



TOETSEN VOLGENS DE TOETSCYCLUS

FASE 1, 2 EN 3 - TOT EN MET GESLOTEN EN OPEN VRAAGVORMEN

Deel 1

Harry Molkenboer

BTB

TOETSEN^{VOLGENS} DE TOETSCYCLUS

FASE 1, 2 EN 3 - TOT EN MET GESLOTEN EN OPEN VRAAGVORMEN

Deel **1**

Auteur

Ing. Harry Molkenboer

Toetsdeskundige, eigenaar van Bureau voor Toetsen & Beoordelen en redactielid van *EXAMENS Tijdschrift voor de toetspraktijk*.

Redactie

- Karin Gerritsen-van Leeuwenkamp MSc
Onderwijskundig adviseur bij de Academie Gezondheidszorg van Saxion en kenniskringlid bij het Kenniscentrum Onderwijsinnovatie van Saxion, buitenpromovendus bij het Welten-instituut van de Open Universiteit en redactielid van *EXAMENS Tijdschrift voor de toetspraktijk*.
- Dr. Desirée Joosten-ten Brinke
Lector Kwaliteit van Toetsen en Beoordelen, Fontys Lerarenopleiding Tilburg, Universitair hoofddocent Open universiteit en hoofdredacteur *EXAMENS Tijdschrift voor de toetspraktijk*.

Critical friends

- Drs. Monique Altemühl-Booltink
Onderwijskundige/toetsdeskundige, eigenaar MAB Onderwijskundige expertise.
- Drs. Manon Bonefaas-Weijenberg
Business development manager/(senior) consultant online examinering, Andriessen.
- Petra Theeuwen M SEN
Docent Engels vo en assessor bij Hogeschool Utrecht en Fontys Lerarenopleiding Tilburg.

INHOUDSOPGAVE DEEL 1

Voorwoord	5
Een omvangrijk boek	7
Inhoudsopgave deel 2 (onder voorbehoud)	11
1. Inleiding	13
1.1. Doel	13
1.2. Doelgroep	14
1.3. Kwaliteitspiramide en toetscyclus als basis voor toetskwaliteit	15
1.4. Leeruitlekomsten en indicatoren BKE	19
2. Toetstaal	21
2.1. Begrippen rond de vorm van een toetstaak	22
2.2. Begrippen rond de aard van een toetstaak	23
2.3. Begrippen rond toetsen en het begrip beoordelen	26
3. Toetsvormen	29
3.1. Ordening van toetsvormen op basis van beoordelingsbekwaamheden	29
3.2. Twee overzichten van toetsvormen	31
4. De toets in een breder perspectief	35
4.1. De toets als uitgangspunt voor het leren	35
4.2. Toetsprogramma en het formatief en summatief gebruik van toetsen	36
FASE 1 BESCHRIJVEN VAN DE CONTEXT VAN EEN TOETS	41
5. Beschrijven van de context van een toets	43
5.1. Beschrijven van de beroepscontext	43
5.2. Toetsbeleid	48
5.3. Toetsprogramma	50
5.4. Rol competenties	52
5.5. Rol onderwijsvisie	55
FASE 2 ONTWIKKELEN VAN TOETSTERMEN EN EEN TOETSMATRIJS	57
6. Ontwikkelen van toetstermen	59
6.1. De toetsterm in relatie tot eindkwalificaties en leerdoelen	59
6.2. Handreikingen voor het formuleren van toetstermen	64
6.3. Gebruik van een overzichtelijke structuur	67
6.4. Nuanceren bij het gebruik van toetstermen	72
7. Toepassen van een taxonomie	75
7.1. Wat is een taxonomie?	75
7.2. Een taxonomie kiezen	76
7.3. Taxonomie van Romiszowski	79
7.4. Taxonomie van Bloom	82

7.5.	Praktijkvoorbeeld met drie gedragsdomeinen	86
7.6.	Taxonomieën op programmaniveau	89
8.	Ontwikkelen van een toetsmatrijs	93
8.1.	Wat is een toetsmatrijs?	93
8.2.	Functies en voordelen van de toetsmatrijs	94
8.3.	Opbouw van een toetsmatrijs voor vragen	96
8.4.	Opbouw van een toetsmatrijs voor opdrachten	101
8.5.	Kiezen van een passende toetsvorm	104
8.6.	De paradox van de toetsmatrijs	107
FASE 3	ONTWIKKELEN VAN VRAGEN EN OPDRACHTEN	109
9.	Gebruiken van kwaliteitscriteria voor vragen en opdrachten	111
9.1.	Kwaliteitscriteria voor vragen en opdrachten	111
9.2.	Gebruik van kwaliteitscriteria	115
10.	Ontwikkelen van gesloten vragen	117
10.1.	Wat is een gesloten vraag en welke vormen komen veel voor?	117
10.2.	Opbouw van een gesloten vraag	118
10.3.	Voorbeelden van de belangrijkste gesloten vraagvormen	120
10.4.	Een gesloten vraag meet meer dan alleen reproduceerbare kennis	137
10.5.	Papieren of digitale afname van gesloten vragen	138
10.6.	Voor- en nadelen van gesloten vragen	142
10.7.	Ontwikkelaanpak en taalkundige aanbevelingen voor gesloten vragen	143
10.8.	Constructievoorschriften voor gesloten vragen	146
10.9.	Toepassen van constructievoorschriften	148
11.	Ontwikkelen van open vragen	221
11.1.	Wat is een open vraag?	221
11.2.	Voor- en nadelen van open vragen	224
11.3.	Opbouw van open vraagvormen	227
11.4.	Ontwikkelaanpak en taalkundige aanbevelingen voor open vragen	244
11.5.	Constructievoorschriften voor open vragen	246
11.6.	Formuleren van de vraag of kleine opdracht bij een open vraag	250
12.	Toetsvormen met gesloten en open vragen	263
12.1.	De casustoets	263
12.2.	De voortgangstoets	277
12.3.	De Script Concordance Test	291
12.4.	De mondelinge toets	304
Referenties		319
Over de auteur		325
Register figuren en tabellen		327
Index		331



INHOUDSOPGAVE DEEL 2 (ONDER VOORBEHOUD)

FASE 3 - ONTWIKKELEN VAN VRAGEN EN OPDRACHTEN (VERVOLG)

- 13. Ontwikkelen van opdrachten in de vorm van een product
- 14. Ontwikkelen van opdrachten in de vorm van een handeling
- 15. Ontwikkelen van opdrachten in de vorm van een gesprek

FASE 4 - SAMENSTELLEN TOETS

- 16. Kwaliteitscriteria voor een toets
- 17. Samenstellen van een toets

FASE 5 - AFNEMEN

- 18. Afnemen van een toets

FASE 6 - BEOORDELEN

- 19. Beoordelen van een toets
- 20. Beoordelaarsbetrouwbaarheid

FASE 7 - ANALYSEREN EN VASTSTELLEN RESULTATEN

- 21. Analyseren en vaststellen van resultaten

FASE 8 – EVALUEREN, RAPPORTEREN EN VERBETEREN

- 22. Evalueren, rapporteren en doorvoeren van verbeteringen

AANVULLEND

- 23. Gebruiken van een toetsprogramma
- 24. Toepassen van toetsbeleid



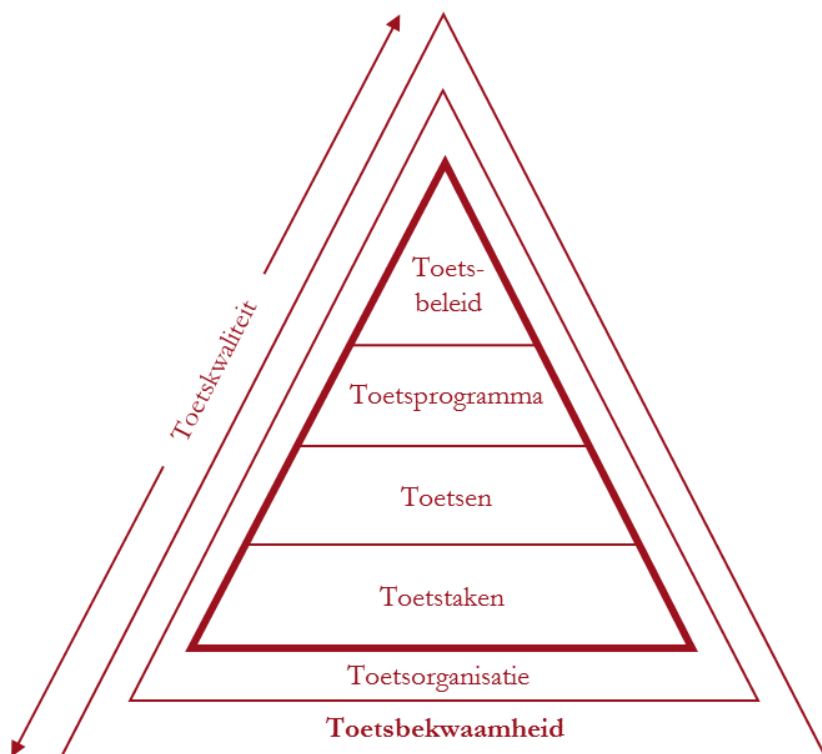
1.3. Kwaliteitspiramide en toetscyclus als basis voor toetskwaliteit

Het boek sluit aan bij alle leeruitkomsten en indicatoren die van toepassing zijn binnen de Basiskwalificatie examinering (zie Expertgroep BKE/-SKE, 2013), en is daarmee geschikt als voorbereiding op een BKE-certificeringstraject.

In dit boek worden conform het rapport van de Expertgroep BKE/SKE twee constructen als uitgangspunt genomen:

- de kwaliteitspiramide voor toetsen en beoordelen (Joosten-ten Brinke, 2011; Sluijsmans, Peeters, Jakobs & Weijzen, 2012);
- een toetscyclus met bijbehorende producten.

De kwaliteitspiramide voor toetsen en beoordelen



Figuur 1.1 Kwaliteitspiramide voor toetsen en beoordelen (Joosten-ten Brinke, 2011; Sluijsmans, Peeters, Jakobs & Weijzen, 2012)

Kwaliteitspiramide vat alles voor toetskwaliteit samen

De kwaliteitspiramide in figuur 1.1 bestaat uit twee delen. De twee buitenste lagen (Toetsbekwaamheid en Toetsorganisatie) zijn voorwaardelijk om tot goede toetskwaliteit te komen. Toetsbekwaamheid gaat over de toetsdeskundigheid van alle betrokkenen binnen het toetsproces. Toetsorganisatie gaat over wat nodig is om goede toetsen te ontwikkelen, af te nemen, te beoordelen, te analyseren en te verbeteren. Bij de vier gestapelde binnenlagen gaat het om de te onderscheiden aspecten die ieder op orde moeten zijn om te komen tot toetskwaliteit. De vier binnenlagen staan hiërarchisch tot elkaar, waarbij de kwaliteit van elke laag bijdraagt aan de totale toetskwaliteit. De laag Toetstaken betreft de afzonderlijke vragen (gesloten en open) en opdrachten. De laag Toetsen betreft de toetsen bedoeld voor de afname, beoordeling, analyse en verbetering. De laag Toetsprogramma verbindt de losse toetsen onderling en met het curriculum van een opleiding. De laag Toetsbeleid verbindt de toetsen met de visie op onderwijs die bestaat ten aanzien van leren en toetsen en met aspecten van buiten die inwerken op de wijze van toetsen.

De kwaliteitspiramide laat zien dat kwaliteit van toetsing een breed begrip is. De piramide biedt de mogelijkheid om vanuit verschillende perspectieven op toetskwaliteit de lagen onder de loep te nemen (Joosten-ten Brinke, 2011; Sluijsmans, Peeters, Jakobs & Weijzen, 2012).

De toetscyclus met bijbehorende producten

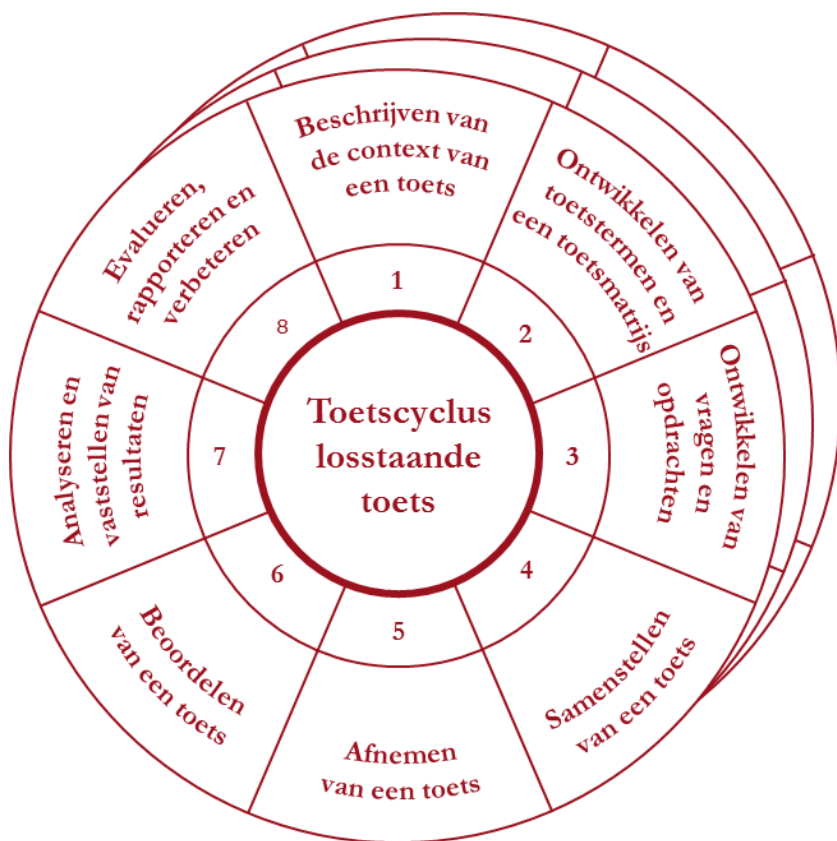
Een toetscyclus is een model dat de belangrijkste fasen beschrijft die nodig zijn om goede toetsen te ontwikkelen, af te nemen en de kwaliteit ervan te borgen. Dit boek gaat uit van een toetscyclus van acht fasen.

Toetscyclus van acht fasen

Door een toetscyclus van acht fasen te gebruiken, ontstaat een goed en logisch onderscheid tussen verschillende fasen om tot kwalitatief goede toetsen te komen; toetsen die zoveel mogelijk valide en betrouwbaar zijn. De acht fasen van de toetscyclus zijn voor elke toetsvorm toepasbaar. De kwaliteit van een toets is afhankelijk van de kwaliteit van uitvoering binnen elke fase van de toetscyclus. De kwaliteit van de gehele toets is daarmee zo hoog als die van de zwakste schakel. Als bijvoorbeeld de toetstermen niet goed zijn ontwikkeld in fase twee, heeft dat gevolgen voor alle volgende fasen.

De toetscyclus met acht fasen ziet er als volgt uit:



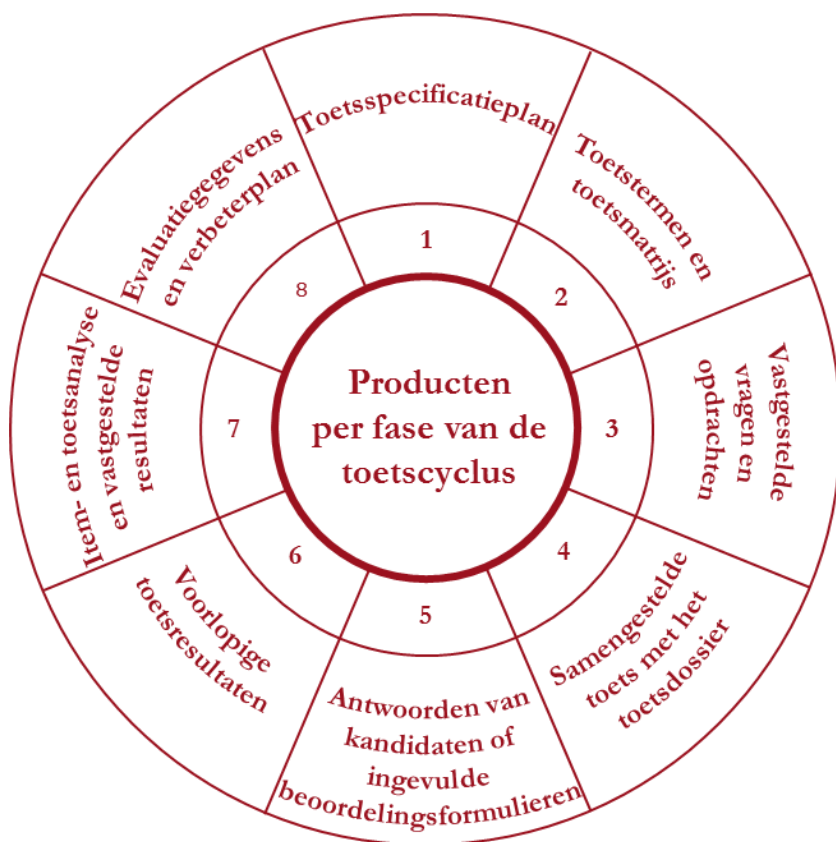


Figuur 1.2 Toetscyclus per toets

In figuur 1.2 stelt elke cirkel een losstaande toets voor. De twee achterliggende cirkels in figuur 1.2 geven weer dat er een verbinding kan bestaan tussen afzonderlijke toetsen. Bij een toets als onderdeel van een toetsprogramma zal een toets ook een relatie hebben met één of meer andere toetsen in het programma. Zo zal bijvoorbeeld een op kennis gerichte toets gerelateerd kunnen zijn aan een opdracht waarbij de kennis moet worden toegepast. Die toetsen zullen dan bij voorkeur op elkaar zijn afgestemd. Maar zo kan een toets binnen de ene onderwijseenheid ook leerstof omvatten van een andere onderwijseenheid vanwege de integratie van de twee afzonderlijke onderwijseenheden. Dit zal als het goed is zichtbaar worden binnen het toetsprogramma en bij de beschrijving van de context van de toets.

