

Tips bij Toetsen

Colofon

Deze handreiking is geschreven door:

- Maarten Bergwerff, Faculteit der Geesteswetenschappen, Universiteit Leiden
- Marlous Dekker, Faculteit der Geesteswetenschappen, Universiteit Leiden
- Floris van Blankenstein, ICLON, Universiteit Leiden
- Daan Romein, ICLON, Universiteit Leiden
- Marie Klaren, KlarenConsult, Leiderdorp

Ben je docent bij de Faculteit der Geesteswetenschappen en heb je vragen over het ontwikkelen, beoordelen of afnemen van toetsen? Neem dan contact op met de [onderwijsadviseur](#) van het team Onderwijsadvies en Kwaliteitszorg van de faculteit: Marlous Dekker (m.l.dekker@hum.leidenuniv.nl, 071-527 2153).

Het ICLON biedt [scholing en advies](#) over toetsen en beoordelen. Mocht je meer willen weten hierover, neem dan contact op met Floris van Blankenstein: f.m.van.blankenstein@iclon.leidenuniv.nl

Inhoud

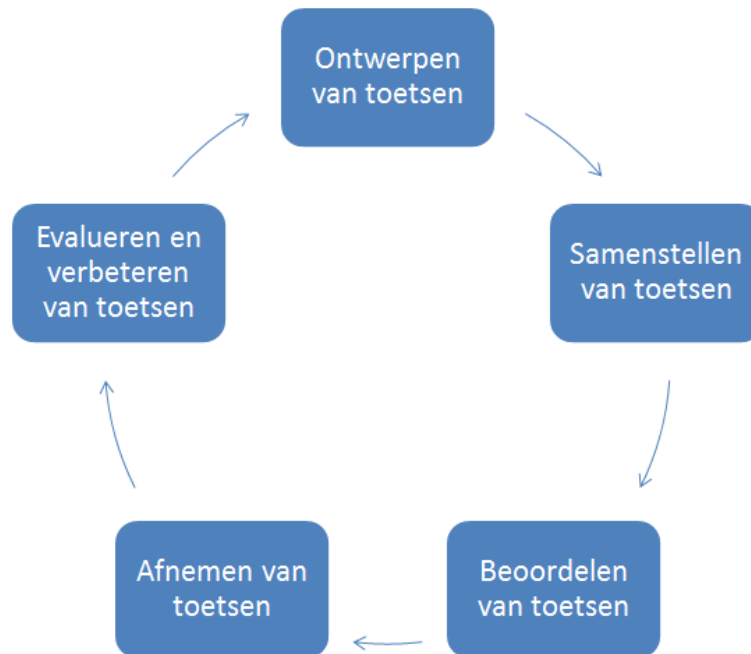
Inleiding.....	4
1. Ontwerpen van toetsen	5
2. Samenstellen van toetsen.....	7
2.1 Gesloten toetsvragen.....	7
2.2 Open toetsvragen	9
2.3 Schriftelijke werkstukken: essay, verslag en artikel.....	10
2.4 Mondelinge toetsen: presentaties.....	11
2.4 Toetsen van groepswerk: schriftelijke werkstukken of mondelinge presentaties	13
3. Samenstellen van beoordelingsmodellen.....	16
3.1 Beoordelingsmodellen voor open toetsvormen.....	16
3.1.1 Beoordeling van schriftelijke werkstukken (zoals essays en papers)	20
3.1.2 Beoordeling van eindwerkstukken (scripties).....	21
3.1.3 Beoordeling van mondelinge presentaties	22
3.1.4 Beoordeling van groepsopdrachten	23
3.2 Cesuurbepaling	24
4. Afnemen van toetsen.....	25
5. Evalueren en verbeteren van toetsen	27
6. Praktijkvoorbeelden Faculteit Geesteswetenschappen	28
Voorbeeld van leerdoelen	28
Voorbeelden van toetsvragen	28
- Voorbeelden van open toetsvragen	28
- Voorbeeld van open toetsvraag met modelantwoord	28
Voorbeeld van opdracht voor schriftelijk werkstuk	29
- Schrijfopdracht.....	29
Voorbeelden van beoordelingsformulieren	30
- Voorbeelden van beoordelingsformulieren voor schrijfopdrachten.....	30
- Voorbeeld van beoordelingsformulier inzet BA-werkcollege.....	33
- Voorbeeld van beoordelingsformulier BA-scriptie	34
Geraadpleegde bronnen	37
Bijlage I Opstellen van leerdoelen en passende toetsvormen	38
Bijlage II Maken van een toetsmatrijs.....	41
Bijlage III Checklist voor het samenstellen, beoordelen en redigeren van open vragen	42

