende prestatieniveaus zijn uitgewerkt. Door de concrete uitwerking van de beoordelingscriteria kunnen studenten zichzelf en elkaar gericht 'inschalen' in de diverse rubrieken. Dat maakt een rubric ook geschikt als instrument voor een tussentijdse evaluatie of zelfbeoordeling. Voor een examiner heeft een rubric als voordeel dat de feedback al is uitgewerkt in diverse beoordelingsschalen: er staat al welk oordeel hoort bij een bepaald zichtbaar resultaat.

Nadelen

Het kost veel tijd om de beoordelingscriteria op verschillende prestatieniveaus uit te werken. Zeker omdat het vaak erg moeilijk is om elk prestatieniveau op een betekenisvolle manier te beschrijven. Doordat de onderbouwing al beschreven is, heeft de examiner minder flexibiliteit bij de beoordeling. Beoordelingen met rubrics komen regelmatig uit in de tussenliggende rubric, omdat de beoordelingscriteria verschillende aspecten bevatten. De cijfers liggen dan ongewenst vaak tussen het cijfer 6 en 7.

HOE MAAK JE RUBRICS?

1. Op basis van de leerdoelen (in termen van meetbaar gedrag en het beoogde resultaat) formuleer je specifieke beoordelingscriteria: wanneer is het gewenste gedrag of het beoogde resultaat goed?
2. Voor elk beoordelingscriterium brainstorm je over bijbehorende kenmerken, zodanig dat je voor elk criterium tot niveaubeschrijvingen komt. Kijk of je tot voldoende gedragskenmerken komt om de verschillende prestatieniveaus te vullen. Merk je dat je enkel het gewenste en ongewenste gedrag kunt beschrijven, kies er dan voor om een 'gewoon' beoordelingsformulier met alleen het gewenste prestatieniveau te maken.
3. Formuleer vervolgens nauwkeurige gedragsbeschrijvingen voor het laagste en het hoogste niveau. Gebruik hierbij geen woorden als beter, vaker of excellent. Vermijd negatieve beschrijvingen, maar definieer precies wat de student op een bepaald niveau moeten laten zien.
4. Maak daarna de gedragsbeschrijvingen voor de overgebleven niveaus. Beperk het schema tot één A4-tje.
5. Voeg de normering en cesuur toe.
6. Bespreek de rubric met collega's en studenten om te bepalen of alles duidelijk is beschreven.
7. Verzamel voorbeelden of bewijzen van ingeleverd werk voor elke niveaubeschrijving, zodat je die later kunt gebruiken als ijkpunt.
8. Evalueer de rubric en pas het eventueel aan.

Tabel 1. Voorbeeld van Rubrics¹¹.

	KAN BETER (1 PUNT)	VOLDOENDE (2 PUNTEN)
PROBLEEMSTELLING Een echt probleem heeft veel kanten. Een duidelijke probleemafbakening is essentieel.	De probleemstelling is niet of onduidelijk beschreven.	Je neemt de probleemstelling over uit de beschrijving van de praktische opdracht.
PROBLEEMVERKENNING Welke factoren spelen een rol? Hangen die factoren samen? Welke resultaten leveren concrete berekeningen op?	Je onderzoekt de twee gegeven vraagstellingen met een voorbeeld en trekt hieruit een (verkeerde of geen) conclusie	Je onderzoekt de twee gegeven vraagstellingen met enkele voorbeelden en tekeningen en trekt daaruit een conclusie (waar, niet waar, waar als...)
PLAN MAKEN Op basis van de probleemstelling stel je een aantal onderzoeksvragen op.	De onderzoeksvragen zijn niet of onduidelijk beschreven.	Je geeft de twee onderzoeksvragen.
PLAN UITVOEREN Je onderzoekt je vermoedens verder, geeft voorbeelden van diverse functies en zoekt naar uitzonderingen.	Je onderzoekt de twee vraagstellingen met een voorbeeld.	Je onderzoekt de twee vraagstellingen met enkele voorbeelden, tekeningen en berekeningen.
CONCLUSIES TREKKEN Wat heb je nu ontdekt? Wat geldt bij welke aannames? Welke uitzonderingen?	Er is geen of een verkeerde conclusie getrokken.	Je trekt conclusies (waar, niet waar, waar als...).
VERDIEPING Vervolgonderzoek met dezelfde stappen als hiervoor.	Je onderzoekt de gegeven vraagstelling met een voorbeeld en trekt geen of een verkeerde conclusie	Je onderzoekt de gegeven vraagstelling met enkele voorbeelden en tekeningen en trekt daaruit een conclusie.
PRESENTATIE	Het werkstuk ziet er onverzorgd uit. Het onderzoek wordt niet duidelijk omschreven.	Het werkstuk bevat een weergave van het voorwerk, het onderzoek, de complexe opdracht en de uitwerking hiervan. En het ziet goed verzorgd uit.
PLANNING	Het verslag is te laat.	
TOTAAL		

Formule: subscore per onderdeel = aantal punten keer absolute wegingspercentage. Eindscore = optelsom van subscores. Maximaal te behalen punten: $4 \times 100 = 400$, minimaal te behalen punten: $1 \times 100 = 100$. 5,5 bij 200 punten. Cijfer berekenen a.d.h.v. normeringstabel.

GOED (3 PUNTEN)	UITSTEKEND (4 PUNTEN)	WEGING	SCORE
Je formuleert de probleemstelling in een paar algemene zinnen.	Op basis van het verkennen van het probleem, formuleer je het probleem in een paar heldere volzinnen.	5%	
Je onderzoekt de vraagstellingen met een reeks voorbeelden die systematisch zijn gekozen en geeft een volledige conclusie.	... en je geeft vanuit een analyse van algemene formules van lijn en parabool een algemeen geldende wiskundige verklaring.	15%	
Je geeft twee onderzoeksvragen waarbij je ook aangeeft welk vermoeden je hebt over de uitkomst.	... en je geeft ook nog andere onderzoeksvragen en vermoedens die bij je opkomen, in samenhang met de gegeven vragen.	20%	
Je onderzoekt de vraagstellingen met een reeks voorbeelden die systematisch zijn gekozen.	... en je geeft vanuit een analyse van algemene formules van lijn en parabool een algemeen geldende wiskundige verklaring.	20%	
Je geeft volledige conclusies.	... en je trekt ook conclusies over de manier waarop je het onderzocht hebt.	10%	
Je onderzoekt de vraagstelling met voorbeelden die systematisch zijn gekozen en geeft een heldere wiskundige verklaring en een volledige conclusie.	... en je geeft vanuit een analyse van algemene formules van lijn en parabool een algemeen geldende wiskundige verklaring en conclusie.	15%	
... en het is overzichtelijk ingedeeld. De vragen van de complexe opdracht zijn helder geformuleerd. De taal is correct. De lay-out is overzichtelijk.	... en het is een leesbaar stuk voor medestudenten die de opdracht niet gedaan hebben.	10%	
	Het verslag is op tijd.	5%	
		100%	

3g OPSTELLEN VAN NORMERING EN CESUUR

WAT IS SCOREN EN NORMEREN?