Het doorlopen van de toetscyclus levert per fase verschillende toetsproducten op (figuur 1.3). De producten maken de toetscyclus concreet en

de uitvoering beoordeelbaar, waarmee de kwaliteit van toetstaken, toetsen en beoordelingen geborgd kan worden.

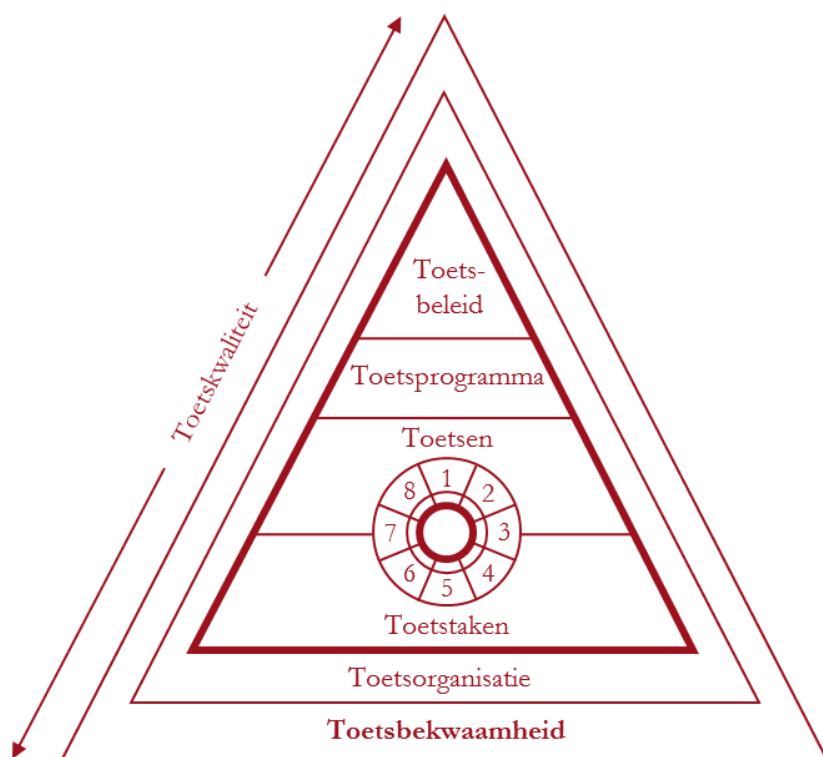


Figuur 1.3 Producten per fase van de toetscyclus

Toetscyclus als onderdeel van de kwaliteitspiramide

Toetscyclus is onderdeel van de kwaliteitspiramide

Figuur 1.4 laat zien dat de toetscyclus eigenlijk een vanzelfsprekend onderdeel is van de kwaliteitspiramide. De toetscyclus geeft inhoud aan de uitwerking van de lagen Toetstaken en Toetsen. De figuur die daarvoor ontstaat, omvat de essentie waar het bij het ontwikkelen, afnemen, beoordelen, analyseren en verbeteren van kwalitatieve toetsen om gaat.



Figuur 1.4 De toetscyclus als bouwsteen binnen de kwaliteitspiramide voor toetsen en beoordelen

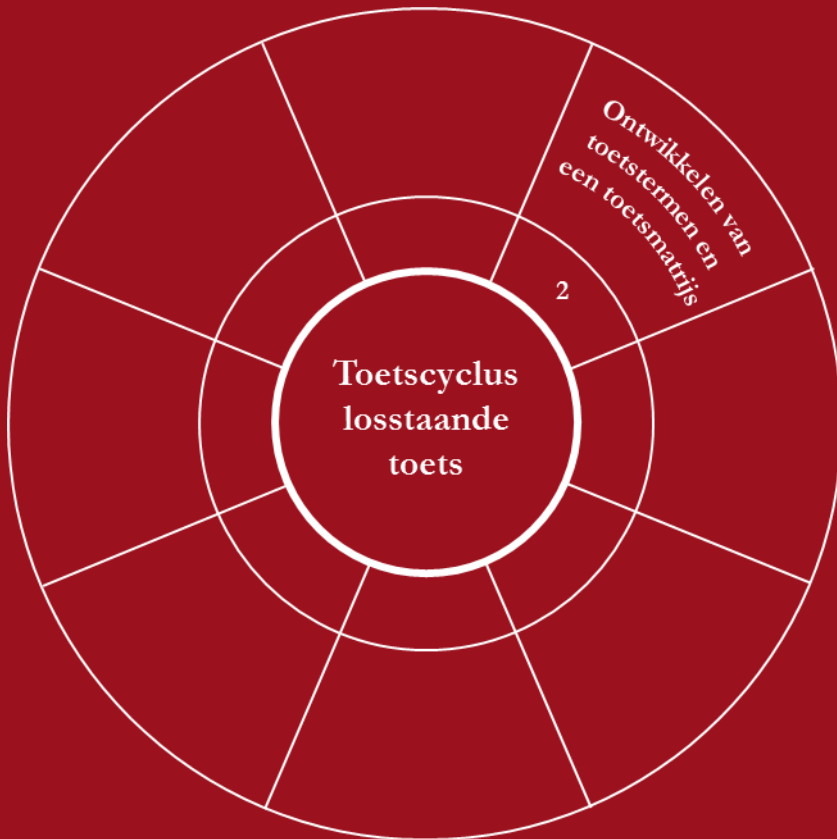
1.4. Leeruitkomsten en indicatoren BKE

In het rapport van de Expertgroep BKE/SKE worden zowel voor BKE als voor SKE leeruitkomsten en indicatoren beschreven. De indicatoren maken duidelijk wat van een BKE-/SKE-er toetstechnisch wordt verwacht. Aan de hand van toetstechnisch producten waar hij binnen de hogeschool mee te maken heeft, moet de BKE-/SKE-er met eigen toetsproducten aantonen dat hij de betreffende indicatoren beheerst. Voor indicatoren waar hij niet of minder mee te maken heeft, moet hij aannemelijk maken dat hij die adequaat uit zou kunnen voeren.

Om de toetstechnische producten bekwaam te realiseren of om aannemelijk te maken dat ze adequaat uitgevoerd zouden kunnen worden, moet de BKE-/SKE-er beschikken over toetstechnische kennis. Dit boek en deel 2 bieden de BKE-er alle informatie die daarvoor nodig is.

FASE 2

ONTWIKKELEN VAN TOETSTERMEN EN EEN TOETSMATRIJS



7. TOEPASSEN VAN EEN TAXONOMIE

Het bepalen van het juiste niveau van een vraag of een opdracht is een moeilijk aspect bij het ontwikkelen. Een taxonomie biedt hierbij ondersteuning. Met een taxonomie worden leerdoelen/toetstermen geïnclassificeerd in niveaus van toenemende complexiteit.

In paragraaf 7.1 wordt toegelicht wat een taxonomie is en wat de voordelen zijn van het gebruik. Paragraaf 7.2 gaat over het selecteren van een passende taxonomie. In paragraaf 7.3 en 7.4 worden de twee meest gebruikte taxonomieën binnen het hoger onderwijs, die van Romiszowski en Bloom, toegelicht. Van Bloom worden zijn taxonomie voor het cognitieve, het affectieve en psychomotorische domein uitgewerkt. Aan de hand van een integratief praktijkvoorbeeld wordt in paragraaf 7.5 duidelijk hoe de drie domeinen in verband met elkaar gebracht kunnen worden. Tot slot worden in paragraaf 7.6 nog twee andere taxonomieën weergegeven. Dit betreffen taxonomieën op programmaniveau, het HvA-model en het ZelCommodel.

7.1. Wat is een taxonomie?

Definitie

Taxonomie

Een taxonomie is een structuur voor het hiërarchisch indelen van objecten en abstracte elementen in groepen om ze te classificeren, te ordenen of te rangschikken.

Denk bij een taxonomie bijvoorbeeld aan de ordening van het dierenrijk en plantengemeenschappen.

Bij toetsing gaat het om het classificeren van leerdoelen in niveaus van complexiteit binnen het cognitieve, het affectieve of het psychomotorische domein. Hierbij moet elk lager niveau worden beheerst voordat een hoger niveau ontwikkeld kan worden.

De betekenis van het cognitieve, affectieve en psychomotorische domein is:

- *Cognitieve domein: gericht op kennis, intelligentie.*
- *Affectieve domein: gericht op gevoel, emotie, houding en sociale, reactieve en interactieve vaardigheden.*
- *Psychomotorische domein: gericht op fysieke vaardigheden en vaardigheden zoals presenteren en notuleren.*

Bij een leerdoel is een zuiver onderscheid tussen de domeinen soms niet goed te maken. Denk bijvoorbeeld aan het kunnen schrijven en presenteren van een beleidsnotitie. Daarvoor zijn alle drie de domeinen van toepassing. Als het goed is, zal dat leiden tot de inzet van verschillende toetsvormen die met elkaar de ontwikkelde kennis en inzicht, vaardigheden en houdingen vaststellen.

Het gebruik van een taxonomie heeft verschillende voordelen:

- Een taxonomie stimuleert om bij het formuleren van leerdoelen niet alleen vanuit de inhoud te denken, maar ook na te denken over de vorm van het gewenste gedrag (cognitief, affectief en psychomotorisch) en het niveau van dat gedrag.
- De niveaus van een taxonomie en de taxonomiedomeinen bieden een handvat voor het formuleren van de leerdoelen/toetstermen. Ook de suggesties voor werkwoorden die bij een niveau worden gegeven (zie tabel 7.2 en de figuren 7.7, 7.6 en 7.7) bieden een handvat.
- De keuze voor het domein in combinatie met het niveau geeft een indicatie voor de te gebruiken toetstaak.
- Een taxonomie biedt mogelijkheden om na een toets te onderzoeken waarmee studenten moeite hebben, met bijvoorbeeld meer op kennis gerichte vragen of meer op inzicht of toepassingsgerichte vragen.
- Binnen een leerlijn biedt de taxonomie een handvat om de opbouw in niveau inhoud te geven mede in combinatie met de integratie van domeinen. Denk aan een opdracht waarin zowel cognitieve, affectieve en psychomotorische aspecten op een hoog niveau integratief worden getoetst. Maar het biedt ook een handvat om discrepanties te ontdekken in doorlopende leerlijnen.
- Een taxonomie biedt leden van een examencommissie en externe auditoren handvatten voor het beoordelen van het niveau van afzonderlijke toetsen en hun onderlinge relatie alsmede voor de toetsing als geheel. Hierbij speelt de didactische en toetstechnische onderbouwing voor de keuze van de taxonomie met de eventuele domeinen en de niveaus, vanzelfsprekend een rol.

7.2. Een taxonomie kiezen

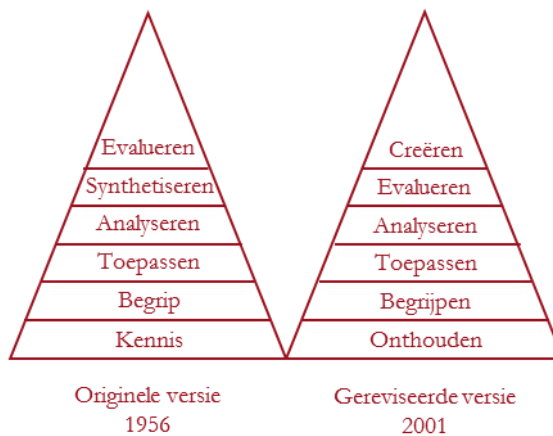
De keuze voor een taxonomie voor leren en toetsen is afhankelijk van de omgeving waarin men verkeert. Er zijn verschillende taxonomieën voor leren en toetsen ontwikkeld. De meest bekende zijn die van Bloom (1956), de gereviseerde versie ervan (Anderson, Krathwohl & Bloom, 2001) en de taxonomie van Romiszowski (1981).

lastig, hoewel de voorbeeldwerkwoorden een goede ondersteuning bieden. Romiszowski wordt daardoor vaak alleen gebruikt om leerdoelen te classificeren op het tweede niveau (feitelijke kennis, begripsmatige kennis, reproductieve vaardigheid en productieve vaardigheid).

7.4. Taxonomie van Bloom

Het cognitieve domein van Bloom

De taxonomie van Bloom is voornamelijk bekend vanuit het cognitieve domein. Hiervan bestaan inmiddels twee varianten: de oorspronkelijke versie (1956) en de gereviseerde versie (2001).

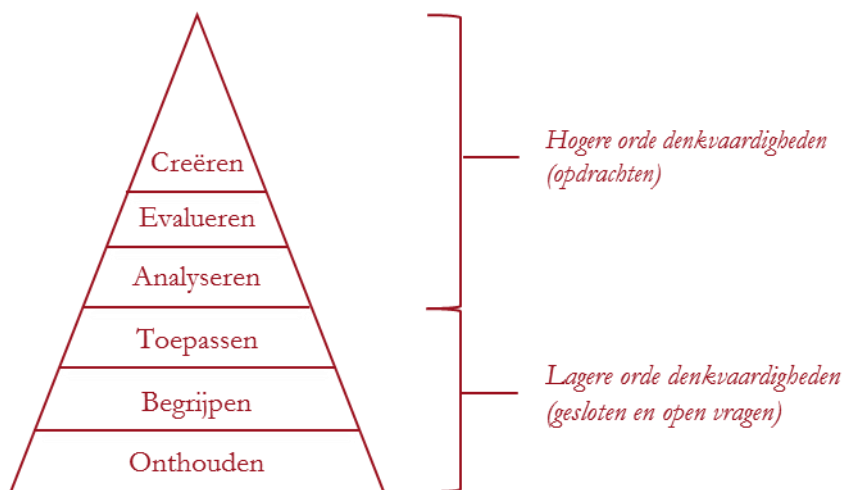


Figuur 7.2 De twee varianten van de cognitieve taxonomie van Bloom

Het verschil van de gereviseerde versie ten opzichte van de originele versie bestaat uit het volgende:

- Alle niveaus hebben een werkwoord als aanduiding gekregen.
- *Synthetiseren* is qua aanduiding veranderd in *Creëren* en is het hoogste cognitieve niveau geworden.
- *Evalueren* is een niveau gezakt.

De taxonomie kent een oplopende schaal van lagere naar hogere denkvaardigheden. Denkvaardigheden zijn cognitieve processen die zich afspelen binnen de verschillende niveaus.

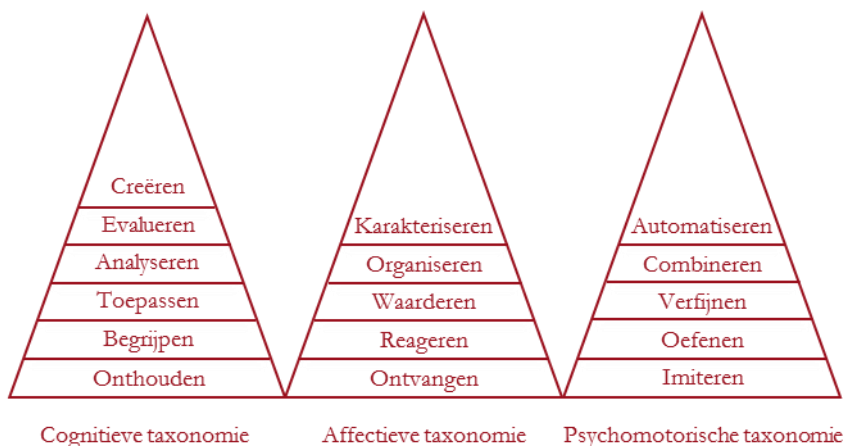


Figuur 7.3 Verdeling lagere en hogere orde denkvaardigheden

Lagere orde denkvaardigheden kunnen worden getoetst met vragen. Voor het toetsen van hogere orde denkvaardigheden lenen opdrachten zich beter. De *Piramide van Miller* (zie figuur 2.1), komt de verdeling van de toetsvormen in hogere en lagere orde denkvaardigheden ook voor. Een hoofdtoetsvorm kan goed worden gekozen op basis van Miller. Op basis van de taxonomie kan de toetsvorm verder gepreciseerd worden.