Inleiding

'Tips bij toetsen' is geschreven voor en door de faculteit Geesteswetenschappen (in samenwerking met het ICLON).

'Tips bij toetsen' is bedoeld om docenten te helpen bij het opstellen en beoordelen van toetsen. Uit gesprekken met docenten is gebleken dat veel docenten het lastig vinden om toetsen te maken die de behandelde vakinhoud dekken. Ook geven docenten aan dat zij het moeilijk vinden om toetsen op een goede manier te beoordelen. In deze handreiking staan concrete tips voor docenten met verwijzingen naar voorbeelden en extra uitleg.

'Tips bij toetsen' behandelt vijf stappen van de toetscyclus die hieronder is weergegeven, te weten: Ontwerpen van toetsen (hoofdstuk 1), Samenstellen van toetsen (hoofdstuk 2), Beoordelen van toetsen (hoofdstuk 3), Afnemen van toetsen (hoofdstuk 4) en Evalueren en verbeteren van toetsen (hoofdstuk 5). Tot slot zijn in hoofdstuk 6 voorbeelden opgenomen van toetsvormen, toetsvragen en beoordelingsinstrumenten van de faculteit Geesteswetenschappen.



Figuur 1. Toetscyclus

1. Ontwerpen van toetsen

Wat is het ontwerp voor een toets?

Toetsen kun je gebruiken om inzicht te verkrijgen over het niveau van studenten. Dit kan zowel formatief, bijvoorbeeld om studenten te helpen hun leerproces te structureren, als summatief, bijvoorbeeld om te bepalen of een student de vakinhoud beheerst. Welke functie een toets ook heeft, de leerdoelen staan centraal bij het ontwerpen van toetsen. De basis van het toetsontwerp is een overzicht van leerdoelen en de daarbij behorende toetsonderwerpen. Vervolgens bepaal je de toetsvorm. Ook houd je rekening met het toetsmoment en de organisatie van de toets.

Hoe maak je het ontwerp?

Bij het ontwerpen van een toets staan drie vragen centraal, namelijk:

1. Waarom wordt er getoetst?
2. Wat wordt er getoetst?
3. Hoe wordt er getoetst?

Ad 1. Toetsfunctie

Voor het ontwerp van de toets is het belangrijk te weten waarom je toetst. Als je een toets ontwerpt omdat je tijdens het onderwijs feedback wilt verzamelen over de voortgang van het leerproces bij studenten (en dus op een z.g. formatieve wijze toetst), hanteer je mogelijk een andere toets dan wanneer je aan het eind van het onderwijsonderdeel het prestatieniveau van de student beoordeelt (en op een z.g. summatieve wijze toetst).

Ad 2. Leerdoelen

Leerdoelen, inhoud en toetsing zijn onlosmakelijk met elkaar verbonden. Helder geformuleerde leerdoelen zijn essentieel als je een toets ontwerpt. Wanneer je de toets gaat ontwerpen, grijp je terug op die leerdoelen; specifieke en meetbare leerdoelen geven veel richting aan het kiezen van een passende toetsvorm en toetsvragen¹. Dit werkt het beste met leerdoelen met actieve werkwoorden die specifiek en meetbaar zijn, bijvoorbeeld: beschrijven, uitleggen, verklaren, vergelijken etc. Woorden als kennis, inzicht en begrip zijn niet specifiek en meetbaar. Een leerdoel is bijvoorbeeld: “de student beschrijft drie oorzaken voor het ontstaan van de Tweede Wereldoorlog”. Gedurende het vak vergaren studenten kennis en inzicht om ‘drie oorzaken voor de Tweede Wereldoorlog’ te beschrijven en dit wordt aan het eind van het vak getoetst.