Om tot cijfers voor een toets te komen, moet je de scoring en normering bepalen. Uit de score blijkt hoeveel punten de student per vraag of beoordelingscriterium kan halen. Normeren is het proces waarbij scores worden omgezet naar cijfers of waarderingen²³. Een belangrijke vraag bij het normeren is, waar de cesuur komt te liggen.

WAT IS CESUUR?

De cesuur is de zak/slaaggrens, het aantal punten dat een student tenminste moet behalen om een voldoende te halen.

HOE BEREKEN JE CIJFERS VOOR EEN TOETS?²³

1. Bepaal de scoring en maximale score van de toets.
2. Bepaal de cesuur: welke score moet gehaald worden voor het cijfer 5,5?
3. Corrigeer een eventuele gokkans bij een meerkeuzetoets
4. Werk een normeringstabel uit, waarin je scores koppelt aan cijfers.

1. Bepalen van de scoring per onderdeel en maximumscore van de toets

Voor de scoring bepaal je op basis van welke criteria studenten punten kunnen verdienen, gevolgd door toepassing van die criteria op de toets²³. De score is niet automatisch 1 punt per vraag of per beoordelingsaspect. Sommige vragen of beoordelingsaspecten zijn complexer en vereisen meer denkstappen of handelingen waardoor er meer punten voor kunnen worden gegeven. De punten van alle vragen of beoordelingsaspecten samen, vormen het totaal: de maximumscore van de toets. De maximumscore ligt niet automatisch op 100 punten, maar is dus afhankelijk van het aantal punten dat per vraag of per beoordelingsaspect kan worden behaald. Een toets kan elke denkbare reikwijdte en daarmee maximumscore hebben, bv. van 0 tot 10, van 1 (voor het opschrijven van je naam) tot 43, van 500 tot 550 (zoals bij de CITO-toets basisonderwijs), enz.²³.

2. Bepalen van de cesuur

Vervolgens wordt bepaald waar de zak/slaaggrens, de cesuur, komt te liggen. Binnen Avans wordt vrijwel altijd een absolute cesuur gebruikt. Dat houdt in dat de beheersingsgraad (in de vorm van een percentage van het totaal aantal punten dat behaald moet worden voor een voldoende) voorafgaand aan de afname is vastgesteld. De cesuur is niet automatisch 55%. Dit percentage wordt door de examinerator (al dan niet in overleg met collega's) bepaald o.b.v. de beoogde beheersingsgraad en de complexiteit van de vraag, het beoordelingsaspect of de totale toets. In het geval van het toetsen van de handeling 'injecteren' door een student verpleegkunde is voor te stellen dat je vereist dat de beheersing 100% is, waardoor de cesuur dus op 100% ligt. Alle beoordelingsaspecten moeten dan voldoende zijn om te slagen voor de toets.

3. Corrigeren van de gokkans

Bij een meerkeuzevraag bestaat de mogelijkheid dat een student door gokken/raden voor het juiste antwoord kiest. De gokkans is gelijk aan 1 gedeeld door het aantal antwoordalternatieven. Bij een driekeuzevraag is de gokkans dus $1/3$, bij een vierkeuzevraag $1/4$, bij een vijfkeuzevraag $1/5$ enz. De gokscore wordt berekend door de gokkansen van alle meerkeuzevragen bij elkaar op te tellen. Als een toets bestaat uit 40 vierkeuzevragen, is de gokscore $40 \text{ maal } 1/4 = 10$, wat betekent dat een student door uitsluitend te gokken statistisch gezien 10 vragen juist zou kunnen beantwoorden. Als de cesuur bij deze toets op 55% ligt, moet een student naast deze 10 vragen die door gokken juist beantwoord kunnen worden, 55% (beheersingsgraad) van de overige 30 vragen goed hebben. 55% van 30 is afgerond 17 vragen. In dit voorbeeld moet de student dus $10 + 17 = 27$ vragen goed hebben voor een voldoende. We zeggen dan dat de op gokkans gecorrigeerde cesuur bij 27 ligt.

4. Uitwerken van een normeringstabel

Wanneer de maximumscore en de cesuur van de toets zijn bepaald, moeten scores (behaald aantal punten) omgezet worden naar cijfers m.b.v. een normeringstabel. Om tot een bewuste en beargumenteerde normering te komen, verdient het de voorkeur te werken met een normeringstabel waarbij de 1-, 10- en cesuur-score worden gebruikt. Dit wordt de niet-lineaire methode genoemd. Je bepaalt hoeveel punten er behaald moeten zijn om het cijfer 1 (1-score), het cijfer 10 (10-score) of het cijfer 5,5 (cesuur-score) te verkrijgen.

- De 1-score bepaal je door je af te vragen bij welk aantal punten je vindt dat de student geen noemenswaardige prestatie heeft geleverd²³ of bij een meerkeuzetoets hoe groot de gokkans

is en het aantal goed gegokte antwoorden als 1-score te gebruiken. Het hoeft niet automatisch zo te zijn dat alleen bij het behalen van 0 punten een 1-score wordt gegeven, wellicht is het behalen van 10% van de maximaal te behalen score nog steeds reden om een 1-score te geven.

- De 10-score bepaal je door je af te vragen bij welk aantal punten je vindt dat de student een uitmuntende prestatie heeft geleverd en dus het cijfer 10 verdient. Moet de toets perfect gemaakt zijn of is het behalen van (bv.) 90% van de maximaal te behalen score genoeg voor een 10-score?
- De 1-, 10, en cesuur-score vormen het uitgangspunt voor de normeringstabel en op basis hiervan worden alle tussenliggende scores omgezet naar een cijfer.

Tabel 1. Normeringstabel.

40	10	10-SCORE	
39	10		
38	9,7		
37	9,4		
36	9,1		
31	7,6		
30	7,3		
26	6,1		
25	5,8		
24	5,5		CESUUR-SCORE
23	5,3		
22	5,1		
21	4,9		
11	2,9		
10	2,7		
9	2,5		
8	2,3		
4	1,5		
3	1,3		
2	1		
1	1	1-SCORE	

UITGEWERKT VOORBEELD

Een toets bestaat uit 40 open vragen waarbij per vraag 1 punt kan worden behaald. De maximumscore van de toets is 40 punten. De examinator wenst een beheersingsgraad van 60% en dit betekent dat de cesuur-score ligt op 24 punten. De examinator besluit dat de 1-score op 2 punten ligt en de 10-score op 39 en 40 punten. Onder de cesuur-score moet je 22 stappen maken om van het cijfer 5,5 naar cijfer 1 te komen. Je deelt dan de 4,5 cijferstappen door 22 scorestappen ($4,5/22=0,204$). Boven de cesuur-score moet je 15 stappen maken om van cijfer 5,5 naar cijfer 10 te komen. Je deelt 4,5 cijferstappen door 15 scorestappen ($4,5/15=0,3$). Dit levert de volgende (gecomprimeerde) normeringstabel (tabel 1) op.