Het affectieve en psychomotorische domein van Bloom

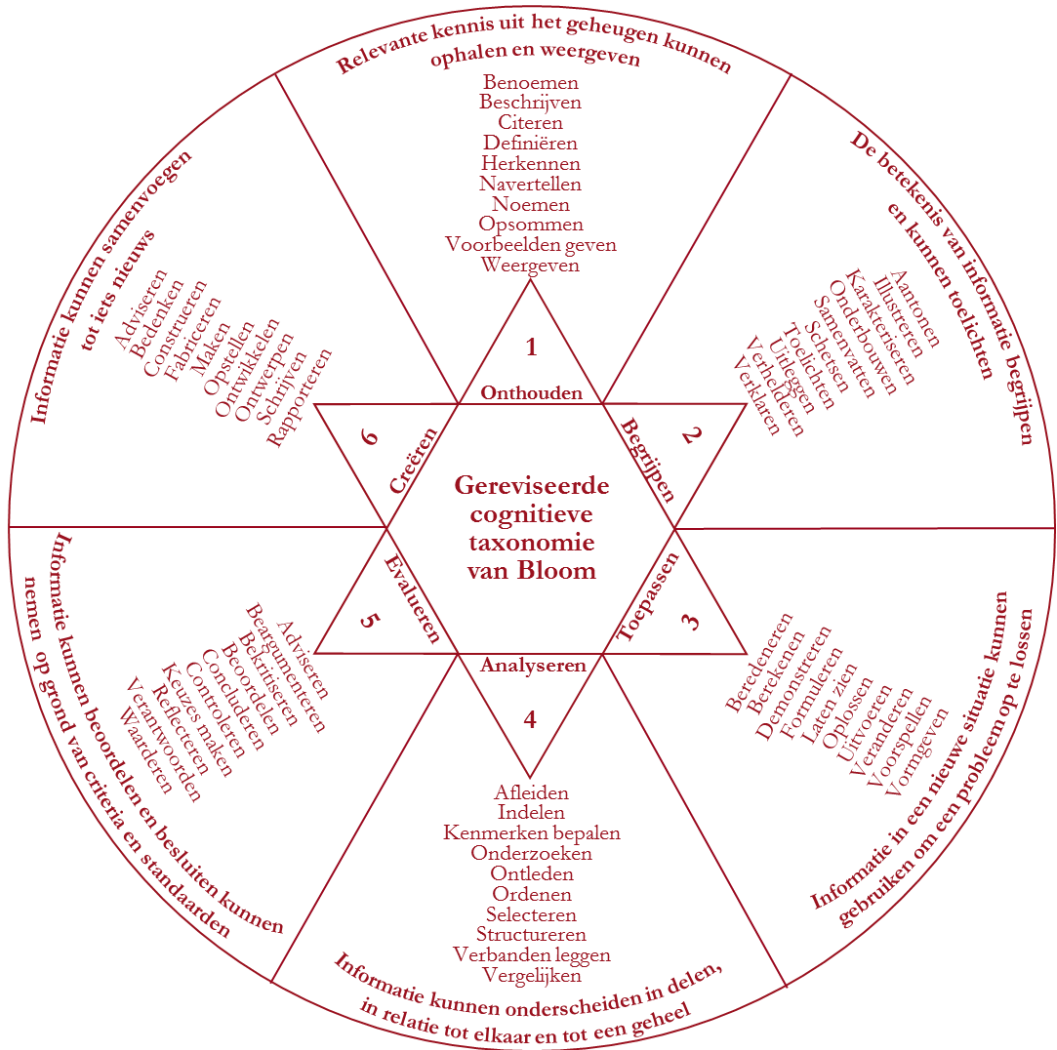
In figuur 7.4 is de affectieve taxonomie van Krathwohl, Bloom & Masia (1964) en de psychomotorische taxonomie van Dave (1967/1970) naast de cognitieve taxonomie van Bloom weergegeven.



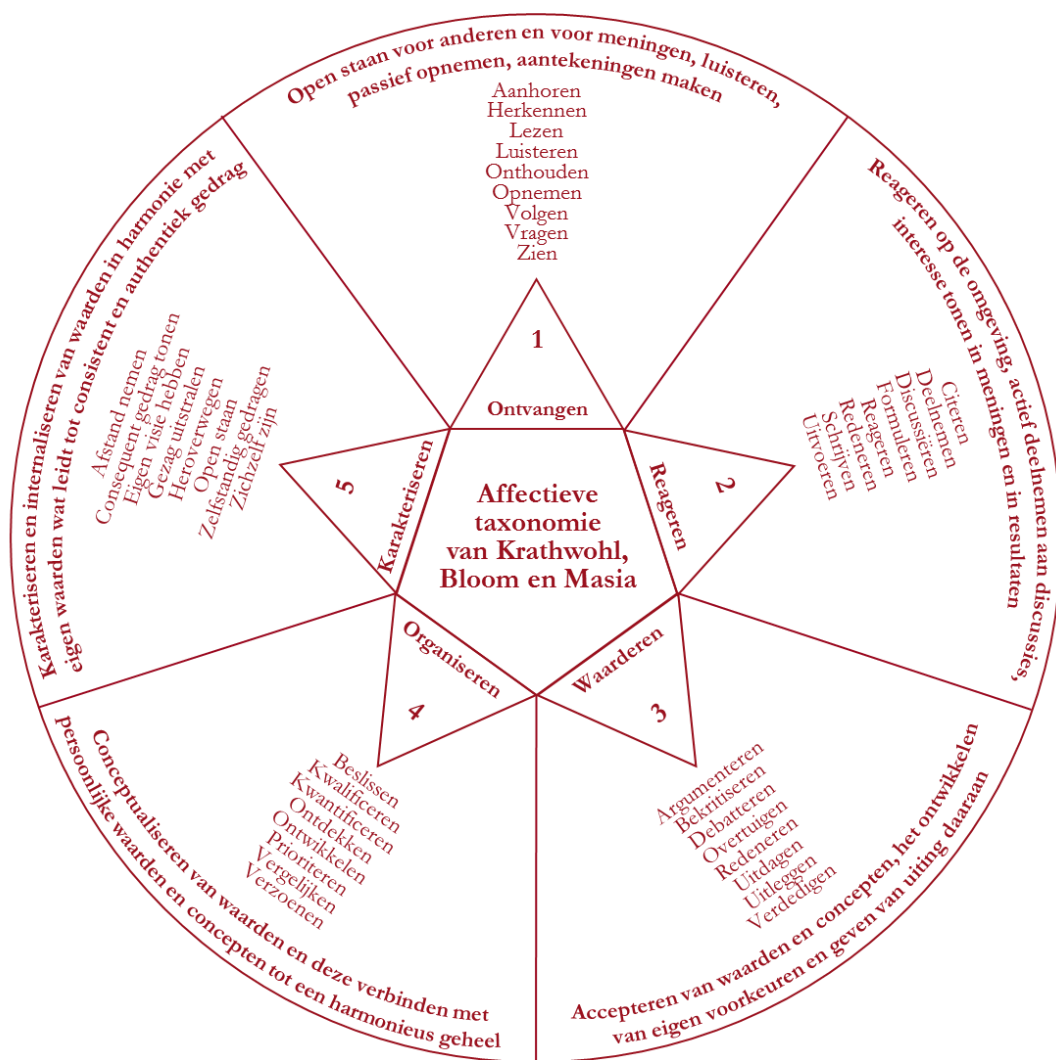
Figuur 7.4 De drie domeintaxonomieën

Beschrijvingen en werkwoorden

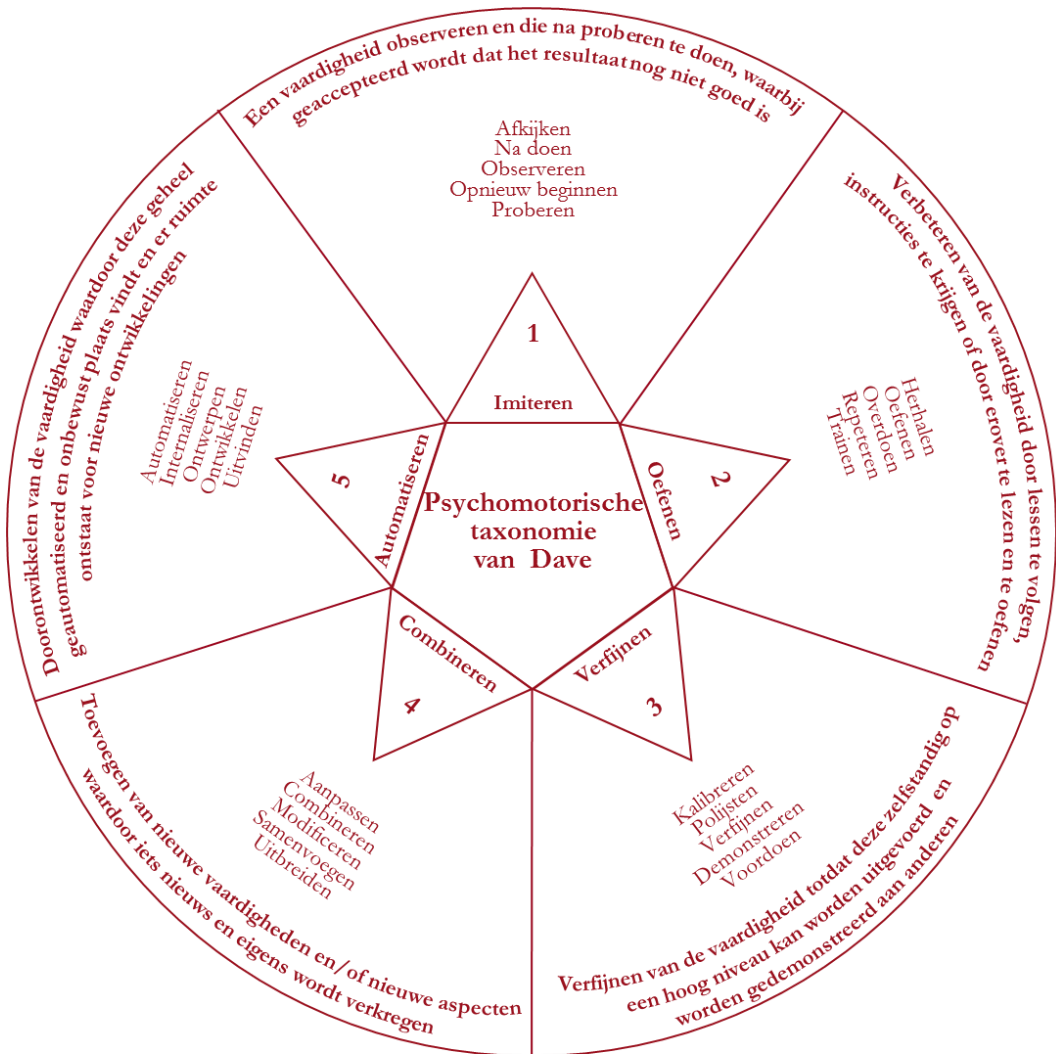
Figuren 7.5, 7.6 en 7.7 beschrijven per domein de taxonomie-niveaus en geven voorbeelden van waarneembare en actieve werkwoorden die passen bij een niveau.



Figuur 7.5 Beschrijvingen en werkwoorden behorende bij de gereviseerde cognitieve taxonomie van Bloom



Figuur 7.6 Beschrijvingen en werkwoorden behorende bij de affectieve taxonomie van Krathwohl, Bloom en Masia



Figuur 7.7 Beschrijvingen en werkwoorden behorende bij de psychomotorische taxonomie van Dave

7.5. Praktijkvoorbeeld met drie gedragsdomeinen

Bij kwalificaties zijn de domeinen vaak met elkaar verweven. Aan de hand van een voorbeeld uit het domein van watermanagement wordt dat duidelijk gemaakt. In dit voorbeeld wordt per domein (cognitief, affectief en psychomotorisch) en daarbinnen per gedragsniveau een bijpassend leerdoel uitgewerkt. Dit wordt gedaan aan de hand van de volgende -casus:

- Voor het cognitieve domein gaat het om een artikel als product; wanneer is een artikel een goed artikel.
- Voor het affectieve domein gaat het om het proactief deelnemen aan een gastcollege.
- Voor het psychomotorische domein moet de studenten een presentatie geven en hun mening verdedigen.

Dit leidt tot een toets waarin in een geïntegreerde omgeving wordt gewerkt aan het uitdiepen van kennis over een bepaald onderwerp (cognitief) waarover de studenten een eigen visie moeten ontwikkelen door actief deel te nemen aan het gastcollege (affectief). Dit moet leiden tot een goed artikel (cognitief), waarmee ze leren wat schriftelijk rapporteren betekent (kennis) en wat ze moeten presenteren (psychomotorisch) en verdedigen (affectief).

De leerdoelen voor het affectieve en psychomotorische domein zijn verwoord als acties die leiden tot een bepaald gedrag. Beoordelingscriteria moeten worden ontwikkeld, om dat gedrag meetbaar te maken. Dat is hier niet gebeurd, zie hiervoor deel 2.

Voorbeeld *Het cognitieve domein: wanneer is een artikel een goed artikel?*

Gedragsniveau	Voorbeeld van een leerdoel
<i>Onthouden</i>	<i>De student kan benoemen aan welke eisen een inhoudelijk artikel voor een vakblad gericht op watermanagement moet voldoen.</i>
<i>Begrijpen</i>	<i>De student kan verklaren welke gevolgen de eisen hebben voor de omvang, de opbouw, de inhoud, de lay-out en de vormgeving van het artikel.</i>
<i>Toepassen</i>	<i>De student kan op basis van drie bestaande artikelen laten zien of de artikelen wel of niet voldoen aan de eisen.</i>
<i>Analyseren</i>	<i>De student kan ontleden wat de hoofd- en bijzaken zijn van een artikel.</i>
<i>Evalueren</i>	<i>De student kan vanuit verschillende dimensies een artikel beoordelen en suggesties doen voor verbeteringen.</i>
<i>Creëren</i>	<i>De student kan zelf een artikel ten behoeve van een vakblad over watermanagement schrijven, dat voldoet aan de eisen die aan een artikel worden gesteld</i>

Het affectieve domein: pro-actief deelnemen aan een gastcollege

Gedragsniveau	Voorbeeld van een leerdoel
<i>Ontvangen</i>	<i>De student volgt met aandacht het gastcollege en is er merkbaar bij betrokken.</i>
<i>Reageren</i>	<i>De student laat door de vragen die hij stelt, zien dat hij zich heeft voorbereid op het gastcollege door het vergaren en lezen van informatie over de autoriteit, het lezen van het betreffende artikel dat in het gastcollege wordt uitgediept en het lezen van enkele andere gerelateerde artikelen over het onderwerp. De student neemt deel aan discussies die tijdens het gastcollege ontstaan en reageert vanuit een betrokken houding proactief op wat door de autoriteit en/of medestudenten wordt gezegd.</i>
<i>Waarderen</i>	<i>De student laat blijken of hij het eens of oneens is met wat naar voren wordt gebracht en beargumenteert dat.</i>
<i>Organiseren</i>	<i>De student weegt op basis van alle informatie die naar voren is gekomen af waarin hij zich kan vinden in de visie van de autoriteit en waarin niet en kan dit helder en duidelijk benoemen.</i>
<i>Karakteriseren</i>	<i>De student is trots op zijn kijk op de visie van de autoriteit (eens, oneens of gedeeltelijk eens, dat maakt niet uit) en kan zijn visie enthousiast en onderbouwd weergeven. Hij staat daarbij open voor meningen van anderen en respecteert dat anderen een andere visie hebben. Hij is zelfs in staat zijn visie waar nodig aan te passen of aan te scherpen.</i>

Het psychomotorische domein: schrijven van een eigen artikel over de eigen visie met een presentatie en een verdediging

<i>Gedragsniveau</i>	<i>Voorbeeld van een leerdoel</i>
<i>Imiteren</i>	<i>De student schrijft naar analogie van het artikel van de autoriteit zelf een artikel op basis van de eigen visie en houdt daarin rekening met de eisen die aan een artikel worden gesteld.</i>
<i>Oefenen</i>	<i>Na beoordeling van het artikel door de docent past de student het artikel aan, conform de aanwijzingen van de docent. Waar de student het niet eens is met de aanwijzingen van de docent beargumenteert de student dat en overlegt dat met de docent, waarna het artikel wordt opgeleverd aan de docent.</i>
<i>Verfijnen</i>	<i>Na redactie door de docent voert de student de laatste wijzigingen door en voegt de student illustraties toe en zorgt voor een eigen opmaak alsof het artikel in een tijdschrift geplaatst gaat worden en levert het in voor een eindbeoordeling.</i>
<i>Combineren</i>	<i>De student ontwikkelt een presentatie over het artikel dat voldoet aan de eisen die aan presentaties worden gesteld, presenteert dit en verdedigt het.</i>
<i>Automatiseren</i>	<i>Een vergelijkbare opdracht vindt later nog een keer plaats als vervolg op een zelf gekozen onderwerp dat de student moet presenteren voor zijn medestudenten als ware hij zelf de autoriteit.</i>

7.6. Taxonomieën op programmaniveau

In de voorgaande paragrafen zijn taxonomieën gepresenteerd die gebruikt kunnen worden om een goede niveau-aanduiding te geven aan leerdoelen/toetstermen. Om een meer generieke uitspraak te doen over de opbouw van het niveau in een opleiding zijn twee taxonomieën op programmaniveau interessant. Dit betreffen het HvA-model (n.d.) en het ZelCommodel (Bulthuis, 2013). Beide taxonomieën benadrukken dat de student gedurende de opleiding steeds complexer en steeds zelfstandiger moet kunnen werken.

8. ONTWIKKELEN VAN EEN TOETSMATRIJS

De toetsmatrijs is het hart van de toetscyclus. Het vormt het resultaat van wat tijdens het ontwikkelen van het toetsspecificatieplan, de leerdoelen en het gebruik van de taxonomie is gerealiseerd. Op basis van de toetsmatrijs wordt de uiteindelijke keuze voor de toetsvorm bepaald alsmede de inhoud, het niveau en de omvang van de toets. Het is niet alleen het vertrekpunt voor de ontwikkelaars van vragen en opdrachten, maar ook voor de ontwikkelaars van leermiddelen.

In paragraaf 8.1 wordt de definitie van een toetsmatrijs gegeven. In paragraaf 8.2 wordt aangegeven wat de functies en de voordelen zijn van een toetsmatrijs. Paragraaf 8.3 en 8.4 gaan over de opbouw van een toetsmatrijs voor vragen respectievelijk voor opdrachten. In Paragraaf 8.5 wordt ingegaan op het kiezen van een toetsvorm. Tot slot wordt in paragraaf 8.6 (hopelijk) het bezwaar weggenomen dat sommigen hebben tegen een transparante en volledige toetsmatrijs.