Ad 3. Toetsvorm

Welke toetsvorm je het beste kunt kiezen, hangt af van een aantal factoren. De belangrijkste zijn²:

- Leerdoelen: met de toetsvorm moet je daadwerkelijk kunnen meten of de leerdoelen zijn behaald;
- Benodigde en beschikbare constructietijd: het construeren van meerkeuzevragen kost meer tijd dan het opstellen van open vragen of opdrachten;
- Benodigde en beschikbare beoordelingstijd: meerkeuzetoetsen zijn sneller te beoordelen dan essays;

¹ Zie Bijlage I

² Bron: Kallenberg et al. (2014).

- Benodigde en beschikbare afnametijd: studenten moeten de toets kunnen maken binnen de gestelde afnametijd.

In principe kun je verschillende cognitieve niveaus (zoals onthouden, begrijpen, toepassen, analyseren) toetsen met verschillende toetsvormen. De ene toetsvorm is echter meer geschikt (en makkelijker te gebruiken) voor een bepaald cognitief niveau dan de andere. Wil je bijvoorbeeld dat de student twee visies op onderwerp X met elkaar kan vergelijken? Dan ligt het meer voor de hand om open vragen te stellen, of bijvoorbeeld een essay-opdracht te geven. Moet de student aspecten X, Y en Z kunnen benoemen? Kies dan bijvoorbeeld voor gesloten vragen. In bijlage I wordt een overzicht gegeven van passende toetsvormen bij verschillende cognitieve niveaus.

Het kan nuttig zijn om voor je begint aan het samenstellen van de toets een toetsmatrijs (ook wel specificatietabel) te maken. Dit kan meer of minder uitgebreid zijn. In de simpelste variant zet je de leerdoelen op een rij die je met de toets wilt beoordelen (en die ook tijdens het vak aan de orde zijn gekomen) met daarachter hoe zwaar/ belangrijk dit onderwerp is, door punten toe te kunnen aan de onderwerpen. Per onderwerp formuleer je minimaal één vraag. In een uitgebreidere variant geef je ook nog aan op welk cognitief niveau de vragen geformuleerd moeten worden; handig bij het evenwichtig verdelen van toetsvragen (bijvoorbeeld als je wilt voorkomen dat je alleen 'kennisreproductie-vragen' stelt). Ook kun je in een toetsmatrijs opnemen welke toetsvorm(en) je gebruikt; dat kan van pas komen als je meerdere toetsvormen los van elkaar gebruikt of juist combineert. Voor meer informatie over toetsmatrijzen: zie bijlage II.

Waarom is het van belang een toetsontwerp te maken?

Het ontwerp voor een toets geeft je inzicht in de samenhang tussen leerdoelen en toetsvragen en toetsvormen. Het helpt je om alle relevante leerdoelen in een toets op te nemen. Bovendien geven de leerdoelen richting aan de manier van lesgeven. Denk daarom al bij het ontwikkelen van de opzet van je vak na over het toetsontwerp. Studenten moeten tijdens de lessen tenslotte oefenen met het kunnen halen van de leerdoelen en feedback krijgen op hun voortgang. Daarom is het belangrijk om al bij de opzet van een vak te werken aan de ontwikkeling van een toetsontwerp.

Meer weten...

- Over het opstellen van specifieke en meetbare leerdoelen en bijpassende toetsvormen: Bijlage I
- Over het maken van een toetsmatrijs (incl. voorbeelden): Bijlage II
- Over [soorten leerdoelen](#): Knevel, R. (2013). Taxonomieën zijn hot en handig. *Bureau ICE*.

2. Samenstellen van toetsen

Bij het samenstellen van een toets kun je verschillende toetsvormen onderscheiden. In dit hoofdstuk komen de volgende toetsvormen aan de orde:

- gesloten toetsvragen
- open toetsvragen
- schriftelijke werkstukken
- mondelinge presentaties
- groepsopdrachten

2.1 Gesloten toetsvragen

Wat?