4a BETROUWBARE AFNAME EN BEOORDELING

WAT IS EEN BETROUWBARE AFNAME?

Bij een betrouwbare afname gaat het om het creëren van de juiste omstandigheden voor de afname van de toets, zodanig dat de prestatie van studenten zo eerlijk en betrouwbaar mogelijk kan worden beoordeeld. De toets moet de leerdoelen meten en geen andere factoren, zoals concentratie en stressbestendigheid.

Voordat de toets wordt afgenomen zorgt de examinerator ervoor dat studenten onder de juiste condities getoetst kunnen worden. Deze condities hebben zowel te maken met de objectiviteit en vergelijkbaarheid van de studentprestaties, als met de voorwaarden en omstandigheden tijdens de toetsafname.

Binnen Avans Hogeschool geldt bij afname van geroosterde toetsen (veelal schriftelijke toetsen) het protocol van RET¹². Bij niet-geroosterde toetsen, zoals een beroepsproduct, stage, presentatie e.d., dien je als examinerator(en) zelf zorg te dragen voor de juiste condities zodat studenten objectief, vergelijkbaar en eerlijk kunnen worden beoordeeld.

HOE DRAAG JE ZORG VOOR EEN BETROUWBARE TOETSAFNAME?

- Heldere toets-instructies voor de student, zoals:
 - o welke hulpmiddelen wel of niet zijn toegestaan (bv. wetboek, rekenmachine);
 - o lengte en duur van de toets;
 - o normering van de toets, per vraag of opdracht en de cesuur (slaag/zakgrens).
- Voldoende materialen om de toets uit te kunnen voeren;
 - o voldoende exemplaren van de toets bij analoge afname of computers bij digitale afname
- Geschikte toetslocatie:
 - o voldoende ruimte, normale temperatuur en goede ventilatie;
 - o geen storende factoren (glazenwasser of een luid pratende surveillant).

- Gelijkwaardigheid tussen studenten:
 - o gelijkwaardige of dezelfde toets
 - o goed geïnstrueerde examinatoren en helder beoordelingsmodel
 - o goed geïnstrueerde surveillanten (wat is wel of niet toegestaan tijdens de afname, hoe te handelen bij problemen)
 - o iedereen krijgt dezelfde – werkende – hulpmiddelen tot zijn beschikking (bij practica).
- Gelijkwaardige omstandigheden voor studenten:
 - o beschikbare/toegestane tijd
 - o inrichting van de toetslocatie
 - o voorwaarden voor deelname (ontvankelijkheidseisen)
 - o complexiteit van context/setting van bijvoorbeeld stageplek of handelen van een spelacteur in een rollenspel
 - o de vooraf aangereikte informatie en de aard van de begeleiding en feedback gedurende het maken van o.a. beroepsproduct, stages
- Fraudebestrijding door surveillanten:
 - o identiteitscontrole in het lokaal
 - o presentielijsten, ook af laten tekenen bij vertrek inclusief tijdstip
 - o voorvallenregistratieformulieren door surveillant in te vullen bij onregelmatigheden
 - o afspraken omtrent: toiletbezoek, voortijdig verlaten van lokaal, specifieke eigendommen (tas, horloge) innemen.

WAT IS EEN BETROUWBARE BEOORDELING?

Bij betrouwbaar beoordelen gaat het erom dat de beoordelaar op zo'n manier een toets beoordeelt, dat hij onder dezelfde omstandigheden weer tot hetzelfde eindoordeel komt. Dit geldt voor één examinerator die een toets beoordeelt, maar ook voor verschillende examinatoren die (al dan niet samen) een zelfde toets beoordelen. Als gevolg van een betrouwbare beoordeling leidt eenzelfde antwoord of prestatie van verschillende studenten tot hetzelfde aantal punten.

HOE DRAAG JE ZORG VOOR EEN BETROUWBARE BEOORDELING?

Als de examiner de studenten kent, bestaat het gevaar dat hij hun eerdere prestaties of bijvoorbeeld zijn persoonlijke ervaringen met hen tijdens de lessen, laat meetellen bij de (eind)beoordeling. Zinnvolle maatregelen hierbij zijn: het niet zichtbaar maken van de namen van studenten of het beoordelen van een andere klas studenten. Ook kan een beoordelaar zich (ongemerkt) gaan aanpassen aan de geleverde prestaties van studenten. Hij kan bijvoorbeeld minder streng worden als na een aantal beoordelingen blijkt dat de meeste studenten eenzelfde vraag fout beantwoorden. Of hij kan de gemiddelde prestatie of zijn eigen gedrag tot norm gaan verheffen. Een voorzorgsmaatregel is hiervoor om per vraag na te kijken en in overleg te gaan (voorafgaand of tijdens het nakijken) met een andere examiner van dezelfde toets.

Om betrouwbaar te kunnen beoordelen is een goed beoordelingsmodel essentieel.

- Voor een kennistoets is dit een uitgewerkt antwoordmodel met beoordelaarsinstructie, uitgewerkte antwoorden, scoringsvoorschrift, normering en cesuur.
- Voor een competentietoets is dit een beoordelingsformulier met specifieke beoordelingscriteria, een beoordelingsschaal, normering en cesuur en een beoordelaarsinstructie.

Als de examinatoren voorafgaand aan de beoordeling antwoord hebben op de volgende vragen, is de kans groot dat hun beoordeling betrouwbaar is:

- Welk gedrag moet een student laten zien?
- Welke bewijsvoering moet hij hiervoor leveren?
- Welke (kwantitatieve en kwalitatieve) eisen stellen we aan de betreffende antwoorden, producten of prestaties?
- Aan welke beoordelingscriteria (gedragskenmerken) of indicaties moet de bewijsvoering of het antwoord voldoen?

- Wanneer is het gedrag en het geleverde bewijs in een bepaalde context goed (cesuur)?

Om de betrouwbaarheid tussen verschillende beoordelaars te verhogen is het zinvol om een kalibreersessie (**zie kaart 4b**) te organiseren. Tijdens zo'n sessie beoordeel je met mede-examinatoren in een paar stappen een bepaalde toets om de beoordelingscriteria en de gewenste normering te verduidelijken, zodanig dat alle examinatoren op een meer dezelfde en eenduidige wijze de prestatie van student beoordelen. Dat komt de betrouwbaarheid van de beoordeling ten goede.

4b HOUDEN VAN EEN KALIBREERSESSIE

WAT IS EEN KALIBREERSESSIE?

Een kalibreersessie is een gezamenlijke bijeenkomst van examinatoren waarin één of meerdere werkstukken van studenten worden besproken. Het doel van een kalibreersessie is het verduidelijken van de beoordelingscriteria en de gewenste normen uit het beoordelingsmodel om op die manier betrouwbaarder te kunnen beoordelen. Het is niet bedoeld om een oordeel te vellen over de student, de begeleidend docent of de examinator van het werkstuk.