8.1. Wat is een toetsmatrijs?

Definitie

Toetsmatrijs

Een toetsmatrijs is een blauwdruk van een toets. Het is een tabel waarin zichtbaar wordt hoe de vragen of beoordelingsaspecten (bij opdrachten) worden verdeeld over de leerdoelen en de gedragsniveaus.

Een toetsmatrijs geeft de systematische onderbouwing van een toets weer en maakt de validiteit van de toets zichtbaar, de juiste inhoud en op het juiste niveau (= inhoudsvaliditeit) .

Een toetsmatrijs wordt ook wel een specificatietabel genoemd.

Een goede toets kan niet zonder een toetsmatrijs worden ontwikkeld.

In een toetsmatrijs voor vragen worden de leerdoelen/toetstermen in een matrix uitgezet tegen het niveau van een leerdoel/toetsterm. Hiermee is goed vast te stellen op welk niveau de vragen en daarmee de toets zich bevinden. Het is ook een check of de leerdoelen/toetstermen passen bij het niveau dat ze beschrijven en of ze tot de juiste taxonomiekolom behoren. Door bij de leerdoelen/toetstermen aan te geven hoeveel vragen per leerdoel/toetsterm aan de orde moeten komen wordt duidelijk wat belangrijke leerdoelen/toetstermen zijn en wat de omvang van de gehele toets wordt.

Voor opdrachten ligt het abstractieniveau van de toetsmatrijs meestal hoger, maar dit is sterk afhankelijk van de aard van de opdracht. Bij opdrachten spelen de beoordelingscriteria een nadrukkelijke rol in de toetsmatrijs.

Uit de toetsmatrijs kan worden afgeleid welke vraag- op opdrachtvormen passend zijn (zie tabel 10.1 voor een overzicht van gesloten vraagvormen, tabel 11.1 voor een overzicht van open vraagvormen en tabel 3.1 voor een overzicht van toetsvormen).

Bij veel toets-
termen meerdere
toetsmatrijzen
gebruiken

Bij een grote onderwijseenheid met veel toetstermen kan het voorkomen dat de toetsmatrijs aanleiding geeft om verschillende toetsen in te zetten. Het is voor het overzicht beter om dan verschillende toetsmatrijzen te ontwikkelen, die met elkaar de hele onderwijseenheid afdekken. Dit doet zich ook voor als tot de onderwijseenheid zowel kennis als vaardigheid behoren waarbij het wenselijk is om die beide afzonderlijk te toetsen.

Een toetsmatrijs kan worden opgenomen in het toetsspecificatieplan, waardoor dit plan een steeds vollediger beeld geeft van de toets als geheel.

8.2. Functies en voordelen van de toetsmatrijs

Een toetsmatrijs heeft verschillende belangrijke en krachtige functies en voordelen voor vele betrokkenen bij een toets:

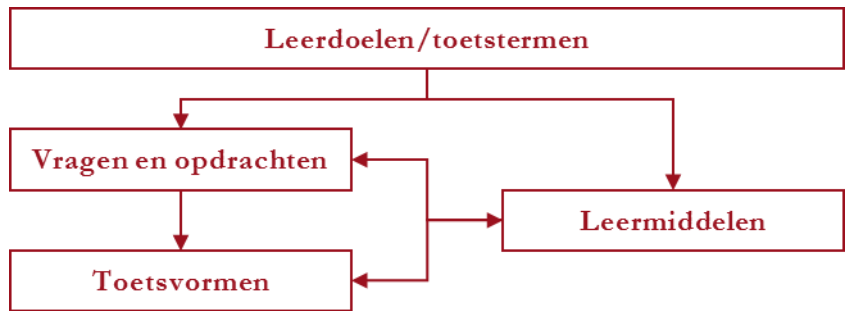
- Voor de student geeft de toetsmatrijs:
 - een concreet en eenduidig handvat voor het leren;
 - inzicht in wat getoetst kan worden, wat belangrijke onderwerpen zijn en wat de te verwachte toetsvorm is;
 - inzicht waarop wordt beoordeeld als het gaat om open vragen en opdrachten.
- Voor de ontwikkelaar van vragen en opdrachten geeft de toetsmatrijs:
 - precies weer hoeveel vragen, welke beoordelingsaspecten, voor welke leerdoelen/toetstermen en op welk niveau ontwikkeld moeten worden en hoe de onderlinge verdeling moet zijn;
 - handvatten voor de te gebruiken vraag- of opdrachtvormen en de toetsvorm;

- een goede basis om de inhoudsvaliditeit van een toets te vergroten, doordat een bewuste verdeling van de vragen/beoordelingsaspecten op basis van de leerdoelen/toetstermen wordt gerealiseerd en het niveau bewust wordt gekozen;
- de mogelijkheid om min of meer verschillende gelijkwaardige parallelle toetsen samen te stellen, bijvoorbeeld voor herkansingen of voor andere groepen, mits er voldoende vragen/opdrachten voorhanden zijn. Hiervoor wordt steeds dezelfde toetsmatrijs gebruikt, maar vragen/opdrachten zijn verschillend.
- de mogelijkheid om verschillende toetsen qua niveau met elkaar te vergelijken door vast te stellen hoe vaak bepaalde niveaus voorkomen binnen vergelijkbare toetsen;
- de basis om besprekingen te voeren met verschillende betrokkenen over de te ontwikkelen, gerealiseerde of afgenomen toets.
- Voor de curriculumontwikkelaar, de ontwikkelaar van leermiddelen en de docent geeft de toetsmatrijs:
 - handvatten voor het ontwikkelen van een didactisch ontwerp, de leermiddelen en het voorbereiden van lessen;
 - handvatten voor de structuur van de leermiddelen en de lessen door de structuur waarop leerdoelen/toetstermen/beoordelingsaspecten zijn geordend;
 - een indicatie hoe de student getoetst gaat worden.
- Voor de examencommissie en externe auditoren is de toetsmatrijs een instrument om te controleren of de gerealiseerde (en afgenomen) toets aansluit bij het ontwerp van de toets, of de toets valide is, de toetsvorm passend is en de toets de juiste omvang en het juiste niveau heeft.
- Voor de werkgever en de maatschappij geeft de toetsmatrijs weer wat werkgever/de maatschappij mag verwachten van een geslaagde student qua niveau en kennis en inzicht, vaardigheden en houdingen.

Leerdoelen/
toetstermen
geven richting
aan de toets-
én de leerom-
geving

Een belangrijk aspect van deze opsomming is dat deze laat zien dat leerdoelen/toetstermen zowel richting geven aan de toetsomgeving als aan de leeromgeving. Het ontwikkelen van toetstaken kan plaatsvinden, zonder dat er gebruik wordt gemaakt van leermiddelen. Dit verklaart ook waarom (vooral in de private sector) opleiding en examinering zonder problemen van elkaar gescheiden kunnen worden. Het is dan wel noodzakelijk dat de leerdoelen concreet, eenduidig en meetbaar, als toetstermen, zijn geformuleerd.





Figuur 8.1 Leerdoelen/toetstermen vormen de basis voor vragen, opdrachten, toetsvormen en leermiddelen

Leerdoelen zijn openbaar

Het is gebruikelijk dat in de private sector de leerdoelen worden ontwikkeld door een examen afnemende organisatie. De leerdoelen worden openbaar gemaakt voor de studenten en de opleiders. Hiermee wordt duidelijk dat de toets het leren bepaalt. In het reguliere onderwijs is dit niet altijd het geval en wordt een toets soms pas samengesteld aan het einde van een leerperiode op basis van de door de docent behandelde leerstof. Dit belemmert de transparantie. Het is wenselijk dat toetsen worden ontwikkeld op basis van vooraf vastgestelde en openbare leerdoelen.

Nadeel toetsmatrijs: het kost veel capaciteit

Tot slot van deze paragraaf een antwoord op de vraag of een toetsmatrijs ook nadelen kent? Het belangrijkste en ook grootste nadeel is dat het ontwikkelen van een toetsmatrijs, waartoe ook het ontwikkelen van leerdoelen en toetstermen behoort, arbeidsintensief en dus tijdrovend is. Dit is een belangrijke reden waarom leerdoelen en een toetsmatrijs (evenmin als een toetsspecificatieplan) niet worden ontwikkeld en direct wordt begonnen met het ontwikkelen van vragen/opdrachten. Management en ontwikkelaars van toetsen dienen zich te realiseren dat het ontwikkelen van toetsen veel capaciteit vergt. Beschikbare capaciteit zou goed verdeeld moeten worden over het ontwikkelen en verzorgen van de lessen en het ontwikkelen, afnemen, beoordelen, analyseren en verbeteren van toetsen.

8.3. Opbouw van een toetsmatrijs voor vragen

Een toetsmatrijs kan op verschillende manieren worden opgebouwd. De meest gebruikte vormen zijn op basis van vaste waarden, op basis van percentages of op basis van marges.

Toetsmatrijs op basis van vaste waarden

Een toetsmatrijs op basis van vaste waarden is de meest eenvoudige en transparante vorm van een toetsmatrijs. Deze vorm is alleen mogelijk als vooraf bekend is uit hoeveel vragen een toets moet bestaan. Deze toetsmatrijs laat zien:

- welke leerdoelen en toetstermen de basis vormen voor de toets;
- hoe de vragen over de leerdoelen en toetstermen precies worden verdeeld;
- op welk gedragsniveau de leerdoelen en toetstermen moeten worden uitgewerkt in vragen;
- uit hoeveel vragen de toets bestaat.

Tabel 8.1 geeft de basisvorm van een toetsmatrijs weer waarbij een deel van de taxonomie van Bloom wordt gehanteerd voor de niveauweergave.

Leerdoel	Toetsterm	Aantal Vragen	Onthouden	Begrijpen	Toepassen
1.	1.1	2	2		
	1.2	1		1	
2.	2.1	3	2	1	
	2.2	2	2		
	2.3	2	1	1	
3.	3.1	1		1	
	3.2	4	1	1	2
Totaal		15	8	5	2

Tabel 8.1 Basisvorm van een toetsmatrijs voor vragen op basis van de cognitieve taxonomie van Bloom

Toetsmatrijs op basis van percentages

In plaats van met vaste aantallen vragen te werken, kan ook met percentages worden gewerkt of de combinatie van aantallen vragen en percentages.

Leer- doel	Toets- term	Aantal vragen		Onthouden		Begrijpen		Toepassen	
1.	1.1 1.2	20%	2 1	13%	2	7%	1		
2.	2.1 2.2 2.3	46%	3 2 2	33%	2 2 1	13%	1 1		
3.	3.1 3.2	34%	1 4	7%	1	14%	1 1	13%	2
Totaal		100%	15	53%	8	34%	5	13%	2

Tabel 8.2 Eenvoudige toetsmatrijs met het aantal vragen waaruit de leerdoelen, toets termen en de toets bestaan en een overzicht van de verdeling van de vragen in percentages (percentages zijn afgerond om ze kloppend te maken)

De toetsmatrijzen in tabel 8.1 en 8.2 laten zien dat de toets bestaat uit 15 vragen, dat de nadruk ligt op Onthouden (53%), dat alle toets termen in de toets terug te vinden zijn en dat leerdoel 2 met 46% van de vragen het belangrijkste leerdoel is. Van de overige vragen gaat 20% over leerdoel 1 en 34% over leerdoel 3. Daarnaast blijkt dat toets term 3.2 belangrijk is. Niet alleen behoren de meeste vragen daartoe (4), maar het leerdoel wordt getoetst met één vraag op het niveau van Onthouden, één vraag op het niveau van Begrijpen en met twee vragen op het niveau van Toepassen.

De bovenstaande analyse moet in overeenstemming zijn met de hoeveelheid leerstof, de leertijd en/of het belang van het betreffende leerdoel in de opleiding. Door de toetsmatrijs met collega's te ontwerpen wordt voorkomen dat in een toets een stokpaardje van een docent te veel aandacht krijgt.

Toetsmatrijs op basis van marges

Een andere manier voor het samenstellen van een toetsmatrijs is werken met marges per toets term. Het eerste cijfer geeft het minimum aantal vragen aan en het tweede cijfer het maximum aantal vragen. In plaats hiervan kan gewerkt worden met minimum en maximum percentages.

Leerdoel	Toetsterm	Aantal Vragen	Onthouden	Begrijpen	Toepassen
1.	1.1	1-2	1-2	1-1	
	1.2	1-1			
2.	2.1	2-4	1-2	1-2	
	2.2	2-4	2-4	0-1	
	2.3	0-2	0-1		
3.	3.1	0-2	1-3	0-2	2-2
	3.2	3-6		0-1	
Totaal		15	8	5	2

Tabel 8.3 Toetsmatrijs met marges per toetsterm met het aantal vragen waaruit de toets bestaat

Let op! Bij een toetsmatrijs met marges moet de toetsmatrijs qua totalen zowel horizontaal als verticaal kloppen. Dat is bij de toetsmatrijs in tabel 8.3 het geval.

Werken met marges geeft voordelen

Werken met marges biedt meer mogelijkheden voor variatie bij het selecteren van geschikte vragen voor een toets. Het biedt binnen vastgestelde marges ook ruimte om te variëren met de belangrijkheid van leerdoelen en toetstermen om daarmee de samensteller van de toets persoonlijke ruimte geven.

Door de marges tussen het minimum en het maximum te beperken, wordt voorkomen dat de samensteller te veel vrijheid krijgt. Dit is duidelijk te zien bij toetsterm 1.2. Daar is zelfs geen enkele ruimte voor de samensteller overgelaten.

Bij de toetstermen 2.3 en 3.1 is wel veel ruimte voor de ontwikkelaars van de vragen. Het is bij deze twee toetstermen namelijk toegestaan om in de toets geen vragen over die toetstermen te stellen. Omdat een toetsmatrijs ook gebruikt wordt voor studenten als leidraad voor het leren, is het nu niet zo dat ze ervan uit kunnen gaan dat ze geen vragen over die toetstermen krijgen. Ze worden min of meer gedwongen de stof wel te bestuderen, zonder dat ze er vragen over gesteld hoeven te krijgen.

Er is nog een manier om ruimte te creëren bij het samenstellen van een toets. Dit kan zinvol zijn als er veel toetstermen zijn. Het is niet nodig om alle toetstermen te bevragen om vast te stellen of een student de

leerstof voldoende beheerst. Er wordt dan bijvoorbeeld alleen op leerdoelniveau aangegeven hoeveel vragen in de toets voor moeten komen. Dat deze methode minder transparant is, is een nadeel.

Leerdoel	Toetsterm	Aantal vragen	Onthouden	Begrijpen	Toepassen
A.	A.1 A.2 A.3 A.4 A.5 A.n	4	2	1	1
B.	B.1 B.n	7	5	2	
---	---	---	---	---	---
Totaal		40	25	10	5

Tabel 8.4 Voorbeeld toetsmatrijs waarbij alleen de leerdoelen worden gebruikt en niet de toetstermen

Een oplossing om wel te zorgen voor meer transparantie is om hier aan te geven wat de marge van het aantal vragen is.

Leerdoel	Toetsterm	Aantal vragen	Onthouden	Begrijpen	Toepassen
A.	A.1 A.2 A.3 A.4 A.5 A.n	4 0-1 0-1 1-3 0-1 1-2 ---	2 1-2 0-1 ---	1 0-1 1-2 ---	1 0-1 0-1 ---
B.	B.1 B.n	7 0-3 ---	2 --- ---	4 --- ---	1 --- ---
---	---	---	---	---	---
Totaal		40	25	10	5

Tabel 8.5 Voorbeeld toetsmatrijs waarbij leerdoelen en toetstermen worden gebruikt



Welke opzet en detaillering van de toetsmatrijs wordt gekozen kan worden vastgelegd in het toetsbeleid.

Het ontwikkelen van de toetsmatrijs geschiedt het beste in overleg en afstemming met collega's.

Weging wel of niet meenemen in een toetsmatrijs

Bij een toets met gesloten vragen wordt afgeraden om in een toetsmatrijs een kolom op te nemen waarin de weegfactor van een leerdoel/-toetsterm kan worden aangegeven. Het opnemen van een weegfactor is een andere manier om het belang van leerdoelen/toetstermen ten opzichte van elkaar weer te geven.

Gebruik liever meer vragen bij een belangrijk leerdoel dan een vraag met een hoge weegfactor

Als een leerdoel/toetsterm belangrijk is, is het beter om daarover meer vragen te stellen, dan bijvoorbeeld één vraag en die vraag qua score een hoge weegfactor mee te geven (bijvoorbeeld drie punten, terwijl de andere vragen één punt opleveren). Er kunnen dan beter drie vragen worden gesteld. Dit geldt zowel voor gesloten als open vragen. Een toets met meer vragen geeft de student meer mogelijkheden om zijn kennis te demonstreren. Bovendien kan de validiteit en de betrouwbaarheid van een toets toenemen als die meer vragen bevat.

8.4. Opbouw van een toetsmatrijs voor opdrachten

Een toetsmatrijs voor opdrachten wordt een opdrachtenontwerpschema genoemd (Joosten-ten Brinke & Draaijer, 2015).