Gesloten vragen zijn toetsvragen waarbij de student uit twee of meer opties moet kiezen. Een daarvan is het juiste of beste antwoord. Met gesloten vragen kunnen studenten aantonen dat ze iets goed hebben onthouden, begrepen, of kunnen toepassen in een nieuwe situatie. Er zijn verschillende soorten gesloten toetsvragen: meerkeuzevragen, juist/onjuist-vragen (meestal aan de hand van een stelling of propositie), ja/nee-vragen of casuïstiekvragen met een aantal antwoordopties. Je kunt gesloten vragen gebruiken in een eindtoets (summatief), maar ook tijdens het onderwijs, bijvoorbeeld door quizvragen te stellen tijdens een werkgroep of college (diagnostisch of formatief).

Met gesloten vragen kun je naast reproductiekennis ook hogere cognitieve niveaus toetsen, zoals inzicht en toepassing. Dit vraagt echter wel veel aandacht voor de constructie van zowel de vraag als de antwoordopties. Het schrijven van 'toepassingsvragen' vraagt naar verhouding de meeste aandacht; meestal zijn dit vragen die worden ingeleid met een situatiebeschrijving (een casus), waarna de student de juiste antwoordoptie moet kunnen beredeneren.

Hoe?

Een gesloten toetsvraag bestaat uit twee delen: de *stam* en de *alternatieven*. De *stam* bevat de context van de vraag en de vraag zelf en de *alternatieven* zijn de antwoordopties. De niet-correcte alternatieven heten *afleiders* en het correcte alternatief heet de *sleutel*³.

Tips voor het formuleren van gesloten toetsvragen

Het schrijven van de stam (=context en de vraag zelf)

1. Geef duidelijke aanwijzingen in de stam. De student zou de vraag ook zonder de antwoordopties moeten kunnen beantwoorden.
2. Beschrijf de context en het centrale idee van de vraag.
3. Houd de stam kort en bondig.
4. Gebruik positieve bewoordingen, vermijd zoveel mogelijk woorden als NIET, GEEN of BEHALVE. Als het noodzakelijk is om negatieve woorden te gebruiken, zorg dan dat deze woorden opvallen door bijvoorbeeld hoofdletters te gebruiken.

³ Bron: Van Berkel, Bax & Joosten-ten Brinke (2014).

Het schrijven van de alternatieven (=de antwoordopties)

1. Zorg dat één antwoord echt juist is.
2. Varieer met de plek van het juiste antwoord of plaats antwoorden altijd in dezelfde logische volgorde (bijvoorbeeld alfabetisch).
3. Houd antwoordopties onafhankelijk; er moet geen overlap tussen antwoordopties zijn.
4. Zorg dat antwoordopties inhoudelijk en grammaticaal vergelijkbaar zijn.
5. Zorg dat de lengte van antwoordopties ongeveer hetzelfde is.
6. Wees terughoudend in het gebruik van bewoordingen als 'geen enkele van bovenstaande...' of 'alle bovenstaande...'
7. Formuleer antwoordopties positief; vermijd negatieve formuleringen zoals NIET.
8. Geef geen aanwijzingen richting het juiste antwoord, zoals:
 - a. Specifieke aanduidingen (altijd, nooit, helemaal, absoluut)
 - b. Associaties met tekst in de stam; woorden die identiek zijn aan woorden in de stam of die lijken op woorden in de stam
 - c. Grammaticale inconsistenties tussen antwoordopties
 - d. Meerdere samenhangende antwoordopties
 - e. Opvallend afwijkende antwoordoptie
 - f. Absurde antwoordoptie
9. Zorg dat alle afleiders in principe plausibel zijn.
10. Gebruik typische denkfouten die studenten maken om afleiders te schrijven (die kun je verzamelen tijdens het onderwijs...)

Waarom gebruiken?

Gesloten vragen worden snel nagekeken. Bij grote aantallen studenten zijn gesloten vragen dus erg bruikbaar. Omdat je een eenduidige antwoordsleutel hebt, beoordeel je iedere student op dezelfde manier. Open toetsvormen zijn vaak minder objectief te beoordelen.