HOE HOUD JE EEN KALIBREERSESSIE?¹⁰

Wanneer één werkstuk wordt besproken duurt de sessie meestal 2,5 uur. De sessie wordt gepland voorafgaand aan het beoordelingsmoment of meerdere keren per jaar als er doorlopende beoordelingsmomenten zijn (niet periodegebonden).

Rollen tijdens de kalibreersessie:

- Organisator: plant en organiseert de sessie en zorgt voor de beschikbaarheid van de werkstukken.
- Examinator: deelnemer van de sessie als betrokkene bij de beoordeling van het werkstuk. Examinatoren zijn inhoudelijk deskundig, weten welke eindkwalificaties dienen te worden getoetst en hebben kennis van de door de student(en) opgeleverde werkstuk(ken).
- Moderator: begeleider van de sessie.
- Notulist: legt verslag van de kalibreersessie.

Voorafgaand aan de sessie:

- De organisator van de sessie zorgt voor een geschikte ruimte en een geschikt tijdstip en stuurt een week voorafgaand aan de sessie het werkstuk aan de deelnemende examinatoren toe. De werkstukken zijn bij voorkeur geanonimiseerd om vooroordelen t.a.v. studenten of examinatoren te voorkomen.
- De examinatoren beoordelen voorafgaand aan de sessie het werkstuk a.d.h.v. het beoordelingsformulier en nemen het

beoordelingsformulier mee naar de sessie.

- De moderator beoordeelt voorafgaand aan de sessie (het) werkstuk(ken) met het beoordelingsformulier. Hij maakt een presentatie met daarin een overzicht van alle prestatiecriteria.

Tijdens de sessie:

- De moderator zorgt voor een U-vormige opstelling, zodat examinatoren en moderator elkaar goed kunnen zien. Daarnaast creëert de moderator een veilige sfeer en open dialoog.
- De moderator vraagt aan de start van de sessie het oordeel van de examinatoren, zo kan een indruk verkregen worden van de overeenstemming die er is en de noodzaak voor kalibratie.
- Vervolgens vraagt de moderator om het oordeel op het eerste prestatie criterium voor het eerste werkstuk. Wanneer er verschillen zijn in oordelen wordt om een argumentatie gevraagd.
- De moderator stimuleert de examinatoren om zoveel mogelijk de woorden van de beoordelingscriteria te gebruiken. Hierdoor ontstaat zicht op de betekenis die de examinatoren geven aan de criteria. De examinatoren en de moderator gaan in de argumentatie op zoek naar aanvullende criteria die worden genoemd.
- De moderator zorgt ervoor dat examinatoren praten over interpretatie van de criteria en niet over de beoordeling van de student.
- Na het bespreken van het eerste prestatie criterium worden achtereenvolgens alle prestatiecriteria besproken. Een notulist notuleert aanvullende criteria die worden benoemd en breed worden gedragen.
- Streef als moderator niet altijd naar besluitvorming. Parkeer zo nodig criteria bij de notulist. De geparkeerde punten worden door een kleiner groepje examinatoren uitgewerkt in een voorstel dat in een volgende kalibreersessie wordt besproken.

De verslaglegging van de kalibreersessie kan worden gedaan in tabelvorm (zie tabel 1). In de eerste kolom is het prestatie criterium weergegeven, in de tweede en derde kolom

aanvullende criteria, uitgesplitst in indicaties (eigenschappen van het werkstuk) voor een positieve score op het criterium en contra-indicaties waaruit blijkt dat het niet aan het prestatie criterium voldoet. Tijdens de sessie ontdekken examinatoren vaak waar het beoordelingsformulier kan worden verbeterd; dit kan men kwijt in een vierde kolom. Eventueel kan een vijfde kolom worden toegevoegd waarin opmerkingen komen over aanpassingen aan het curriculum om studenten beter voor te bereiden om aan het prestatie criterium te kunnen voldoen.

Tabel 1. *Rapportage kalibreersessie (voorbeeld).*

PRESTATIECRITERIUM	CONCLUSIE OVER DE BETEKENIS VAN HET PRESTATIECRITERIUM:		SUGGESTIES VOOR AANPASSING VAN HET FORMULIER	CONSEQUENTIES VOOR HET CURRICULUM
	INDICATIES	CONTRA-INDICATIES		
De onderzoeksvraag is helder	Alle begrippen zijn gedefinieerd	Er staat geen onderzoeksvraag (een zin met een vraagteken) in het werkstuk	Vervang het woord 'helder' door 'precies' zodat verwezen wordt naar de literatuur van Oost en Markenhof (2002)	Er wordt in het curriculum te weinig geoefend met het formuleren van vragen. Inbouwen in jaar 1.

Na afloop van de sessie:

- De notulist stuurt de toelichting die is ontwikkeld tijdens de sessie aan alle deelnemers.

5 UITVOEREN VAN EEN TOETSANALYSE

WAT IS EEN TOETSANALYSE?

Wanneer een toets is afgenomen, de ruwe scores van de afzonderlijke vragen of opdrachten bekend zijn en op basis van de vooraf vastgestelde cesuur de cijfers berekend zijn, kan een analyse plaatvinden. Het doel ervan is om op basis van de toetsscores inzicht te krijgen in de kwaliteit van de toetsvragen of opdrachten en van de toets als geheel¹.

Als bron kunnen gegevens uit Osiris worden gebruikt. Bij toetsen met gesloten vragen kunnen psychometrische analyses worden uitgevoerd met het statistische programma Sonate of met een digitaal toetssysteem zoals Remindo Toets. Bij toetsen met open vragen of competentietoetsen kan Excel hiervoor worden gebruikt. Analyse kan aanleiding geven tot het aanpassen van je onderwijs, je opdracht(en) en/of je beoordelingsmodel. In een uiterst geval kan het ook aanleiding geven tot aanpassing van de huidige beoordelingen als studenten benadeeld zijn door de kwaliteit van de toets. Vaar echter nooit blind op een (psychometrische) toetsanalyse; neem altijd je eigen oordeel en bevindingen mee. Klopt de analyse met wat je had verwacht?

HOE VOER JE EEN TOETSANALYSE UIT?