Bij opdrachten gaat het om hogere orde denkvaardigheden. Ook zullen daarnaast affectieve en/of psychomotorische vaardigheden voorkomen. Op het niveau van hogere orde denkvaardigheden, affectieve en psychomotorische vaardigheden worden leerdoelen abstracter geformuleerd. Daarbij wordt er van uitgegaan dat de student lagere orde denkvaardigheden beheerst en kan toepassen. Die zijn, als het goed is, met andere toetsvormen zoals gesloten en open vragen al getoetst. Als het leerdoel is dat de student een interview kan voorbereiden en houden en er een goed artikel van tweeduizend woorden van kan maken voor een vakblad, dan wordt verondersteld dat hij bijvoorbeeld:

- weet hoe hij een interview moet voorbereiden;
- weet hoe hij een interview moet opbouwen;
- tijdens het gesprek de controle houdt, kan doorvragen, open vragen stelt, samenvat, informatie controleert, de structuur bewaakt, het gesprek op gang houdt, hoofd- en bijzaken scheidt et cetera;

- bij het schrijven grammaticaal juist en qua spelling foutloos te werk kan gaan, een goede opbouw van het artikel weet te realiseren, helder taalgebruik kan toepassen, concreet kan zijn, koppen goed kan gebruiken, et cetera.

Een opdrachten-ontwerpschema bevat beoordelingscriteria

Bij een opdrachtenontwerpschema komt naast het leerdoel en in plaats van een lijst met toetstermen, een lijst met (globale) beoordelingscriteria. De beoordelingscriteria zijn gekoppeld aan de leerdoelen. Van het leerdoel wordt de weging aangegeven en van de beoordelingscriteria de beoordelingsschaal.

Een groot voordeel van deze aanpak is dat beoordelingscriteria bij het leerdoel worden ontwikkeld en niet zoals veel voorkomt pas bij de ontwikkeling van de opdracht. Ontwikkelaars van opdrachten laten zich bij de beoordelingscriteria te veel leiden door de opdracht zelf en vergeten het leerdoel.

Een opdrachtenontwerpschema kan er als volgt uit zien. In dit geval is de taxonomie van het ZelCommodel gebruikt.

Leerdoel	Niveau*	Weging	Opdracht (volgnr.)	Beoordelingscriteria	Scoring*
1	C-Ch/Zl	30%	1	1.1 1.2 1.3 ---	S/O/V/G/U S/O/V/G/U S/O/V/G/U ---
2 en 5	C-Cm/Zm	20%	2	2.1 ---	S/O/V/G/U ---
3, 4, en 6	D-Ch/Zm	50%	3	3.1 ---	S/O/V/G/U ---
Totaal		100			

Tabel 8.6 Voorbeeld opdrachtenontwerpschema voor een opdracht

* De afkortingen in de toetsmatrjjs hebben de volgende betekenis:

- Bij *Niveau* geven C en D de niveaus weer die bepaald worden door:
 - C = complexiteit en Z = zelfstandigheid;
 - l = laag, m = middel en h = hoog.
- Bij *Scoring* hebben de letters de volgende betekenis: S = slecht, O = onvoldoende, V = voldoende, G = goed en U = uitstekend.

Op basis van het voorbeeld van het interview en het artikel zou het opdrachtenontwerpschema er als volgt uit kunnen zien.

Voorbeeld

De student kan een interview voorbereiden, houden en er een goed artikel van maximaal tweeduizend woorden over schrijven

Leerdoel	Niveau	Weging	Opdracht	Beoordelingscriteria (de student heeft ...)	Scoring
Kan een interview op een goede wijze voorbereiden	C-CI/Zh	20%	Interview voorbereiden	■ Een duidelijk doel voor het artikel gekozen en vastgelegd (20%).	S/O/V/G/U
				■ Relevante literatuur verzameld en geordend (10%).	S/O/V/G/U
				■ Literatuur bestudeerd en de belangrijkste punten samengevat (10%).	S/O/V/G/U
				■ Een goede vragenlijst opgesteld (30%).	S/O/V/G/U
				■ De vragenlijst logisch geordend in relatie tot het te voeren gesprek (10%).	S/O/V/G/U
				■ --- (20%).	---
Kan een interview op een professionele wijze houden	D-Cg/Zh	40%	Interview houden	■ De kennismaking op een persoonlijke wijze gedaan (5%).	S/O/V/G/U
				■ Het doel van het interview goed toegelicht (10%).	S/O/V/G/U
				■ Het interview op een natuurlijke wijze uitgevoerd (10%).	S/O/V/G/U
				■ Tijdens het interview overzicht over het interview gehouden (20%).	S/O/V/G/U
				■ Gecontroleerd of alle vragen uit de voorbereiding aan de orde zijn geweest (30%).	S/O/V/G/U
				■ Tijdens het interview samengevat van wat gezegd is en geconcludeerd (5%).	S/O/V/G/U
				■ Afgesproken dat het artikel eerst ter goedkeuring wordt voorgelegd (5%).	S/O/V/G/U
				■ --- (15%).	---

kunnen voeren, dan moet hij dat laten zien in een adviesgesprek. En als hij een beleidsnotitie moet kunnen schrijven dan is het onvoldoende dat hij weet aan welke eisen een beleidsnotitie moet voldoen.

Kennis en ervaring met verschillende toetsvormen noodzakelijk om een passende toetsvorm te selecteren	Om vast te kunnen stellen welke toetsvorm het best passend is, moet kennis omtrent, en bij voorkeur ook doorleefde ervaring met de verschillende toetsvormen, bestaan. Als onbekend is wat een Stationstoets of een Script Concordance Test is, kan ook niet bepaald worden of ze een passende toetsvorm kunnen zijn. Om vast te stellen of een Portfolio-opdracht op zich kan volstaan of dat daarnaast nog een Criterium gericht interview nodig is, dient naast de leerdoelen inzicht te bestaan in wat beide toetsvormen precies kunnen meten.
---	--

Een en ander klinkt logisch, maar het kiezen van een passende toetsvorm blijkt in de praktijk complexer dan het op eerste gezicht lijkt. Helemaal als uit een toetsmatrijs blijkt dat eigenlijk verschillende toetsvormen nodig zijn.

Haalbaarheid

Als uit het Toetsspecificatieplan blijkt dat een doelgroep omvangrijk is (dit is overigens relatief), kan blijken dat bijvoorbeeld het laten voeren van een adviesgesprek met een acteur en twee onafhankelijke beoordelaars te kostbaar wordt en te lang gaat duren waardoor de examens voor alle studenten niet passen binnen een examenplanning. Nog afgezien van kosten en planning kan blijken dat er niet voldoende acteurs of voldoende gekwalificeerde beoordelaars beschikbaar zijn. In de praktijk kan blijken dat de toetstechnisch best passende toetsvorm niet haalbaar is. Met behulp van het Toetsspecificatieplan, de leerdoelen/toetstermen, de toetsmatrijs en in overleg met collega's en anderen dient een wel haalbaar alternatief te worden ontwikkeld. Wellicht wordt de acteur vervangen door een collega en worden twee beoordelaars vervangen door één of krijgen medestudenten een adviserende beoordelaarsrol? Of wordt een digitale casus ontwikkeld, en moet de student in een video-opname via een webcam laten zien hoe hij een klant op basis van de casus te woord staat.

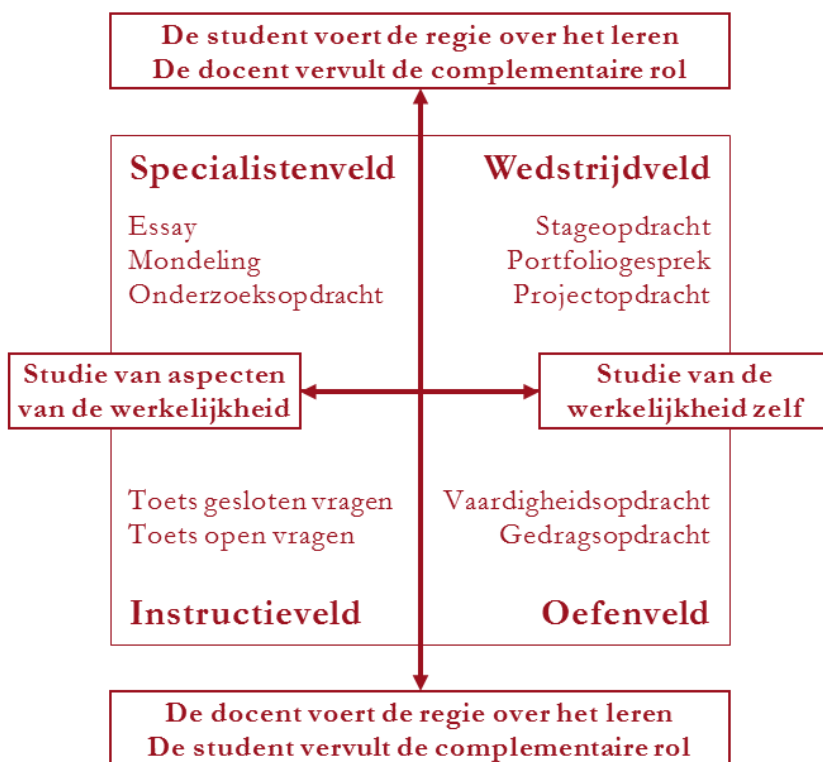
De toetsmatrijs is bepalend voor de toetstechnisch best passende toetsvorm, maar niet alléén bepalend voor de uitvoering daarvan. Didactische en toetstechnische creativiteit kunnen bijdragen aan alternatieve toetstechnische verantwoorde oplossingen binnen toetsvormen.



Rol onderwijsvisie

De onderwijsvisie kan invloed hebben op de te gebruiken toetsvormen. Er zijn opleidingen die ervoor kiezen om reproduceerbare kennis niet expliciet te meten, maar onderdeel te maken van toetsen die hogere orde denkvaardigheden meten. De gedachte daarbij is dat die hogere orde denkvaardigheden alleen bekwaam uitgevoerd kunnen worden als de lagere orde vaardigheden beheerst worden. De betreffende visie op leren en toetsen impliceert dat kennistoetsen niet of minder als toetsvorm gebruikt zullen worden. Dit moet terug te vinden zijn in het toetsbeleid en heeft ook gevolgen voor het Toetsspecificatieplan. Mogelijk niet voor de leerdoelen, want kennis moet nog wel geleerd worden, maar wel voor de toetstermen en dus voor de toetsmatrijs.

Nieweg (2013) heeft het Vierveldenmodel ontwikkeld om daarmee de aard van leerdoelen nader te bepalen en, op basis hiervan, de toetsvorm te kiezen. Het model laat zien hoe sterk een onderwijsvisie mede bepalend kan zijn voor de keuze van een (hoofd)toetsvorm en daarmee van invloed is op het Toetsspecificatieplan en uiteindelijk dus ook op de toetsmatrijs.



Figuur 8.2 Vierveldenmodel met geadviseerde toetsvormen

8.6. De paradox van de toetsmatrijs

In de praktijk blijkt dat ontwikkelaars van een toetsmatrijs moeite hebben met het gebruik van transparante en gedetailleerde toetsmatrijzen zoals in de basisvorm is weergegeven. Ze hebben het gevoel dat zo wordt weggegeven waar de toets precies over gaat, wat toetsgericht studeren in de hand werkt. Dit is echter precies de bedoeling van de toetsmatrijs. Alleen moet dat niet worden gezien als toetsgericht studeren.

Een toetsmatrijs maakt het onderwijs transparant De toetsmatrijs beschrijft precies wat een student moet kennen en kunnen, op welk niveau en wat de belangrijke onderwerpen zijn. Een toetsmatrijs maakt het onderwijs juist transparant. Bovendien worden de vragen zelf niet weggegeven.

Als een docent onderwerpen belangrijk vindt die niet in de toetsmatrijs staan, kunnen er twee zaken aan de hand zijn:

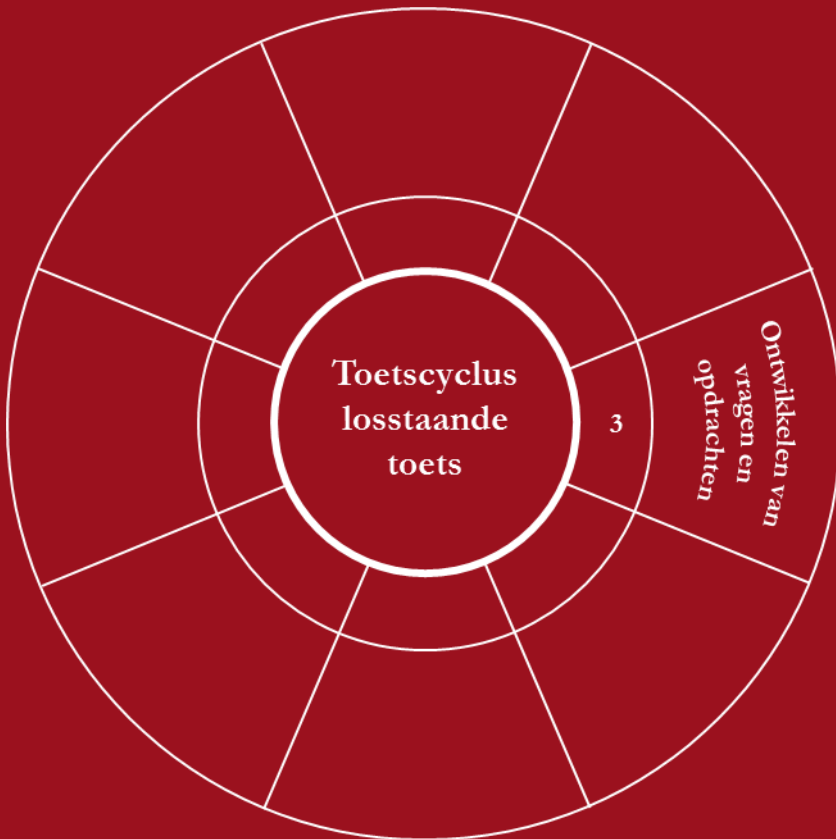
1. De leerdoelen/toetstermen zijn niet compleet. In dat geval moet de toetsmatrijs worden bijgesteld.
2. De docent heeft blijkbaar een eigen stokpaardje. Die mag hij in zijn lessen behandelen en hierbij aangeven waarom hij dat belangrijk vindt, maar het behoort niet tot de te toetsen leerstof. De docent kan in de les zijn verhaal toch kwijt en studenten die het belangrijk vinden, hebben de stof dan wel tot zich genomen.

Om het gevoel weg te nemen dat de toetsmatrijs alles weggeeft wat getoetst wordt, kan de toetsmatrijs met marges uit tabel 8.3 of 8.5 worden ingezet.



FASE 3

ONTWIKKELEN VAN VRAGEN EN OPDRACHTEN



Constructievoorschrift	24
De alternatieven zijn qua lengte, woordgebruik en formulering vergelijkbaar (homogeniteit)	
Kwaliteitseis	Specificiteit

Belang

Alle alternatieven lijken op elkaar, zodat uit de lengte, het woordgebruik en de formulering niet toevallig de sleutel is op te maken.

Toelichting

Nadat de sleutel en de alternatieven zijn ontwikkeld, kan het probleem van homogeniteit relatief eenvoudig worden opgelost, door de afleiders aan te passen en ze te laten lijken op de sleutel.

Valkuil

Ongemerkt wordt de sleutel vaak nauwkeuriger geformuleerd dan de afleiders, omdat die helemaal goed dient te zijn. Hierdoor wordt de sleutel vaak het langste alternatief.

Voorbeelden

Het eerste voorbeeld spreekt voor zich. Het tweede voorbeeld laat zien dat als de alternatieven juist te sterk op elkaar lijken, dat een aanwijzing kan zijn in de richting van het goede antwoord.

Voorbeeld

Niet

Wanneer is sprake van hypoglykemie?

- A. Als bij de patiënt de waarde van bloedglucose daalt onder 4,0 - 3,5 mmol/l en de patiënt begint te zweten, te trillen en zich slecht voelt.**
- B. Als sprake is van een te laag suikergehalte in het bloed.*
- C. Als men niet meer aanspreekbaar is en wegvalt.*

De formulering van de sleutel valt op door de lengte en de nauwkeurige omschrijving. Vergeleken met de twee alternatieven kan hieruit worden afgeleid dat A het goede antwoord is. Overigens is B ook een goed alternatief maar minder precies geformuleerd dan A. C kan het gevolg zijn van een hypo, maar niet elke hypo leidt tot wegvalen.

Verder valt op dat de drie alternatieven de vraag vanuit verschillende invalshoeken beantwoorden: vanuit de patiënt, in het algemeen en vanuit 'men'. Het is beter om de alternatieven vanuit dezelfde invalshoek te beschrijven.

Wel

Wanneer is sprake van een hypoglykemie?

- A. Als bij de patiënt de waarde van de bloedglucose daalt onder 4,0 - 3,5 mmol/l.**
- B. Als bij de patiënt de waarde van bloedglucose stijgt boven de 10 mmol/l.*
- C. Als de patiënt onduidelijk praat, niet meer aanspreekbaar is en wegvalt.*

Voorbeeld

Niet

Wat zijn voorbeelden van tautologie?

Let op! Meerdere antwoorden kunnen goed zijn.

- A. nooit ofte nimmer**
- B. witte sneeuw*
- C. blij en verheugd**
- D. keurig netjes**
- E. groen gras*

De afleiders B en E lijken zoveel op elkaar, dat die twee wel bij elkaar moeten horen. Als de student weet dat dat voorbeelden zijn van pleonasme dan weet hij dat A, C en D de voorbeelden zijn van tautologie. Hier onderscheiden de afleiders zich dus onvoldoende van elkaar.

Wel

Wat zijn voorbeelden van tautologie?

Let op! Meerdere antwoorden kunnen goed zijn.