[De toets- en tentamenservice van het ICLON](#) kan de vragen automatisch nakijken aan de hand van jouw antwoordsleutel. Het ICLON voert dan ook een kwaliteitsanalyse uit (zie ook hoofdstuk 5) en berekent de cesuur (zak/slaaggrens) van het tentamen. Deze analyse en cesuurberekening kan overigens ook op open vragen worden uitgevoerd. Aan de toets- en tentamenservice zijn kosten verbonden.

Meer weten?

- Van Berkel, H., Bax, A. & Joosten-ten Brinke, D. (2014). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15(3), 309-333. doi: 10.1207/S15324818AME1503_5
- Informatie over toetsen via Blackboard (en andere mogelijkheden van Blackboard): <http://hum.leidenuniv.nl/medewerkers/ict-voorzieningen/ontdek-de-tools/blackboard2.html>
- ICLON Toets- en tentamenservice: <http://www.universiteitleiden.nl/iclon/ho/toets--en-tentamenservice>

2.2 Open toetsvragen

Wat?

Open vragen zijn toetsvragen waarbij studenten zelf actief een antwoord moeten formuleren. Met open vragen kun je niet alleen feitenkennis, begrip en het toepassen van kennis toetsen (zoals bij gesloten vragen), maar ook hogere cognitieve niveaus zoals kritisch analyseren en het vormen van een onderbouwde mening.

Hoe?

Tips voor het formuleren van open vragen⁴

- Bepaal hoeveel toetsvragen je wilt schrijven. Hiervoor kun je vooraf inschatten hoeveel tijd studenten nodig hebben om de vragen te beantwoorden. Zie tabel 1 voor een indicatie van beantwoordingstijden voor open vragen.

Tabel 1: Indicatie beantwoordingstijden open vragen.

Lengte open vraag	Beantwoordingstijd
1 woord of zin	1 minuut
¼ pagina A4	5 minuten
½ pagina A4	10 minuten
1 pagina A4	25 minuten
2 pagina's A4	60 minuten

- Gebruik een toetsmatrijs⁵ om het aantal te behalen punten en de niveaus van de vragen te bepalen. Specificeer ook hoeveel punten studenten maximaal per vraag kunnen krijgen.
- Formuleer eerst het modelantwoord en pas daarna de vraag.
- Schrijf de vragen indien mogelijk samen met collega's.
- Splits de vraag indien nodig op in twee delen: een inleiding met achtergrondinformatie en daarna de vraag zelf.
- Wees specifiek over wat je in het antwoord verwacht, bijvoorbeeld: "Geef drie redenen voor...".
- Stel de vraag in de vorm van een opdracht en gebruik hierbij actieve werkwoorden, zoals je idealiter ook in de leerdoelen hebt geformuleerd.
- Gebruik standaardformuleringen voor de vragen om de vraagformulering consistent te maken, zoals:
 - Toon aan
 - Laat zien
 - Leg uit

Waarom gebruiken?

In tegenstelling tot gesloten vragen vereisen open vragen dat studenten zelf actief een antwoord bedenken. Bij open vragen elimineer je de 'gokkans' vrijwel helemaal. Dit komt de validiteit (meet je wat je wilt weten) ten goede, mits de vragen goed geformuleerd zijn.

⁴ Bron: Van Berkel, Bax & Joosten-ten Brinke (2014).

⁵ Zie hoofdstuk 1 en Bijlage II

Meer weten?

- [Jensma, E. & Hubregtse, M. \(2014\)](#). *Vragen beoordelen: Richtlijnen voor het opstellen van theoretische toetsen en antwoordmodellen*. KCH Examens.
- Van Berkel, H., Bax, A. & Joosten-ten Brinke, D. (2014). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.
- Checklist voor het samenstellen, beoordelen en redigeren van open vragen: zie bijlage III

2.3 Schriftelijke werkstukken: essay, verslag en artikel

Wat?