In Osiris of met een digitaal toetssysteem zoals Remindo Toets kunnen onder andere gegevens

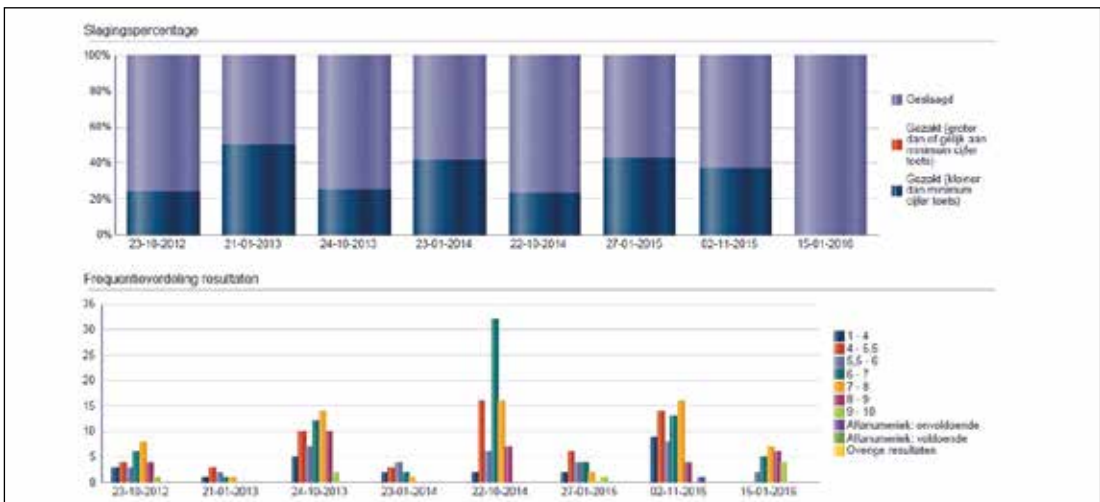
worden verkregen over groeps grootte, slagingspercentage, gemiddelde cijfer, spreiding van cijfers. Deze cijfers geven je informatie en roepen mogelijke vervolgvragen op.

Slagingspercentage

Wat zegt een slagingspercentage van 50%, of 90%? Voldoet dit aan je verwachting? Er is geen 'perfect slagingspercentage'. Dit hangt samen met de functie van de toets en plaats van de toets in de opleiding (drempel om bv. stage te mogen gaan doen), behaalde resultaten door eerdere cohorten (is het slagingspercentage vergelijkbaar met vorige cohorten, of is het ineens veel hoger of lager), afspraken over slagingspercentages binnen de academie e.d..

Spreiding

Ook de spreiding van cijfers geeft informatie. Zitten alle cijfers tussen de vijf en de zes of zijn er ook enen en tieners gevallen? Hoe groter de spreiding, hoe meer onderscheid de toets maakt tussen 'goede' en 'zwakke' studenten. Bij een psychometrische analyse wordt gekeken naar de kwaliteit van de toetsopdrachten en naar de betrouwbaarheid van de toets in totaal. Door invoer van het behaalde aantal punten per student per vraag kunnen Cronbach's Alpha, p-waardes, RIT-waardes en a-waardes worden berekend.



Figuur 1. Voorbeeld van gegevens uit Osiris.

Tabel 1. Toelichting psychometrische grootheden.

GROOTHEID	BETEKENIS	WAARDES	INTERPRETATIE
CRONBACH'S ALPHA	Maat voor de betrouwbaarheid van de toets als geheel. Het is de mate waarin de toets consistent meet.	Waardes tussen 0 en 1	Een waarde dicht bij 0 geeft een lage consistentie/ betrouwbaarheid aan, een getal dicht bij 1 een hoge consistentie/ betrouwbaarheid. Factoren die de betrouwbaarheid beïnvloeden zijn: <ul style="list-style-type: none"> • De itemcorrelatie (Rit) • De moeilijkheidsgraad (p-waarde) • Aantal en keuze van alternatieven • Formuleringsfouten of sleutfouten • Toetslengte en aantal vragen • Te grote gokkans • De mate van heterogeniteit van de samenstelling van de groep • De groepsgrootte • Ondeugdelijke, niet eenduidig geformuleerde toetsvragen • Geen goede spreiding tussen moeilijke en makkelijke toetsvragen • Ongewenste storingen tijdens de toetsafname • De aanwezigheid van herkansers in de populatie • Slechte alternatieven • Vragen die niet relevant zijn voor de toets- en studiestof
P- WAARDE (p)	De moeilijkheidsgraad van een individuele toetsvraag of opdracht. De P-waarde is de proportie studenten die de toetsvraag goed heeft beantwoord of het percentage punten dat gehaald is op de vraag/ opdracht t.o.v. het totaal aantal punten dat was te behalen.	Waardes tussen 0 en 1	Vragen of opdrachten met een P-waarde dicht bij 1 worden goed gemaakt door studenten, vragen of opdrachten met een P-waarde dicht bij 0 worden slecht gemaakt. Dit kan meerdere dingen betekenen: <ul style="list-style-type: none"> • De vraag is verkeerd gesleuteld (sleutel, juiste antwoord aanpassen) • De vraag is foutief geformuleerd (vraag verwijderen) • De vraag is te moeilijk • Het onderwerp is in het onderwijs niet aan bod geweest • Er zijn meerdere antwoorden mogelijk en goed (formulering aanpassen)
A- WAARDE (A)	De A-waarde geeft de aantrekkelijkheid van de afleiders (de foute antwoorden in een meerkeuzetoets) aan. De A-waarde is de proportie studenten die de betreffende afleider heeft gekozen. Dit doe je voor iedere afleider apart.	Waardes tussen 0 en 1	De A-waarde mag niet te laag zijn, want dan vervult het antwoord zijn functie als afleider niet. Daarnaast mag de A-waarde niet hoger zijn dan de P-waarde. In dat geval is de afleider aantrekkelijker dan het goede antwoord. Om te weten wat er dan aan de hand is, moet je terug kijken naar je toetsvraag: is de afleider misschien toch ook goed? Is dit niet aan bod geweest in je onderwijs, of is er iets anders aan de hand?
RIT-WAARDE	De Rit-waarde geeft het onderscheidend vermogen van een vraag of opdracht aan. Dit wil zeggen, in hoeverre maakt de betreffende vraag of opdracht onderscheid tussen goede en slechte studenten?	Waardes tussen -1 en 1	Een negatieve Rit-waarde wil zeggen dat op deze vraag of opdracht studenten die over de gehele toets goed gescoord hebben, deze vraag of opdracht niet goed gescoord hebben en dat de studenten die over het geheel genomen slecht hebben gescoord, deze vraag of opdracht juist wel goed hebben gedaan. Deze vraag meet blijkbaar een heel ander concept dan de rest van de toets. Een (ruim) positieve Rit-waarde wil het omgekeerde zeggen (en dit is wat je wil): een vraag of opdracht wordt goed uitgevoerd door de studenten die de toets als geheel goed hebben gemaakt en slecht uitgevoerd door studenten die de toets over het geheel slecht hebben gemaakt. Je vraag of opdracht meet dan hetzelfde als de rest van de toets.

6 GEVEN VAN FEEDBACK

WAT IS FEEDBACK?

Feedback is de informatie die een student krijgt over waar hij staat in zijn leerproces en hoe hij zich verder kan ontwikkelen. Daarbij worden drie vormen van feedback onderscheiden namelijk¹³;

- Feed-up: welke prestatie moet worden geleverd?
- Feedback: hoe heeft de student een bepaalde prestatie uitgevoerd?
- Feedforward: welke aanpak is nodig om de beoogde prestatie te bereiken? (figuur 1).