- A. nooit ofte nimmer**
- B. een ruiter te paard*
- C. blij en verheugd**
- D. keurig netjes**
- E. groen gras*

Of, en nog beter:

Kruis aan wat een voorbeeld is van tautologie en wat een voorbeeld is van pleonasme.

	<i>Tautologie</i>	<i>Pleonasme</i>
<i>Nooit ofte nimmer</i>	<i>X</i>	
<i>Een ruiter te paard</i>		<i>X</i>
<i>Blij en verheugd</i>	<i>X</i>	
<i>Keurig netjes</i>	<i>X</i>	
<i>Groen gras</i>		<i>X</i>

Constructievoorschrift	25
De alternatieven bevatten geen vage of absolute formuleringen	
Kwaliteitseis	Specificiteit

Belang

Vage of absolute formuleringen geven bijna altijd een indicatie voor het goede antwoord.

Toelichting

Vage formuleringen zijn formuleringen met woorden zoals vaak, soms, meestal, enig, nauwelijks en kan. Absolute formuleringen zijn formuleringen met woorden zoals alleen, alle, allemaal, altijd, geen, niet, nooit en uitsluitend.

Een alternatief met vage formuleringen is een indicatie voor de sleutel. Alternatieven met absolute formuleringen zijn bijna altijd afleiders.

Naar analogie van de vage en absolute formuleringen geven zogenaamde kan-stellingen ook een indicatie voor het goede antwoord. Bij medische toetsen komen kan-stellingen veel voor. In de meerderheid van de gevallen zijn ze juist, omdat het een menselijke eigenschap is om naar de goede oplossing te vragen en niet naar een verkeerde.

Valkuil

De valkuil is dat veel ontwikkelaars niet doorhebben dat ze met de gebruikte vage en absolute formuleringen de studenten juist hints gaven in de richting van het goede antwoord. Met dit constructievoorschrift wordt deze valkuil gedicht.

Voorbeelden

De eerste twee voorbeelden spreken voor zich. Het derde voorbeeld is een voorbeeld uit de medische wereld met de kan-stelling.

Voorbeeld

Niet

Met welke botten vormt het os sacrum een verbinding?

- A. Uitsluitend met het os coxae en het os coccygis.*
- B. Uitsluitend met de vijfde lumbale wervel en het os coccygis.*
- C. Uitsluitend met de vijfde lumbale wervel en het os coxae.*
- D. Met de vijfde lumbale wervel, het os coccygis en het os coxae.**

Uitsluitend is erg absoluut geformuleerd waardoor direct twijfel ontstaat over de juistheid van dat alternatief, maar bij drie keer ga je toch twijfelen. Door ook de botten in de alternatieven te gaan tellen blijkt dat in A, B en C de os coxae, de os coccygis en de vijfde lumbale wervel elk twee keer voorkomen én alle drie ook in D (zie Constructievoorschrift G-30: De sleutel is niet af te leiden uit woordherhalingen in de alternatieven). Hiermee wordt D wel een erg aantrekkelijk alternatief om te kiezen. Niet toevallig is D ook het goede alternatief.

Wel

Met welke botten vormt het os sacrum een verbinding?

Let op! Meerdere antwoorden kunnen goed zijn.

- A. met het os coxae**
- B. met het os coccygis**
- C. met de derde lumbale wervel*
- D. met de vierde lumbale wervel*
- E. met de vijfde lumbale wervel**

Voorbeeld

Niet

Het gebruik van een zendantenne met een vaste lengte op verschillende frequentiebanden is:

- A. Niet mogelijk zonder extra aanpassing omdat elke frequentie zijn eigen golflengte heeft en dit de antennelengte bepaalt.**
- B. Altijd mogelijk zonder extra aanpassingen.*
- C. Alleen mogelijk als de antenne een goede aarding heeft en op het hoogste punt van iets is geplaatst.*

Alternatief B is een onwaarschijnlijk alternatief, juist door het woord 'altijd'. Bij alternatief C trekt het woord 'alleen' direct de aandacht. Die aarding zal wel kloppen, alleen op het hoogste punt, is zeer twijfelachtig. Het hoogste punt zal leiden tot een beter ontvangst maar daardoor niet tot de ontvangst van verschillende frequentiebanden. Bij alternatief A is sprake van extra aanpassingen. Als alternatief B juist daarom af valt moet alternatief A ook om die reden wel het juiste alternatief zijn.

Overigens is de vraag geen echte vraag omdat de 'vraag' niet als een vraag is geformuleerd (Constructievoorschrift G-12: De Stam heeft een duidelijke opdracht of vraag). Het is beter om een concrete vraag te stellen.

Wel

Welke eigenschappen horen bij een antenne met een vaste lengte?

Let op! Meerdere antwoorden kunnen goed zijn.

- A. Kan niet elke frequentie ontvangen.**
- B. Kan elke frequentie ontvangen.*
- C. Moet aarding hebben.**
- D. Heeft geen aarding nodig.*
- E. Moet bij voorkeur op het hoogste punt in de buurt staan.**
- F. Kan overal in de vrije ruimte staan, maar de ontvangstkwaliteit kan per plaats verschillen.**

Voorbeeld**Niet**

Stelling:

Een hoge triglyceridenwaarde in het bloed kan worden veroorzaakt door een traag werkende schildklier.

Is deze stelling juist/onjuist?

- A. juist**
- B. onjuist*

Ook bij deze vraag blijkt de kan-stelling juist te zijn.

Wel

Een hoge triglyceridenwaarde in het bloed kan aan de werking van de schildklier liggen. Is de werking van de schildklier dan versneld of vertraagd?

- A. versneld*
- B. vertraagd**

Constructievoorschrift	26
De alternatieven zijn niet als 'Alle van bovenstaande' of 'Geen van bovenstaande' geformuleerd	
Kwaliteitseis	Relevantie, efficiëntie, objectiviteit

Belang

Vragen met als alternatief 'Alle van bovenstaande' en 'Geen van bovenstaande' zijn, toetstechnisch gezien, bijna altijd kwalitatief slechte vragen. Ze komen helaas vaak voor.

Toelichting

Omdat de stam bij voorkeur als een vraag wordt geformuleerd, kunnen 'Alle van bovenstaande' en 'Geen van bovenstaande' nooit het directe antwoord zijn op de vraag: *Wie is lid van het koninklijk huis? Alternatief D Alle van bovenstaande.*

Als de sleutel 'Alle van bovenstaande' is, hebben de alternatieven een relatie met elkaar (zie het voorbeeld hierna). Als de sleutel 'Geen van bovenstaande' is, wordt niet getoetst of de student het juiste antwoord wel weet! Het is beter om dan het goede antwoord als sleutel gebruiken.

Als 'Alle van bovenstaande' en 'Geen van bovenstaande' niet de sleutel zijn, kunnen deze afleiders beter worden weglaten. Met de meer-uitmeervraag wordt het probleem van 'Alle van bovenstaande' opgelost.

Valkuil

-

Voorbeelden

De voorbeelden spreken voor zich. De situatie bij het tweede voorbeeld dat in geen enkele vraag alternatief D 'Geen van bovenstaande' de sleutel bleek te zijn, is geen uitzondering. Dit geldt helaas voor veel toetsen.

Voorbeeld

Niet

In de darmen van elk mens is een bepaald type enterobacterie dat domineert. Op basis hiervan kan iedereen ingedeeld worden in een bepaalde darmfloragroep. Zo heb je de Bacteroides-, Prevotella- en Ruminococcussoort. Een van de opmerkelijke verschillen tussen deze groepen is de vitamineproductie.

Waar produceert de Bacteroidesgroep het meeste van?

- A. vitamine B2*
- B. vitamine B5*
- C. vitamine C*
- D. vitamine H*
- E. alle van bovenstaande**

'Alle van bovenstaande' is geen antwoord op de vraag, want er wordt naar vitamines gevraagd.

Een student die bijvoorbeeld weet dat twee van de alternatieven van A tot en met D goed zijn, weet dan al dat de sleutel alternatief E is. Hij hoeft dan niet meer te weten of de andere alternatieven juist of onjuist zijn.

Het omgekeerde geldt ook. Stel dat alternatief B goed zou zijn (wat niet zo is) dan zou iemand die weet dat alternatief D niet goed is, ook alternatief E kunnen wegstrepen.

Wel

In de darmen van elk mens is een bepaald type enterobacterie dat domineert. Op basis hiervan kan iedereen ingedeeld worden in een bepaalde darmfloragroep. Zo heb je de Bacteroides-, Prevotella- en Ruminococcussoort. Een van de opmerkelijke verschillen tussen deze groepen is de vitamineproductie.

Van welke vitamine of vitamines produceert de Bacteroidesgroep het meest?

Let op! Meerdere antwoorden kunnen goed zijn.

- A. van B2**
- B. van B5**
- C. van B8*
- D. van B12*
- E. van C**
- F. van H**
- G. van K*

Voorbeeld**Niet**

Bij het verwerken van audiovisuele informatie van media vinden bij kijkers en luisteraars vaak onbewust een aantal automatische selectieprocessen plaats.

*Welke van onderstaande fenomenen behoort **niet** tot automatische selectie van media-informatie?*

- A. dissonantiereductie**
- B. oriënterende reflex*
- C. post arousal blokkade*
- D. geen van bovenstaande antwoorden is juist*

Deze vraag komt uit een toets met tien een-uit-meervragen met vier alternatieven. Bij vier vragen was alternatief D 'Geen van bovenstaande antwoorden is juist'. Bij geen enkele vraag was dit het juiste alternatief. Het werd als opvulling gebruikt bij gebrek aan een goede derde afleider. Dan is met 'Geen van bovenstaande' het vierde alternatief snel gevonden en waarschijnlijk wordt gedacht dat dit geen kwaad kan, want het is niet het goede antwoord. Dat sommige vragen drie alternatieven hebben en de ander vragen vier is geen enkel probleem.

Opbouw kort-antwoordvraag

Studentgedeelte

Vraaggedeelte	■ Een beknopte vraag (meestal geen kleine opdracht). ■ Meestal geen situatieschets.
Informatiegedeelte	■ Geen of een beperkte antwoordinstructie. ■ Aantal te behalen punten met geen of beperkte globale beoordelingscriteria. ■ Eventueel een afbakening van de lengte van het antwoord.
Antwoordgedeelte	Het antwoord van de student zal bestaan uit enkele woorden, een korte losstaande zin, een korte berekening, een eenvoudige schets, een formule of iets dergelijks.

Beoordelaarsgedeelte

Antwoordmodel	■ Een eenvoudig modelantwoord. ■ Geen of enkele ook goed te rekenen antwoord-alternatieven. ■ Geen of enkele niet goed te rekenen antwoord-alternatieven.
Beoordelingsvoorschrift	Een beknopt beoordelingsvoorschrift met geen of slechts enkele criteria en een beperkt aantal te halen punten zoals 0 (fout), 1 (gedeeltelijk goed) of 2 (goed).

Tabel 11.4 Onderscheidende kenmerken kort-antwoordvraag

Voorbeeld kort- antwoord- vraag

Voor de student

Vraag <i>Hoe wordt een indexcijfer van economische gegevens berekend?</i> Aandachtspunten voor de beantwoording ■ <i>Goed antwoord: 1 punt</i> ■ <i>Alleen hele punten worden toegekend, dus 0 (fout) of 1 (goed)</i> Antwoord _____
--

Voor de beoordelaar

Antwoordmodel

Modelantwoord:

Nieuwe waarde / oude waarde x 100

Alternatieve goede antwoorden:

- *Nieuw / oud x 100%*
- *Nieuwe prijs / oude prijs x 100 of 100%*

Niet goed te rekenen alternatieve antwoorden:

De AEX-index bijvoorbeeld, wordt anders berekend, namelijk de gemiddelde beurswaarde van alle koersen op een bepaald moment gedeeld door een divisor. Om deze reden wordt in de vraag de link met economische gegevens gelegd.

Een antwoord dat een index op basis van een gemiddelde waarde van iets wordt verkregen dient dus 0 punten te krijgen.

Beoordelingsvoorschrift

Goed antwoord: 1 punt. Antwoorden gedeeltelijk goed rekenen is niet toegestaan, dus eventueel afronden naar 0 of naar 1 punt.

Opbouw lang-antwoordvraag

Studentgedeelte

Vraaggedeelte	■ Een vraag of een kleine opdracht. ■ Een eventuele korte situatieschets.
Informatiegedeelte	■ Indien nodig een korte antwoordinstructie. ■ Het aantal te behalen punten bij voorkeur gerelateerd aan globale beoordelingscriteria. ■ Een afbakening van de lengte van het antwoord.
Antwoordgedeelte	Het antwoord van de student zal bestaan uit enkele samenhangende zinnen, een uitgebreidere berekening of schets, een onderbouwing van een formule, een opsomming of iets dergelijks.

Beoordelaarsgedeelte

Antwoordmodel	■ Een breder uitgewerkt modelantwoord. ■ Geen of enkele ook goed te rekenen antwoord-alternatieven. ■ Geen of enkele niet goed te rekenen antwoord-alternatieven.
Beoordelingsvoorschrift	Een breder uitgewerkt beoordelingsvoorschrift, met een beoordelingsinstructie, bij voorkeur een opsomming van gedetailleerde beoordelingscriteria en eraan gekoppelde punten.

Tabel 11.5 Onderscheidende kenmerken lang-antwoordvraag

Voorbeeld lang- antwoord- vraag

Voor de student

Opdracht/vraag

Doordat Nederland gedeeltelijk onder de zeespiegel ligt, houden Rijks-waterstaat en de Waterschappen het niveau van de zeespiegel nauwlettend in de gaten. Een belangrijke aspecten hierbij is het springtij.

- Geef een korte beschrijving van springtij.*
- Verklaar beknopt waardoor springtij ontstaat.*
- Wanneer is springtij voor de Nederlandse kust zo gevaarlijk dat dijkbewaking wordt overwogen?*

Aandachtspunten voor de beantwoording

- Voor A, B en C elk maximaal 5 punten. Dus totaal maximaal 15 punten voor deze opdracht/vraag. Er worden alleen hele punten gegeven.
- Voor A, B en C mogen elk vijftig woorden of minder worden gebruikt. Worden meer woorden gebruikt dan wordt het te veel aan woorden niet meegenomen in de beoordeling, ook al bevatten die woorden delen van het goede antwoord.

Antwoord

A. _____

B. _____

C. _____

Voor de beoordelaar**Antwoordmodel**

Modelantwoord:

A. Beschrijving

De getijperiode (1) waarin het verschil tussen hoog water en laag water het grootst is (2). Hierbij is het hoogwater, hoger dan gemiddeld en het laagwater later dan gemiddeld (3). Het treedt eens in de 14 dagen op (4).

B. Ontstaan

Springtij ontstaat als de zon, de aarde en de maan in een rechte lijn staan (5). De invloed van hun zwaartekrachten versterken elkaar dan (6). Dit is het geval bij volle maan en nieuwe maan (7). Het springtij doet zich ongeveer twee dagen na deze situatie voor (8).

C. Dijkbewaking

Als sprake is van springvloed (9) in combinatie met een storm, zware storm, zeer zware storm of een orkaan (10) uit met name Noord-Westelijke richting (11).

Alternatieve goede antwoorden:

Kijk naar de beschrijvingen van de student en beoordeel of de kenmerken 1 tot en met 11 grotendeels of met andere woorden voorkomen. Interpreteer dit zelf.

Niet goed te rekenen alternatieve antwoorden:

-

Beeoordelingsvoorschrift

De punten 2, 3, 5, 6, 9, 10 en 11 mogen met twee punten worden beoordeeld worden, de overige punten met 1 punt. Het maximum per A, B, of C is 5 punten, ook al behaalt de student meer punten.

Bij meer dan 50 woorden mag tot 20% nog in de beoordeling worden meegenomen (studenten mogen een zin afmaken). Bij meer dan 20% mag die tekst niet mee-beoordeeld worden. Ook al staat er goede informatie in die punten oplevert.

Opbouw essay

Studentgedeelte

Vraaggedeelte	■ Een kleine opdracht. ■ Een (uitgebreide) situatieschets.
Informatiegedeelte	■ Een uitgebreidere antwoordinstructie zoals een specificatie van de te behandelen onderwerpen en/of suggesties voor de opbouw of taalkundige aspecten. ■ Een overzicht van de globale beoordelingscriteria en de daaraan gekoppelde puntenverdeling. ■ Een afbakening van de lengte van het essay.
Antwoordgedeelte	Het antwoord van de student zal bestaan uit een uiteenzetting, analyse, concept, visie, ontwerp, samenvatting van iets. Het essay bestaat uit een goed leesbare, qua spelling en grammatica correcte tekst die samenhangend en gestructureerd is. Welke eisen aan goed leesbaar, samenhangend en gestructureerd worden gesteld, dient bekend te zijn bij de studenten.

Beoordelaarsgedeelte

Antwoordmodel	Er is geen antwoordmodel vanwege het divergerende karakter van de opdracht. Wel dient het resultaat van het essay binnen de opdracht te passen. Hiervoor dienen eventueel richtlijnen of eisen genoemd te worden waar de uitwerking aan moet voldoen. Deze kunnen bij het beoordelingsvoorschrift worden opgenomen.
Beoordelingsvoorschrift	Een uitgebreid beoordelingsvoorschrift, met een beoordelingsinstructie en een gedetailleerde opsomming van beoordelingscriteria en eraan gekoppelde punten.

Tabel 11.6 Onderscheidende kenmerken essay

De afkortingen in de laatste kolom verwijzen naar de kwaliteitscriteria (zie tabel 11.10).