Met de toetsvorm van een essay, verslag of artikel wordt met name gedoeld op het korte schriftelijke werkstuk, te onderscheiden van een eindwerkstuk of scriptie. Een schriftelijk werkstuk biedt de mogelijkheid om inzicht te krijgen in de beheersingsgraad van de hogere cognitieve niveaus. Vaak gaat het daarbij om vaardigheden als kritisch evalueren, samenhang aantonen of onderbouwd argumenteren. Een schriftelijk werkstuk kan bijvoorbeeld een beargumenteerde stellingname ten opzichte van een gegeven stelling zijn, een kritische bespreking van een artikel, of een betoog over de maatschappelijke relevantie van een onderzoeksresultaat. Deze schriftelijke werkstukken vervaardigen studenten veelal individueel, op een zelf gekozen tijd en op een eigen plek. De toetsomstandigheden zijn dan ook niet of nauwelijks te controleren.

Hoe?

Voor het opstellen van een heldere toetsopdracht voor een schriftelijk werkstuk is het van belang om duidelijk te hebben welke criteria je hanteert ten aanzien van de inhoud (gerelateerd aan de leerdoelen) en de vorm. Deze criteria kun je gebruiken om de opdracht goed te formuleren.

Tips voor het opstellen van een toetsopdracht voor een schriftelijk werkstuk⁶

- De vraagstelling
Geef in de toetsopdracht aan wat je van de student verwacht. Bijvoorbeeld een duidelijke inleiding van de vraagstelling, eenduidige formulering, inbedding in een bepaalde context of relatie met de actualiteit.
- De methode en/of de analysetechnieken
In de toetsopdracht aan de student kun je aangeven dat je bijvoorbeeld een adequate beschrijving van de methode verwacht, en wellicht ook een onderbouwing van de keuze voor juist deze methode. Hetzelfde geldt voor de eventuele analysetechnieken die in het werkstuk gehanteerd worden.
- De presentatie van de bevindingen
Geef in de toetsopdracht aan of je bijvoorbeeld verwacht dat er een consequent ordeningsprincipe wordt toegepast, of dat de student verantwoord gebruik maakt van schema's en figuren.
- De conclusies
Neem in de toetsopdracht bijvoorbeeld op dat de student conclusies logisch moet laten

⁶ Bron: Van Berkel, Bax & Joosten-ten Brinke (2014).

volgen uit de voorafgaande analyse, en dat de conclusies een antwoord moeten geven op de vraagstelling.

- De vorm

Geef ook aan wat je verwacht van de vorm. Vermeld hierbij bijvoorbeeld criteria over omvang (aantal woorden, lettergrootte, regelafstand), correct taalgebruik, leesbaarheid, verwijzingen, alinea-indeling, etc. In de praktijk van het onderwijs wordt wel de regel gehanteerd dat het niet voldoen aan de vormeisen direct tot een onvoldoende leidt ongeacht de kwaliteit van de inhoud (ook wel knock-outcriteria genoemd). Als dit van toepassing is, geef dit dan duidelijk aan bij de opdrachtformulering.

- Goede voorbeelden

Je kunt de student verwijzen naar enkele goede voorbeelden van een vergelijkbaar schriftelijk werkstuk, ter illustratie van de aanwijzingen bij de opdracht.

Waarom gebruiken?

De student heeft veel vrijheid om te laten zien wat hij weet en kan. Het maken van een werkstuk doet een beroep op zelfstandigheid, intellectuele capaciteiten en vermogen tot (zelf)reflectie. Medestudenten kunnen betrokken worden bij het geven van feedback, zodat uitwisseling van kennis en ervaring tussen studenten ontstaat. Zo kan het schriftelijk werkstuk een waardevol onderdeel van het onderwijs zijn.

Een belangrijk aandachtspunt is de betrouwbaarheid van beoordeling bij schriftelijke werkstukken. Het blijkt lastig om heldere beoordelingscriteria op te stellen voor juist de hogere cognitieve niveaus. Als gevolg daarvan worden in de praktijk soms met name vormeisen expliciet gemaakt en niet zozeer de inhoudelijke eisen. Het effect hiervan is dat de feedbackfunctie van de beoordeling beperkt wordt en de student weinig kan bijsturen aan zijn leerproces. Bovendien ontstaat bij het ontbreken van duidelijke inhoudelijke beoordelingscriteria het gevaar, dat studenten met name op vormeisen gaan werken en minder op de hogere leerdoelen.

Meer weten?