Effectieve feedback

Alleen een cijfer geven is niet voldoende: de student weet nu wat zijn prestatie waard is, maar niet waarop hij zich zou kunnen verbeteren. Om het gat te kunnen overbruggen heeft de student informatierijke feedback nodig: welke onderdelen van de prestatie waren goed (feedback)? Hoe kan de student zich verbeteren (feedforward)? De combinatie tussen feedback en feedforward is groeigerichte feedback: het stelt de student in staat zich verder te ontwikkelen.

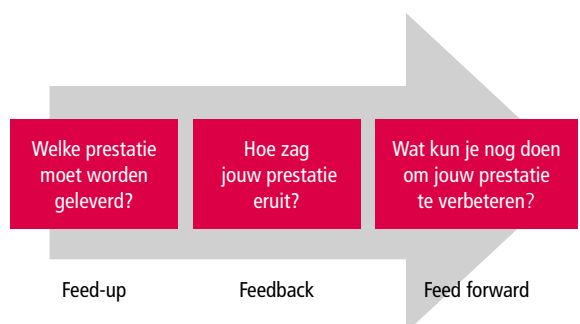
HOE GEEF JE FEEDBACK NAAR AANLEIDING VAN EEN TOETS?

Feedback geef je om de kloof te overbruggen tussen wat de student wel beheerst en wat nog niet¹⁴. Feed-up, feedback en feedforward geef je op verschillende niveaus, namelijk op;

- De taak: de prestatie die de student levert
"De probleemstelling van je onderzoek is nog onvoldoende afgebakend en eenduidig."
- Het proces; de wijze waarop de prestatie tot stand komt "Je zou je presentatie nog beter kunnen maken door gebruik te maken van een heldere structuur."
- De zelfregulatie: gericht op het leerproces van de student "Goed dat je om extra uitleg hebt gevraagd."

Feedback op de eigenschappen van de student (slim van je!) helpt de student niet om beslissingen te nemen over vervolgstappen in het leerproces en is dus weinig effectief^{13, 14}. Feedback op de taak, het proces of zelfregulatie wordt gezien als het meest waardevol voor de leeropbrengst¹⁵.

Feedback kan worden gegeven door de docent, maar daarnaast ook door medestudenten (peerfeedback) of door de student zelf (self-feedback) ondersteund door een goed uitgewerkt beoordelingsmodel. Door de concrete uitwerking van de beoogde prestatie in beoordelingscriteria heb je als docent, maar ook als student, houvast voor de te leveren prestatie (feed-up). Wanneer de beoordelingsschaal uitgewerkt is naar rubrics (zie kaart 3f), leent het beoordelingsmodel zich nog beter voor feedback. De rubrics bevatten informatie over de verschillende prestatieniveaus. Studenten kunnen zichzelf en elkaar gericht 'inschalen' in de diverse rubrieken en krijgen zicht op hoe zij zich kunnen verbeteren en ook als docent gebruik je de informatie uit de rubric(s) om feedforward te kunnen formuleren. Maak gebruik van self-feedback door de student, dit maakt de kans dat de student de feedback opvolgt groter. Feedback geven is geen monoloog, maar een dialoog¹⁶.



Figuur 1. Feedback is een verzamelnaam voor zowel feed-up, feedback als feedforward¹⁵.

Tabel 1. *Tips per toetsvorm.*

Kennistoets met gesloten vragen	Organiseer het inzagemoment op zo'n manier dat het leergericht is door bijvoorbeeld met de studenten alle vragen te doorlopen en aan te geven waarom de sleutel(s) goed is/zijn en afleiders fout. Bij de digitale afname van een kennistoets kan de student direct na afname zien wat goed was en wat niet met informatie over de reden van het goede of foute antwoord. Bij digitale afname kan de feedback direct plaatsvinden, maar zijn mondelinge aanvullingen niet mogelijk.
Kennistoets met open vragen	Het antwoordmodel geeft studenten inzicht in de modelantwoorden. Ook hier kun je met studenten de vragen doorlopen en aangeven welke foutieve antwoorden werden gegeven en waarom het modelantwoord de volledige punten geeft.
Stage	Dit is bij uitstek een toets waar veel tussentijdse feedback wordt gegeven, maar vaak door een praktijkbegeleider. Als docent heb je beter zicht op de gewenste prestatie (namelijk het Hbo-niveau) en moet je je ervan vergewissen dat feedback ook informatie geeft over huidige prestatie t.a.v. de gewenste prestatie. Je moet dus ook (middels stageopdrachten en/of stagebezoeken) inzicht verkrijgen in het 'gat' tussen huidige prestatie en gewenste prestatie. Stage leent zich voor mondelinge feedback, maar het is belangrijk om ook schriftelijk feedback te geven ¹⁷ .
(Project/ afstudeer)-opdrachten	Deze toetsvorm leent zich voor veel tussentijdse feedback gericht op het product, het proces en zelfregulatie. Hierbij kunnen ook vormen van self-, en peerfeedback worden ingezet. Bij dergelijke toetsen moet je als docent zelf ook de informatie over de prestatie van de student vergaren om informatierijke feedback en feedforward te kunnen geven.

Tabel 2. *Tips bij variatie in studentfactoren.*

(Ze)er slechte prestatie van de student	Het is beter om mondelinge feedback te geven naast schriftelijke (bijvoorbeeld in geval van projectopdrachten).
De docent kent de student en voorgaande prestaties goed	Laat de feedback aansluiten bij de student en voorgaande prestaties. Dit geeft blijk van een doorgaande ontwikkeling.
De docent kent de student en voorgaande prestaties niet (of slecht)	Feedback blijft vaak oppervlakkiger. Vraag mondeling door of de student de feedback herkent uit eerdere prestaties. Probeer hiermee aan te sluiten bij de doorgaande ontwikkeling van de student.
Mate van verplichting tot het ontvangen van feedback	Wanneer iemand er zelf om vraagt staat hij/zij er meestal meer voor open. In een gesprek kan ook bij 'verplichte' feedback nagegaan worden waar de student behoefte aan heeft.
Mindset	Heeft de student de overtuiging dat groei en ontwikkeling mogelijk is (groiegerichte mindset ¹⁸), dan staat deze meer open voor feedback dan wanneer hij/zij er van overtuigd is dat de prestatie afhankelijk is van een vaststaande intelligentie of mate van creativiteit (fixed mindset). Mindset is ontwikkelbaar.

7 REFLECTIE OP TOETSCYCLUS

WAT IS REFLECTIE OP DE TOETSCYCLUS?

Reflecteren betekent jezelf of anderen een spiegel voorhouden, met een (positief) kritische blik kijken naar jezelf en je handelen in bepaalde situaties. Het doel van reflecteren is om tot meer inzicht of een verbetering van je gedrag te komen.