Constructievoorschriften voor open vragen geordend op basis van de opbouw van een open vraag inclusief antwoordmodel en beoordelingsvoorschrift

De vraag als geheel (studentgedeelte én beoordelaarsgedeelte)

O-1	sluit aan bij de inhoud en het gedrag van het leerdoel en het gedragsniveau uit de toetsmatrijs.	Re*
O-2	is zoveel mogelijk authentiek.	Re
O-3	is als vraagvorm (kort-, lang-antwoord of essay) het meest geschikt.	Ef
O-4	is inhoudelijk, qua formulering en taalkundig correct.	Ob/Re
O-5	is niet afhankelijk van een andere vraag.	Sp
O-6	bevat geen inhoud die antwoord geeft op een andere vraag. Let hierbij ook op informatie in situatieschetsen.	Ov/Sp
O-7	is qua lay-out zowel voor de student als de beoordelaar zodanig opgezet dat de student en de beoordelaar direct overzien hoe een en ander is opgebouwd, wat hun taak inhoudt en hoe ze die moeten uitvoeren.	

Studentgedeelte

De vraag of kleine opdracht inclusief eventuele situatieschets

O-8	is begrijpelijk, concreet en eenduidig. Bij een vraag is de vraag als een vraag met een vraagteken geformuleerd.	Ob
O-9	bevat niet een te omvangrijk probleem, moet in zijn geheel te maken zijn binnen de beschikbare tijd.	Ob
O-10	is kort, bondig en niet te ingewikkeld geformuleerd.	Ef/Mo/Sp
O-11	bevat geen vage aanduidingen zoals <i>enkele, een aantal, beschrijf in het kort</i> .	Ee
O-12	bevat geen overbodige informatie.	Ef/Mo/Re/Sp
O-13	vraagt niet naar de mening van de student.	Ob
O-14	bevat geen bias.	Ov/Sp
O-15	maakt zinvol en op een goede wijze gebruik van media (afbeeldingen, audio en/of video).	Re

O-16	bevat zo min mogelijk negatieve formuleringen zoals <i>niet</i> , <i>nooit</i> of <i>geen</i> .	Ob
O-17	bevat geen dubbele ontkenning.	Ob

De antwoordinstructie

O-18	is aanwezig, indien zinvol, en bevat dan informatie over het aantal woorden en de vorm.	Ob
O-19	is begrijpelijk, concreet, eenduidig en volledig.	Ob

De globale beoordelingscriteria met puntenverdeling

O-20	zijn aanwezig, indien zinvol. Een punttoekenning is altijd wel aanwezig.	Ob
O-21	zijn logisch en begrijpelijk.	Ob

De antwoordruimte

O-22	is aanwezig of wordt naar verwezen.	Ob
O-23	is qua omvang in overeenstemming met het aantal woorden dat wordt verwacht.	Ob

Beoordelaarsgedeelte

Het modelantwoord (bij kort- en lang-antwoordvragen)

O-24	is aanwezig.	Ob
O-25	is begrijpelijk, concreet en eenduidig geformuleerd.	Ob
O-26	is vastgesteld door de ontwikkelaar, betrokken deskundigen en eventuele anderen en wordt daarmee uniform gedragen qua inhoud, omvang en vorm.	Ob
O-27	is niet te algemeen of te globaal waardoor van een uniforme beoordeling weinig terecht komt.	Ob
O-28	laat geen ruimte over voor individuele interpretaties.	Ob

De alternatieve goede modelantwoorden

O-29	zijn aanwezig, indien zinvol.	Ob
O-30	voldoen ook aan de eisen die zijn genoemd bij het modelantwoord.	Ob

De niet goed te rekenen alternatieve antwoorden

O-31	zijn aanwezig, indien zinvol.	Ob
O-32	zijn qua inhoud en betekenis duidelijk en betrokken deskundigen en anderen zijn het allemaal eens dat die antwoorden niet goed gerekend kunnen worden.	Ob

De beoordelingsinstructie

O-33	is aanwezig.	Ob
O-34	is begrijpelijk, concreet, eenduidig en volledig.	Ob

De gedetailleerde beoordelingscriteria met puntenverdeling

O-35	zijn aanwezig.	Ob
O-36	zijn logisch en begrijpelijk voor elke beoordelaar en bieden een eenduidig en concreet handvat voor de beoordeling zodat een goede puntentoekenning per criterium mogelijk is.	Ob
O-37	zijn zodanig opgezet dat alle beoordelaars de beoordelingscriteria en de puntenverdeling op eenzelfde wijze toepassen.	Ob
O-38	bevatten ruimte voor het maken van opmerkingen.	Ob

Tabel 11.9 Constructievoorschriften voor open vragen geordend op basis van de opbouw van een open vraag

Re	Relevantie
Ob	Objectiviteit
Sp	Specificiteit
Ef	Efficiëntie

Mo	Moeilijkheid
Ov	Onderscheidend vermogen

Tabel 11.10 Afkortingen voor kwaliteitscriteria voor gesloten en open vragen en casussen

In paragraaf 12.1 wordt de casustoets uitgebreid besproken. Daar worden constructievoorschriften gegeven voor het ontwikkelen van casussen. Omdat bij open vragen korte casussen vaak een rol spelen worden deze constructievoorschriften hieronder genoemd. Voor de toelichting op de constructievoorschriften en voor meer informatie over casussen zie paragraaf 12.1.

Constructievoorschriften voor casussen

C-1	Casussen worden ontwikkeld op basis van leerdoelen en een toetsmatrijs.	Re*
C-2	Casussen worden ontwikkeld op basis van authentieke beroepssituaties of beroepsproducten.	Re
C-3	De casusbeschrijving is feitelijk, objectief, volledig, geplaatst in een context en gestructureerd.	Ef/Ob
C-4	De casus bevat geen overbodige informatie.	Ef/Mo/Re/Sp
C-5	De casus wordt beoordeeld door één of meer collega's en eventueel ook door deskundigen van buiten.	Re/Sp

Tabel 11.11 Constructievoorschriften voor casussen

11.6. Formuleren van de vraag of kleine opdracht bij een open vraag

In tegenstelling tot het hoofdstuk over gesloten vragen worden niet alle constructievoorschriften voor open vragen toegelicht vanwege de overlap die bestaat met de constructievoorschriften voor de gesloten vragen. Het advies is om paragraaf 10.8 *Constructievoorschriften voor gesloten vragen* eerst door te nemen. Daarnaast zijn veel nieuwe constructievoorschriften voor de open vraag logisch, en spreken ze voor zich.

Hierna worden enkele voorbeelden gegeven van de meest relevante constructievoorschriften voor open vragen.

O-8: De vraag of kleine opdracht inclusief eventuele situatieschets is begrijpelijk, concreet en eenduidig

Belang

Het goed formuleren van de vraag (of de kleine opdracht) is het belangrijkste constructievoorschrift bij een open vraag. Als de bedoeling van de vraag niet duidelijk is, ontstaan uiteenlopende antwoorden en wordt de beoordeling lastig.

Toelichting

Ontwikkel direct nadat de vraag is geschreven het antwoordmodel. Neem afstand en beoordeel of het antwoordmodel het enige antwoord is dat een student zou geven op de vraag, ervan uitgaande dat hij de stof beheerst.

Valkuil

Teveel denken vanuit het eigen referentiekader en onvoldoende afstand nemen van de vraag. Het is verrassend om achteraf te ontdekken, hoe studenten de vraag soms hebben gelezen. Zie ook het voorbeeld bij constructievoorschrift G-4: *De vraag als geheel is inhoudelijk, qua formulering en taalkundig correct.*

Voorbeelden

Hierna worden drie voorbeelden gegeven. Voorbeeld twee laat goed zien hoe verschillende studenten tot verschillende antwoorden komen.

Voorbeeld

Niet

Waarom controleert de douane op Schiphol de identiteit van personen aan de hand van het paspoort?

Binnen deze vraag zijn vijf aspecten te herleiden die elk tot een goed antwoord kunnen leiden als de vraag niet verder gespecificeerd wordt:

- 1. Nadruk op waarom**
Omdat daarmee personen kunnen worden ontdekt die gezocht worden.
- 2. Nadruk op douane**
Omdat de douane daarvoor de geautoriseerde organisatie is.
- 3. Nadruk op Schiphol**
Omdat Schiphol een buitengrens betreft voor bestemmingen naar niet-Schengenlanden.

12.3. De Script Concordance Test

De Script Concordance Test (SCT) is een toetsvorm die binnen het specifieke terrein van klinisch redeneren binnen het medisch onderwijs naam heeft verworven en steeds meer wordt gebruikt. Deze toets is een speciale vorm van een meerkeuzetoets.

Deze paragraaf begint met een beschrijving wat een SCT op hoofdlijnen is. Tevens worden enkele voorbeelden gegeven. De begrippen script en concordance worden toegelicht omdat die veel zeggen over de achtergrond van deze toetsvorm. Vervolgens worden de onderdelen waaruit de SCT bestaat uitgediept waarna enkele belangrijke aspecten van de SCT worden toegelicht. Deze paragraaf sluit af met de voor- en nadelen van de SCT.

In deze paragraaf over de SCT wordt geen overzicht van constructievoorschriften of criteria gegeven. Informatie over het ontwikkelen van een SCT is opgenomen in de alinea's die gaan over de onderdelen waaruit een SCT bestaat.

Wat is een Script Concordance Test?

De SCT wordt in het medische domein gebruikt om klinisch redeneren te toetsen. Medici krijgen te maken met patiënten met (vage) klachten en met klachten waarvoor niet zomaar een diagnose is te stellen zonder dat ze meer van de patiënt weten of nader onderzoek hebben uitgevoerd. Door klinisch redeneren worden door de medicus, medische kennis en inzicht, en ervaringen gekoppeld aan waarnemingen (observaties en interpretaties) bij de patiënt, om daarmee te beredeneren welke diagnose, welk onderzoek of welke behandeling het meest waarschijnlijk is ten behoeve van de (verdere) behandeling van een patiënt.

De SCT wordt als volgt gedefinieerd.

Definitie

Script Concordance Test

De Script Concordance Test (SCT) is een toets met een gesloten vraagvorm waarmee in cognitief complexe en niet eenduidige of onzekere situaties, de student op basis van een korte realistische praktijkschets, de waarschijnlijkheid van enkele hypothesen beoordeelt.

Een hypothese is een verwachting van een bepaalde situatie. Deze wordt steeds beoordeeld nadat aan de student een aanvullend gegeven is verstrekt. Het aantal hypothesen dat per praktijkschets aan een student wordt voorgelegd is twee tot vijf en veelal drie.

De waarschijnlijkheid van een hypothese wordt door de student beoordeeld door een keuze te maken op een vijfpunts Likertschaal die loopt van -2 via -1, 0 en +1 naar +2.

Gemeten wordt wat per hypothese de afwijking is van de door de student gekozen waarschijnlijkheid van de hypothese, op de door een groep experts gekozen waarschijnlijkheid van die hypothese. Een groep van experts heeft hiervoor antwoord gegeven op dezelfde vraag (hypothese plus aanvullend gegeven op basis van de praktijkschets). Hierbij heeft de beantwoording van de groep van experts op de Likertschaal juist **niet** tot consensus geleid. De experts verschillen dus onderling van mening.

De praktijkschets wordt bij de SCT een vignet genoemd. Het geheel van het vignet, de hypothese, het aanvullende gegeven en de antwoordmogelijkheid op basis van Likertschaal wordt een casus genoemd.

Het volgende voorbeeld licht de definitie toe.

Voorbeeld

Vignet						
<i>Je ziet een 60-jarige man op de spoedeisende hulp met een linkszijdige verlamming.</i>						
Als je denkt aan de diagnose ...	En je verneemt dat ...	Dan is deze diagnose ...				
<i>Hersenabces</i>	<i>De patiënt tien dagen geleden bij de tandarts een behandeling heeft gekregen</i>	-2	-1	0	+1	+2
<i>Stroke</i>	<i>De klachten twee uur geleden plots zijn opgekomen</i>	-2	-1	0	+1	+2
<i>Herstentumor</i>	<i>De patiënt gisteren betrokken was bij een dronkenmansgevecht</i>	-2	-1	0	+1	+2

-2 = Uitgesloten of bijna uitgesloten

-1 = Minder waarschijnlijk

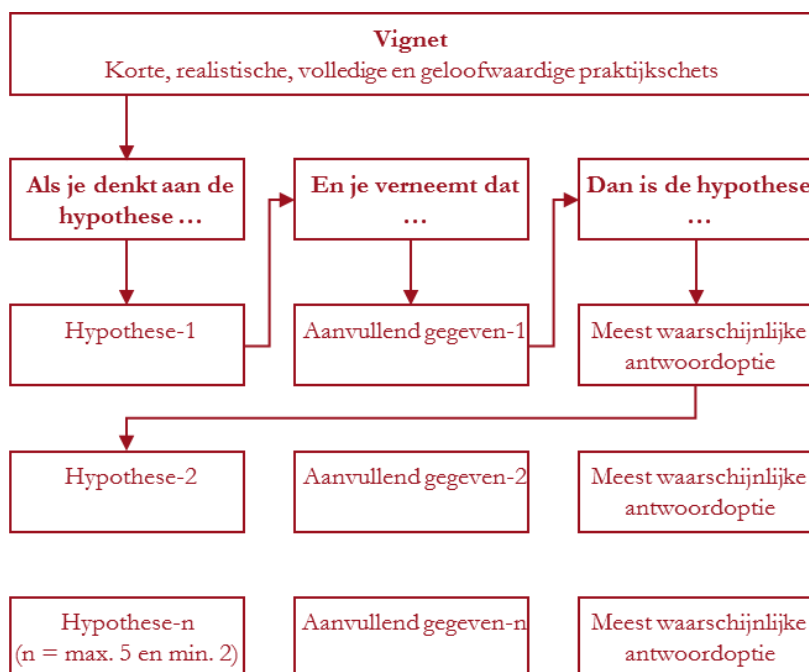
0 = Niet minder of niet meer waarschijnlijk

+1 = Meer waarschijnlijk

+2 = Zeker of bijna zeker juist

Bron: Katholiek Universiteit Leuven. (2012).

Uit het voorbeeld is op te maken dat elke casus van een SCT het volgende patroon kent en op de volgende wijze wordt doorlopen:



Figuur 12.8 Opbouw en het te doorlopen antwoordpatroon van een casus uit een SCT

SCT ook geschikt voor niet medische domeinen

Hoewel de SCT is ontwikkeld als toetsvorm voor klinisch redeneren binnen het medisch onderwijs, lijkt deze toetsvorm ook toepasbaar binnen andere opleidingen. Dit betreft opleidingen voor beroepen waarbinnen de beroepsbeoefenaren veel te maken krijgen met onzekerheden en niet eenduidige situaties. Daarbinnen zijn verschillende oplossingen mogelijk waarbij de ene oplossing weliswaar beter is dan een andere, maar een andere oplossing niet fout hoeft te zijn.

Gedacht kan worden aan ethische, politieke en maatschappelijke kwesties. Maar ook gedragswetenschappers, bedrijfs- en bestuurskundigen en adviseurs krijgen te maken met complexe en onzekere situaties waarin ze moeten kunnen redeneren op basis van opgedane kennis en inzicht, en ervaringen. Wellicht dat zelfs bij juristen de SCT van toepassing kan zijn als ze moeten leren pleiten? Voorwaarde is wel dat bepaalde scenario's van handelen (scripts) beschikbaar zijn.

De naamgeving van een SCT nader toegelicht

De begrippen script en concordance, onderdeel van de benaming van een SCT, vragen om een nadere toelichting.

Script

Een script is in algemene zin een hulpmiddel om te gebruiken als gespreksleidraad, zoals medewerkers van een klantcontactcenter dat toepassen.

Binnen het medische domein heeft een ervaren arts door zijn studie en ook door zijn ervaring met veel patiënten, in veel verschillende patiënt-situaties met ziektebeelden, onderzoeken, oplossingen en ook niet-oplossingen veel kennis en inzicht, en ervaringen verworven. Hierdoor is deze informatie zo impliciet en geïnterneerd geraakt (tacit knowledge), dat hij in een split-second al een eerste waarschijnlijke diagnose heeft bij een patiëntsituatie die hem wordt voorgelegd. Deze tacit knowledge geeft de ervaren arts, zo goed als onbewust het script om medisch te kunnen redeneren, waarmee hij tot zijn eerste diagnose komt, om vervolgens datzelfde script te gebruiken voor het stellen van vragen, te veranderen van diagnose, het aanvragen van onderzoeken, het voorschrijven van medicijnen, het aanbevelen van een bepaald gedrag of het aanreiken van een therapie.

Medisch studenten beschikken nog niet over scripts en moeten nog stap voor stap een klinisch redeneerproces doorlopen, waarbij ze per stap de waarschijnlijkheid van een mogelijke hypothese moeten afwegen op basis van steeds nieuwe aanvullende informatie. Door op een scriptgebaseerde wijze vragen te stellen wordt de student niet alleen op zijn medische kennis getoetst maar tevens getoetst of hij in staat is klinisch te redeneren.

Concordance

Concordance betekent overeenstemming. Binnen de SCT heeft dit te maken met de mate van overeenstemming tussen de door de studenten gebruikte scripts en de daaruit voortvloeiende waarschijnlijkheidskeuze van de student en de scripts van de experts uit de expertgroep en hun waarschijnlijkheidskeuzen. De mate van overeenstemming bepaalt de feitelijke score van de student.