- Tips voor het opstellen van schriftelijke opdrachten:
<https://www.uantwerpen.be/nl/faculteiten/antwerp-school-of-education/deelentiteiten/expertisecentrum-hoger-onderwijs/didactische-tips/onderwijstips/archief-onderwijstips/tip-34--drie-kwalite/>
- SCALA - Scaffolding Assessment for Learning (2012). Tools voor digitaal feedback geven en beoordelen van schriftelijke werkstukken: trends en ontwikkelingen.
<https://www.surf.nl/binaries/content/assets/surf/nl/kennisbank/2014/literatuurstudie-tools-voor-digitaal-feedback-geven-en-beoordelen-van-schriftelijke-werkstukken.pdf>
- Van Berkel, H., Bax, A. & Joosten-ten Brinke, D. (2014). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.
- ExpertiseCentrum Hoger Onderwijs (2013). *Vijftig onderwijstips*. Antwerpen-Apeldoorn: Garant.

2.4 Mondelinge toetsen: presentaties

Wat?

Mondelinge presentaties toetsen doorgaans zowel de inhoud van de presentatie als presentatievaardigheden. Door middel van een mondelinge presentatie beoordeel je of de student

de stof beheerst, maar ook of de student bijvoorbeeld een standpunt kan beargumenteren, kritisch kan denken en een discussie kan voeren. Een andere term voor de mondelinge presentatie is 'referaat'.

Deze paragraaf gaat alleen in op mondelinge presentaties, omdat andere mondelinge toetsen (zoals een situatie waarin docent(en) toetsvragen stellen aan een student) minder breed gangbaar zijn in situaties waarin studenten beoordeeld worden.

Hoe?

Vaak is er een specifieke reden om voor mondelinge presentaties als toetsvorm te kiezen: je wilt iets toetsen dat niet op een andere manier kan of het past bijvoorbeeld bij de 'leerlijn academische vaardigheden' dat studenten oefenen met (en beoordeeld worden op) het geven van mondelinge presentaties.

Tips voor het opstellen van een toetsopdracht voor een mondelinge presentatie⁷

- Plan vooruit. Mondelinge presentaties kosten tijd wat betreft voorbereiding en uitvoering. Afhankelijk van het niveau en de ervaring van studenten moet je eventueel instructie geven over presentatietechnieken en tijd geven voor voorbereiding en oefening door studenten.
- Maak voor jezelf helder wat je precies wilt toetsen met de mondelinge presentaties en hoe je dit het beste kunt vormgeven. Bepaal of studenten individuele- of groepspresentaties moeten geven, hoe lang moeten de presentaties duren om de gestelde doelen te kunnen bereiken en binnen welke tijdsperiode de presentaties moeten plaatsvinden (allemaal op dezelfde dag of verspreid over een langere periode).
- Formuleer een heldere en complete toetsopdracht voor studenten, zodat studenten weten wat er van hen wordt verwacht en op welke punten ze beoordeeld worden:
 - Beschrijf de doelen van de presentatie
 - Geef aan hoe lang de presentatie moet duren (evt. met een marge, bijvoorbeeld 10-15 minuten) en uit welke onderdelen de presentatie moet bestaan (bijv. introductie, beschrijving van een onderzoek, conclusie)
 - Beschrijf alle onderdelen van de toetsopdracht, ook wat studenten ter voorbereiding moeten doen (bijv. het vooraf indienen van een concept of literatuurlijst). Geef ook aan of je wilt dat studenten de presentatie inleveren. Je kunt ervoor kiezen om de presentaties te laten opnemen, zodat je later studenten op hun eigen presentatie kunt laten reflecteren.
 - Specificeer belangrijke deadlines (bijv. voor het inleveren van onderdelen van de opdracht en voor het geven van de presentatie)
 - Geef aan op welke punten je de presentatie zult beoordelen

⁷ University of Pittsburgh, Tips for Assigning Oral Presentations <http://speaking.pitt.edu/instructor/oral-assignments.html>

Waarom gebruiken?

Mondelinge presentaties worden vaak gebruikt in leerlijnen voor academische vaardigheden. Met presentaties leren studenten zich goed uitdrukken en kunnen zij aantonen dat ze niet alleen kennis en begrip hebben, maar ook zelf ideeën kunnen genereren. Daarnaast kunnen presentatievaardigheden studenten in latere beroepen goed van pas komen.