Bij de laatste fase van de toetscyclus reflecteer je op alle fasen van de doorlopen cyclus, zodanig dat je tot sterke en verbeterpunten van de toetsontwikkeling komt. Hierdoor kun je vóór de volgende onderwijsperiode een kwaliteitsslag maken met bijvoorbeeld de toets(onderdelen), de verzorgde onderwijsactiviteiten, de feedbackmomenten of het toetsproces.

HOE REFLECTEER JE OP DE TOETSCYCLUS?

Om tot sterke en verbeterpunten van je toets te komen, kijk je met enige afstand naar de gehele toetscyclus. Het is hierbij aan te raden om samen met een collega jouw bevindingen door te spreken. De volgende vragen kunnen je helpen om tot concrete verbeteracties te komen.

- In hoeverre weerspiegelt en meet mijn toets het gewenste gedrag van onze studenten in de beroepspraktijk?
- Hoe sterk is de samenhang tussen de leerdoelen, de toets(vorm) en de leeractiviteiten?
- Welke afstemming met collega's is nodig als ik bepaalde aanpassingen wil doen?
- Hoe kan ik mijn studenten nog beter voorbereiden op de toets? Wat kan ik hen tijdens de lessen en thuis (nog meer) laten doen in lijn met het beoogde toetsresultaat?
- Hoe kan mijn feedback (voor en na de toets) hen helpen om zich verder te ontwikkelen?

WAT DOE JE MET DE UITKOMSTEN?

Als je tot zinvolle aanbevelingen komt voor bijvoorbeeld andere leerdoelen en/of een andere toetsvorm, ga je hierover in gesprek met je (vak)collega's en de betreffende onderwijs- of curriculumcommissie van je academie. Zij zijn verantwoordelijk voor het totale toetsprogramma en het curriculum van je opleiding.

TOETSVORMEN

Op deze kaart worden definities gegeven van in de toolkit genoemde toetsvormen^{9, 10, 19, 20, 21, 22}.

Deze indeling is niet uitputtend en bevat onderlinge overlap.

Beroepsproduct: het resultaat van een door de student uitgevoerde beroepsopdracht, bedoeld om zijn/haar bekwaamheid met betrekking tot een kerntaak van het beroep te demonstreren. Een beroepsproduct kent verschillende verschijningsvormen, zoals een advies, ontwerp, eindproduct, handeling of onderzoek.

Criteriumgericht interview: een individueel vraaggesprek ten behoeve van een beoordeling van competenties (en onderliggende kennis) op basis van een uitgevoerde opdracht, gebeurtenis of handeling. Deze toetsvorm staat meestal niet op zichzelf, maar gaat vaak gepaard met een observatie, documentanalyse of portfolio, en biedt dan aanvullende informatie.

Essay: een korte of lange beschouwing aan de hand van een vraagstelling, waarin de student mondeling of schriftelijk bronnen behandelt, een eigen analyse/argumentatie/mening geeft en conclusies formuleert.

360° Feedback: een toetsvorm waarbij door de student aan personen vanuit verschillende rollen feedback wordt gevraagd op competenties op de werplek, in afstudeeropdrachten en projecten.

Kennistoets: een schriftelijke of mondelinge toets gericht op cognitieve doeleinden, bestaande uit open en/of gesloten vragen. Een schriftelijke toets kan digitaal of op papier worden afgenomen. Door het gebruik van casussen (ook wel: casustoets) kunnen ook hogere cognitieve niveaus worden gemeten.

Observatie in dagelijkse praktijk: het waarnemen van de houding/attitude en handelingen/vaardigheden van een student in de alledaagse authentieke beroepssituatie. Deze toetsvorm is uitermate geschikt om een uitspraak te doen over de beroepsbekwaamheid van een student.

Onderzoeksvoorstel: het plan van aanpak voor een onderzoek. Het onderzoeksvoorstel wordt geschreven voorafgaand aan de uitvoering.

Paper: een betoog aan de hand van een vraagstelling, waarin bronnen worden behandeld en conclusies worden geformuleerd. Het verschil met een essay is klein; een paper is enkel gebaseerd op feiten, waardoor hier geen mening/visie in is verwerkt.

Peer-assessment: een toetsvorm waarbij medestudenten elkaars werk beoordelen en relateren aan de leerdoelen. Studenten geven elkaar hierbij feedback.

Performance assessment: een toetsvorm waarin een student gedrag moet laten zien tijdens het uitvoeren van kenmerkende, kritische beroepstaken in een authentieke of gesimuleerde beroepssituatie. Voorbeelden: presentatie verzorgen, (advies)gesprek voeren, geven van een les, rollenspel, behandeling van een patiënt. Een vaardigheidstoets, waarin een handeling wordt uitgevoerd, is een vorm van een performance assessment.

Portfolio: deze term wordt op twee manieren gebruikt. Het kan een (papieren of digitaal) dossier zijn waarin een student diverse bewijsmaterialen verzamelt, om met dat geheel zijn competenties aan te tonen (veelal gevolgd door een criteriumgericht interview, dan ook wel portfolio-assessment genoemd). Het wordt ook gebruikt voor een op zichzelf staande verslagvorm, waarin een student moet aantonen, dat hij over de gevraagde kennis en vaardigheden beschikt.

Practicum: een (begeleide) praktische opdracht of proef in een veelal gesimuleerde situatie, bedoeld om leerstof beter te begrijpen of vaardigheden te oefenen.

Presentatie: een mondelinge toelichting aan een publiek, vaak ondersteund met audiovisueel materiaal. Wordt ook gebruikt als mondelinge verdediging van een product of prestatie.

Project: een project is bedoeld om een of meer concrete resultaten tot stand te brengen. Een project kan op basis van een beroepsopdracht resulteren in een beroepsproduct als dan niet op basis van een real-life casus/ vraagstuk uit de beroepspraktijk. De weg er naartoe is redelijk voorspelbaar en er zijn vooraf bindende afspraken over kwaliteit, opleverdatum en beschikbaar budget. Het resultaat van een project is concreet; je kunt het voelen, zien, horen en/of ruiken.

Rapport: een document waarin verslag wordt gedaan van een activiteit, toestand, gebeurtenis of proces, bijvoorbeeld een onderzoeksrapport. Een rapport of rapportage is vaak formeler dan een verslag.

Reflectie: het (al dan niet begeleid) terugkijken op eigen leerprocessen en producten om daarvan te kunnen leren. Veelal wordt de reflectie meetbaar middels een (proces)verslag of mondeling gesprek.

Scriptie: het verslag van een (afstudeer)onderzoek, waarin methodisch vragen worden beantwoord die leiden tot relevante kennis, conclusies en aanbevelingen.

Self-assessment: een methode waarin studenten hun eigen werk relateren aan de leerdoelen, door hun eigen werk te beoordelen, verschillen tussen hun werk en de gewenste uitkomsten te constateren en nieuwe activiteiten in de richting van de leerdoelen uit te voeren.

Stage beoordeling: een verzamelnaam voor de toetsing van het werkplekleren/het leren in de beroepspraktijk. Hiervoor kunnen andere toetsvormen worden gebruikt, zoals een portfolio, 360° feedback, performance assessment, verslagen, reflecties, naargelang de leerdoelen en beoogde competenties van de stage.