Dat tussen experts niet altijd consensus bestaat over een vignet (Fournier, Demeester & Charlin, 2008) heeft onder meer te maken met het volgende:

- niet alle gegevens zijn beschikbaar;
- elke expert maakt gebruik van zijn eigen script;
- er bestaat in de literatuur nog geen consensus over de diagnose, onderzoek, behandeling of medicijn.

12.4. De mondelinge toets

Een toetsvorm waar voornamelijk gewerkt wordt met open vragen is de mondelinge toets. Deze paragraaf begint met de definitie van een mondelinge toets. Omdat een examinerator een grote rol speelt bij een mondelinge toets wordt ook de definitie van een examinerator gegeven. Daarna wordt aangegeven waarvoor een mondelinge toets wordt gebruikt. Vervolgens worden de vier fasen beschreven waaruit de afname van een mondelinge toets is opgebouwd en worden de belangrijkste aspecten van een mondelinge toets beschreven. Deze paragraaf wordt afgesloten met achttien criteria voor mondelinge toetsen.

Wat is een mondelinge toets?

De mondelinge toets wordt als volgt gedefinieerd.

Definitie

Mondelinge toets

Een mondelinge toets is een toets die in een één op één relatie tussen een student en een examinerator (en bij voorkeur twee examineratoren), aan de hand van een afwisseling van vragen van de examinerator en antwoorden van de student, mondeling wordt afgenomen.

De examineratoren stellen het resultaat van de toets vast.

Mondelinge toetsen zijn over het algemeen veel minder gestandaardiseerd dan schriftelijke toetsen. Een mondelinge toets is nog minder gestandaardiseerd als de mate van interactie tijdens een mondeling toeneemt en de mate van structuur afneemt (Van Loopik, 2012).

De examinerator is vaak de docent van de student. Er wordt in de definitie gesproken over examinerator en niet over beoordelaar, omdat de docent niet alleen beoordeelt, maar ook vragen heeft ontwikkeld, vragen stelt en doorvraagt.

Definitie

Examinator

Een persoon die een toets afneemt, die deskundig is ten aanzien van de inhoud, het niveau en de vorm van de toets en deskundig is ten aanzien van de wijze van afname van de toets.

Verschil
mondelinge
toets met een
CGI

Een mondelinge toets is een andere toetsvorm dan een criterium gericht interview (CGI). Binnen een CGI worden, vergelijkbaar bij een mondeling, ook vragen gesteld en beantwoordt de student de vragen. Een CGI is echter gekoppeld aan een andere toetsvorm zoals een portfolio- of een gedragsopdracht. Binnen het CGI wordt de student doorgevraagd over

het beroepsproduct dat de product- of vaardigheidsoopdracht heeft opgeleverd ter verantwoording van de uitvoering en de achterliggende denken en beslisprocessen (HVA, Score-website. n.d.). De mondelinge toets wordt ingezet om een set van leerdoelen te toetsen.

Een mondelinge toets heeft bijna altijd een summatieve functie en wordt haast niet formatief gebruikt, vanwege het capaciteitsbeslag dat een mondelinge toets legt op de examinatoren.

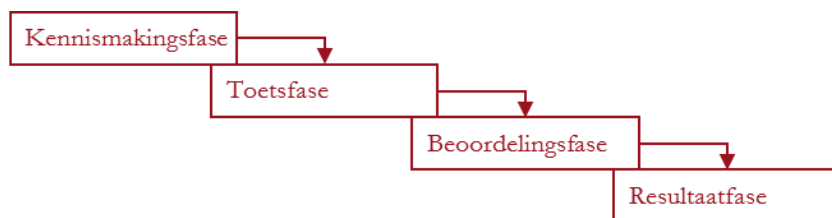
Waarvoor wordt een mondelinge toets gebruikt?

Een mondelinge toets wordt gebruikt in de volgende situaties:

- Bij talen waar het gaat om het meten van spreek- en taalvaardigheden.
- Bij toetsen waaraan weinig studenten (vijf tot zeventien) (Van Berkel, 2014; Ter Burg & Kempers-Warmerdam, 2013) deelnemen vanwege de efficiënte van een mondeling ten opzichte van het ontwikkelen, afnemen en beoordelen van een schriftelijke toets.
- Bij staatsexamens. Een staatsexamen is bedoeld voor studenten die een vmbo, havo of vwo diploma willen halen zonder deel te (kunnen) nemen aan het bekostigde (reguliere) onderwijs of indien ze gezakt zijn in het reguliere onderwijs. Naast het centraal schriftelijke examen krijgen de studenten een mondeling examen voor alle vakken. Dit mondelinge examen komt in de plaats van het zogenaamde schoolexamen dat een school afneemt en voor de helft het eindcijfer bepaalt.
- Voor studenten met een beperking zoals dyslectie, slechtziendheid, abnormale faal- of toetsangst, of taalproblemen. Dit betreft uitzonderingen waarvoor speciale procedures gelden om in aanmerking te komen voor een mondeling.

Opbouw van een mondelinge toets

Een mondelinge toets kent qua opzet van de afname over het algemeen steeds dezelfde opbouw met de volgende vier fasen:



Figuur 12.9 De vier fasen van de opbouw van een mondelinge toets

Kennismakingsfase

Het doel van de kennismakingsfase is de student op zijn gemak stellen en hem een overzicht geven van wat hij kan verwachten. Studenten zijn minder bekend met de mondelinge toets en ze weten niet altijd wie ze tegenover zich krijgen qua persoon (nors, aardig, doorvragend en dergelijke). Dit maakt de student onzeker. De kennismakingsfase bestaat bij voorkeur uit een kort voorstelrondje, een uitleg over hoe een en ander in zijn werk gaat, welke onderwerpen aan de orde zullen komen en of de student nog vragen heeft. Hiermee krijgen de deelnemers de kans om aan elkaar te wennen voordat wordt begonnen met de tweede fase, de toetsfase.

Toetsfase

Het doel van de toetsfase is vaststellen of de student laat zien dat hij de leerdoelen qua inhoud en gedragsniveau voldoende beheerst. De examinatoren stellen hiervoor vragen die de leerdoelen afdekken en vragen door (bredere en dieper) naar gelang de antwoorden die de student geeft. De belangrijke vragen zijn door de examinatoren van te voren ontwikkeld en hebben ze bij de hand. De examinatoren bewaken dat ze binnen de beschikbare toetstijd alle belangrijke vragen hebben gesteld die ze hebben voorbereid. Per vraag hebben de examinatoren ook een antwoordmodel bij de hand. Indien per vraag al het maximum aantal te behalen punten bekend is, kunnen examinatoren dit gebruiken om hun score per vraag te bepalen en te noteren.

Begin de toetsfase met enkele eenvoudige vragen. Dit helpt de student door het eerste moeilijke deel van de toetsfase te komen. De examinatoren noteren in steekwoorden het antwoord van de student bij de vraag. Voorkomen moet worden dat de examiner het antwoord in de mond van de student legt. Omdat (een van) de examinatoren vaak de docent is van de student, wil de docent zijn student wel eens een beetje helpen. Vaak gebeurt dit zelfs onbewust en door non-verbaal gedrag (knikken, knipogen, fronsend kijken). Houd tijdens de toets de rol van examiner en docent goed uit elkaar.

De toetsfase eindigt na de laatste vraag en het antwoord daarop van de student met een korte samenvatting door de examiner(en). Aan de student kan worden gevraagd hoe hij zelf vond dat de toets verliep. Deze informatie kan worden gebruikt voor de wijze waarop het resultaat van de toets en de feedback worden teruggekoppeld.

REGISTER FIGUREN EN TABELLEN

Figuren

0.1	Fragment uit de Vraagbaak 2CV/Dyane	7
1.1	Kwaliteitspiramide voor toetsen en beoordelen (Joosten-ten Brinke, 2011; Sluijsmans; Peeters; Jakobs & Weijzen, 2012)	15
1.2	Toetscyclus per toets	17
1.3	Producten per fase van de toetscyclus	18
1.4	De toetscyclus als bouwsteen binnen de kwaliteitspiramide voor toetsen en beoordelen	19
2.1	Piramide van Miller (1990) in relatie tot kennis en inzicht, vaardigheid, houding, gedrag en vraag-, opdracht- en toetsvormen	24
2.2	Een toets is breder dan een examen	27
4.1	Toetsen en leren, twee zijden van dezelfde medaille	36
4.2	Toetsprogramma en curriculum, innig verbonden	36
4.3	Toetsprogramma en toetsen in samenhang met het curriculum en leermiddelen	38
5.1	Toetsbeleid als contextelement voor toetstaak- en toetsontwikkelaars	50
5.2	Ontwikkeling in de relatie tussen kennis en inzicht, vaardigheden en houding gedurende het vorderen van een opleiding	56
5.3	Geïntegreerde ontwikkeling van kennis en inzicht, vaardigheden en houdingen van een geringe mate van complexiteit en zelfstandigheid naar een grote mate van complexiteit en zelfstandigheid	56
6.1	Hiërarchie begrippen	60
7.1	Schematische voorstelling taxonomie van Romiszowski	79
7.2	De twee varianten van de cognitieve taxonomie van Bloom	82
7.3	Verdeling lagere en hogere orde denkvaardigheden	83
7.4	De drie domeintaxonomieën	83
7.5	Beschrijvingen en werkwoorden behorende bij de gereviseerde cognitieve taxonomie van Bloom	84
7.6	Beschrijvingen en werkwoorden behorende bij de affectieve taxonomie van Krathwohl, Bloom en Masia	85
7.7	Beschrijvingen en werkwoorden behorende bij de psychomotorische taxonomie van Dave	86
7.8	Competentieniveaus binnen het ZelCommodel	91
8.1	Toetstermen vormen de basis voor vragen, opdrachten, toetsvormen en leermiddelen	96
8.2	Vierveldenmodel met geadviseerde toetsvormen	106
10.1	Opbouw en begrippen die horen bij een gesloten vraag waarbij de student een keuze moet maken uit de antwoordmogelijkheden	118

10.2	Begrippen die horen bij een gesloten vraag waarbij het juiste antwoord via enkele muishandelingen tot stand wordt gebracht	119
10.3	Begrippen die horen bij een gesloten vraag waarbij de student iets moet invullen	120
12.1	Opbouw van een lange casustoets (vorm 1)	264
12.2	Opbouw van een casustoets met korte casussen (vorm 2a)	265
12.3	Opbouw van een casustoets met korte casussen (vorm 2b)	265
12.4	Drie-eenheid tussen leerdoelen/toetsmatrijs, casus(sen) en vragen/toets	273
12.5	Voorbeeld cesuur interuniversitaire voortgangstoets geneeskunde met twee cesuurgrenzen voor 6 jaargroepen (bron: www.ivtg.nl)	281
12.6	Voorbeeld momentane feedback (bron: www.ivtg.nl)	282
12.7	Voorbeeld longitudinale en predictieve feedback (witte gedeelte, inclusief betrouwbaarheidsinterval) (bron: www.ivtg.nl)	283
12.8	Opbouw en het te doorlopen antwoordpatroon van een casus uit een SCT	293
12.9	De vier fasen van de opbouw van een mondelinge toets	305



Tabellen

3.1	Ordering beoordelingsbekwaamheid en bijpassende hoofdtoetsvorm	30
3.2	Overzicht hoofdtoetsvormen en de belangrijkste afgeleide toetsvormen	31
3.3	Overzicht toetsvormen website Score en alternatieve benamingen voor de toetsvormen (https://score.hva.nl/docent/15_toetsvormen)	32
3.4	Werkmodel voor toetsvormen (Jaspers & Heijmen-Versteegen, 2004). Kwaliteit van toetsing in het Hoger Onderwijs naar een versterkte rol van examencommissies, 2011)	33
5.1	Voorbeeld vereenvoudigd en verkort toetsprogramma	51
6.1	Toelichting op de SMART-methode	67
7.1	Toelichting schema taxonomie van Romiszowski, bewerkt en overgenomen uit Toelichting taxonomiecodes, Taxonomie van Romiszowski - www.bmwt.nl/files_content/Taxonomiecodes%20Romiszowski.pdf	80
7.2	Veelgebruikte voorbeelden van werkwoorden passend bij de taxonomie van Romiszowski	81
7.3	Niveau-indicaties volgens het HvA-model	90
7.4	Betekenis van laag, middel en hoog ten aanzien van complexiteit en zelfstandigheid bij het ZelCommodel	92
8.1	Basisvorm van een toetsmatrijs voor vragen op basis van de cognitieve taxonomie van Bloom	97
8.2	Eenvoudige toetsmatrijs met het aantal vragen waaruit de leerdoelen, toetstermen en de toets bestaan en een overzicht van de verdeling van de vragen in percentages (percentages zijn afgerond om ze kloppend te maken)	98
8.3	Toetsmatrijs met marges per toetsterm met het aantal vragen waaruit de toets bestaat	99
8.4	Voorbeeld toetsmatrijs waarbij alleen de leerdoelen worden gebruikt en niet de toetstermen	100
8.5	Voorbeeld toetsmatrijs waarbij leerdoelen en toetstermen worden gebruikt	100
8.6	Voorbeeld opdrachtenontwerpschema voor een opdracht	102
9.1	Kwaliteitscriteria voor vragen (gesloten en open) en opdrachten	115
10.1	Overzicht belangrijkste gesloten vraagvormen	118
10.2	Ontwikkelaanpak voor het ontwikkelen van gesloten vragen	144
10.3	Taalkundige aanbevelingen	146
10.4	Constructievoorschriften voor gesloten vragen geordend op basis van de opbouw van een gesloten vraag	148
10.5	Afkortingen van kwaliteitscriteria voor gesloten en open vragen en casussen	148
11.1	Open vraagvormen en beoordelingskenmerken	223
11.2	Beantwoordingstijd open vragen (Van Berkel, 1999)	226
11.3	Opbouw van een open vraag	228

11.4	Onderscheidende kenmerken kort-antwoordvraag	234
11.5	Onderscheidende kenmerken lang-antwoordvraag	236
11.6	Onderscheidende kenmerken essay	239
11.7	Ontwikkelaanpak voor het ontwikkelen van gesloten vragen	245
11.8	Taalkundige aanbevelingen	246
11.9	Constructievoorschriften voor open vragen geordend op basis van de opbouw van een open vraag	249
11.10	Afkortingen van kwaliteitscriteria voor gesloten en open vragen en casussen	250
11.11	Constructievoorschriften voor casussen	250
12.1	Constructievoorschriften voor casussen	272
12.2	Afkortingen voor kwaliteitscriteria voor gesloten en open vragen en casussen	272
12.3	Criteria voor mondelinge toetsen (Van Loopik, 2012)	317
12.4	Afkortingen voor kwaliteitscriteria voor gesloten en open vragen en casussen	317

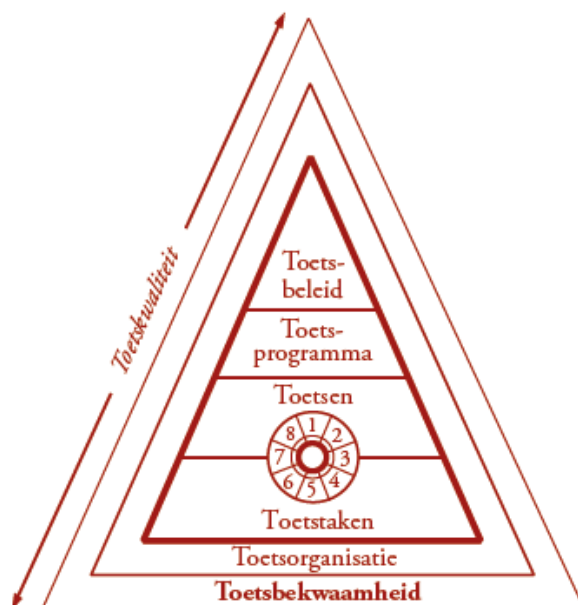


Toetsen volgens de toetscyclus

De kwaliteit van toetsen staat in het hoger onderwijs sterk in de belangstelling. Een van de aanbevelingen in het rapport *Vreemde ogen dwingen van de Commissie Bruijn* is "Zorg voor scholing van docenten die bij toetsing betrokken zijn." "Ter externe validering dient de certificering van examinatoren in het h(b)o vastgelegd te worden in een Basiskwalificatie examinering (BKE) en in een Seniorkwalificatie examinering (SKE)." De uitwerking hiervan is uitgevoerd door de Expertgroep BKE/SKE. Het resultaat van de Expertgroep is een overzicht van gewenste leeruitkomsten en indicatoren voor BKE en SKE. Deze leeruitkomsten en indicatoren dienen als basis voor het opleiden en certificeren van docenten in het h(b)o.

Dit zeer toegankelijke en makkelijk leesbare boek, het eerste van twee delen, biedt docenten vooral praktische handvatten om toetsbekwaam te worden en te blijven. De Kwaliteitspiramide voor toetsen en beoordelen met een toetscyclus met acht fasen biedt hiervoor de basis.

Dit boek bevat maar liefst 36 definities, 36 figuren, 37 tabellen, 87 constructievoorschriften en 119 voorbeelden! Naast de gesloten en open vraagvormen die uitgebreid worden behandeld, worden toetsvormen besproken die zijn gebaseerd op gesloten en/of open vragen: de casustoets, de voortgangstoets, de Script Concordance Test en de mondelinge toets, inclusief constructievoorschriften.



De missie van Bureau voor Toetsen & Beoordelen is het verhogen van de toetsbekwaamheid van medewerkers, het verbeteren van toetsproducten en het verbeteren van de toetsorganisatie. Dit boek geeft daar mede inhoud aan.

BTB BUREAU VOOR
Toetsen & Beoordelen

www.toetsen-beoordelen.nl



ISBN: 978-90-823697-0-0