Verantwoording: een omschrijving, toelichting of verklaring van gemaakte keuzes of de doorlopen ontwikkeling/ proces. Een verantwoording is vaak een aanvullende toetsvorm, naast bijvoorbeeld een beroepsproduct. Een verantwoording kan worden vastgelegd in verslagvorm of rapportage.

Verslag: een schriftelijke weergave (verslaglegging) van een doorlopen proces, uitgevoerde handeling of leerproces waarop gereflecteerd dient te worden. Een verslag is vaak informeler dan een rapport.

Voortgangstoets: een (meestal schriftelijke) toets gericht op cognitieve doeleinden die een vaste opzet en inhoud kent, waarbij de ontwikkeling (groei) in kennis gedurende het studiejaar of door de studie jaren heen wordt gemeten en de beheersingsgraad van de leerstof steeds hoger komt te liggen. Deze vorm van toetsen bevordert een continue wijze van studeren.

Werkstuk: een schriftelijk of tastbaar product of resultaat van een opdracht. Een verslag of rapport zijn vormen van een werkstuk, maar ook een (bijvoorbeeld bouwkundig of creatief) ontwerp is een werkstuk. Een werkstuk kan ook een beroepsproduct zijn, naargelang de opdracht betrekking heeft op de kerntaken uit het beroep waartoe wordt opgeleid.

BRONNEN

- ¹ Expertgroep BKE/SKE (2013). *Verantwoord toetsen en beslissen in het Hoger Beroepsonderwijs. Een voorstel voor een programma van eisen voor een basis- en seniorkwalificatie examinering (BKE/SKE)*. Den Haag: Vereniging van Hogescholen.
- ² Biggs, J. B. (2003). *Teaching for quality learning at university* (second edition). Buckingham: Open University Press/Society for Research into Higher Education.
- ³ Mager, R. F. (1984). *Preparing instructional objectives* (second edition). Belmont, CA: David S. Lake.
- ⁴ Miller, G. E. (1990). The Assessment of Clinical Skills/Competence/Performance. *Academic Medicine*, 65(9), 63-67.
- ⁵ Anderson, L. W. & Krathwohl, D. W. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. New York: Pearson Education.
- ⁶ Milius, J., Doosje, S., & Dekker, K. (2015). *Handout 1a Wanneer kiezen voor Multiple Choice?* Geraadpleegd op 1-9-2016 van <https://tauu.uu.nl/wp-content/uploads/2015/03/Handout-1a-Wanneer-kiezen-voor-Multiple-Choice.pdf>
- ⁷ Expertgroep Brein & Leren (2016). *Bouwstenen voor Leren*. Breda: Avans Hogeschool
- ⁸ Biermond, I., Bos-Horstink, M., Soeting, J., Sugito, A., & Uil-Hoogerwaard, W. (2015). *Hoe maak ik goede toetsen met gesloten en open vragen?* Wilp: Teelen B.V.
- ⁹ Berkel, H. van, Bax, A., & Joosten-ten Brinke, D. (2014). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.
- ¹⁰ Expertgroep Protocol (2014). *Beoordelen is mensenwerk. Bevindingen over de wenselijkheid en mogelijkheid van een gezamenlijk protocol voor het beoordelen van (kern)werkstukken*. Den Haag: Vereniging Hogescholen.
- ¹¹ Spijkerboer, L., & Straatman, P. (2004) Rubrics. Een hulpmiddel bij het beoordelen van praktische opdrachten. *Euclides*, 79(6), 270-273.
- Geraadpleegd van http://vakbladeuclides.nl/archief/pdf/79_2003-04_06.pdf
- ¹² Avans Hogeschool (2016). *Toetsregeling Avans Hogeschool 2016-2017*. Breda: Avans Hogeschool
- ¹³ Fisher, D., & Frey, N. (2009). Feed up, Back, Forward. *Multiple Measures*, 67, 20 – 25.
- ¹⁴ Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77, 81 – 112.
- ¹⁵ Smith, P. L., & Ragan, T. J. (2005). *Instructional design* (3e editie). New York: Wiley.
- ¹⁶ Pelgrim, E. A. M. (2013). *Clarifying observation and assessment feedback in workplace-based learning*. Nijmegen: UMC St. Radboud.
- ¹⁷ Shute V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153 – 189.
- ¹⁸ Dweck, C.S. (2012). *Mindset. Changing the way you think to fulfil your potential*. Hachette: Robinson.
- ¹⁹ Jaspers, M., & Zijl, E. van (2011). *Kwaliteit van toetsing in het Hoger Onderwijs. Naar een versterkte rol van examencommissies*. Eindhoven: Dienst Onderwijs en Onderzoek Fontys Hogescholen.
- ²⁰ Hogeschool van Amsterdam (2012). *15 Toetsvormen*. Geraadpleegd op 19-9-2016 van https://score.hva.nl/docent/15_toetsvormen/Paginas/default.aspx
- ²¹ Avans Hogeschool (z.j.). *Succesvol veranderen. Projectmatig creëren met Commitment*. Breda: Avans Hogeschool.
- ²² Encyclo (2016). *Definities opzoeken* Geraadpleegd op 19-9-2016 van <http://www.encyclo.nl/>
- ²³ Hulshof, C. (2012, 18 december). *Hoe bereken je cijfers voor een toets?* [blog post]. Geraadpleegd van <https://onderzoekonderwijs.net/2012/12/18/hoe-bereken-je-cijfers-voor-een-toets/>





CONTACT

Hogeschoollaan 1
4818 CR Breda
(076) 525 05 00

toetssupport@avans.nl

COLOFON

Dit is een uitgave van Avans Hogeschool

Oplage:

900

Tekst en redactie:

Expertgroep Toetsing, Leer- en Innovatiecentrum, Avans Hogeschool:
Marjolein van Loon, Imke van den Meijdenberg-Brok, Els Roskam-Pelgrim, Allard Visser

Coördinatie:

Diensten Marketing, Communicatie en Studentzaken, Avans Hogeschool
Programma BDB+, Avans Hogeschool

Vormgeving:

HeldHavtig bv, Breda

Druk:

De Bondt grafimedia communicatie, Barendrecht

Versie:

Maart 2017

COPYRIGHT

Je bent vrij om:



- het werk te delen - te kopiëren, te verspreiden en door te geven via elk medium of bestandsformaat
- het werk te bewerken - te remixen, te veranderen en afgeleide werken te maken

Onder de volgende voorwaarden:



Naamsvermelding - De gebruiker dient de maker van het werk te vermelden en aan te geven of het werk veranderd is. Je mag dat op redelijke wijze doen, maar niet zodanig dat de indruk gewekt wordt dat de licentiegever instemt met je werk of je gebruik van het werk.



Niet Commercieel - Je mag het werk niet gebruiken voor commerciële doeleinden.