Meer weten?

- De Grez, L., Valcke, M., & Roozen, I. (2013). The differential impact of observational learning and practice-based learning on the development of oral presentation skills in higher education. *Higher Education Research & Development*, 33(2), 256-271. doi: 10.1080/07294360.2013.832155
- De Grez, L., Valcke, M., & Roozen, I. (2009). The impact of an innovative instructional intervention on the acquisition of oral presentation skills in higher education. *Computers & Education*, 53(1), 112-120. doi: <http://dx.doi.org/10.1016/j.compedu.2009.01.005>
- Tips voor het maken en beoordelen van mondelinge toetsopdrachten: <http://www.speaking.pitt.edu/instructor/oral-assignments.html>

2.4 Toetsen van groepswerk: schriftelijke werkstukken of mondelinge presentaties

Wat?

Bij het toetsen van groepswerk beoordeel je over het algemeen een schriftelijk werkstuk of een mondelinge presentatie waar een groep studenten gezamenlijk aan heeft gewerkt. In theorie kun je ook een groep observeren en beoordelen, terwijl zij met elkaar aan het werk zijn, maar dat komt in de praktijk van het hoger onderwijs weinig voor. Deze paragraaf behandelt daarom het opstellen van een schriftelijke of mondelinge groepsopdracht.

Hoe?

De tips voor het opstellen van schriftelijke werkstukken (2.3) en mondelinge presentaties (2.4) kun je ook gebruiken voor het opstellen van schriftelijke of mondelinge groepsopdrachten. Bij groepsopdrachten moet je naast aanwijzingen voor het schriftelijke of mondelinge 'product' ook beschrijven wat je van de samenwerking tussen studenten verwacht (het 'proces') en hoe je er zorg voor draagt dat er een eerlijke beoordeling tot stand komt. Naast inhoudelijke beoordelingscriteria, zul je bij groepsopdrachten extra letten op *hoe* het product tot stand is gekomen, je neemt dus ook criteria op t.a.v. het groepsproces; idealiter heeft elk groepslid een evenredig aandeel in het groepsproduct gehad. Duidelijkheid (vooraf) over het groepsproces en de beoordeling van de groep en van individuen draagt in hoge mate bij aan de effectiviteit van de (groeps-)leerervaring van studenten.

Tips voor het opstellen van groepsopdrachten

- Geef aan wat het doel van de groepsopdracht is. Beschrijf niet alleen de inhoudelijke doelen, maar - indien van toepassing - ook doelen op gebied van samenwerking en communicatie.
- Bepaal hoe groepen worden gevormd; door studenten onderling of door docent.
- Bepaal of studenten zelf afspraken moeten maken over het groepsproces óf dat je als

docent richtlijnen geeft over de organisatie van het groepsproces (bijv. hoe groepsrollen verdeeld worden, welke inzet van elk groepslid wordt verwacht, dat studenten elkaar feedback geven).

- Geef handvatten aan studenten voor het invullen van verschillende groepsrollen en verantwoordelijkheden. De rollen kunnen ook rouleren tussen studenten, zodat elke student dezelfde ervaring kan opdoen. Afhankelijk van de opdracht gaat het bijvoorbeeld om het coördineren van de opdracht, het notuleren van groepsoverleg, opzoeken van literatuur over verschillende thema's, het samenvatten van artikelen, het uitvoeren van (onderdelen van) een onderzoek, het houden van (onderdelen van) een presentatie.
- Geef aan hoeveelheid tijd er voor de groepsopdracht moet worden uitgetrokken.
- Beschrijf op welke punten je de groepsopdracht beoordeelt. Belangrijk om hierbij aan te geven of je de groep als geheel of individuen beoordeeld (zie ook 3.2.4).
- Overweeg om studenten tussenproducten op te laten leveren of logboeken over de voortgang (en ieders bijdrage) bij te laten houden, om zo zicht te houden op het groepsproces.
- Ruim eventueel tijd in het onderwijs in om groepsopdrachten tussentijds te bespreken.

Waarom gebruiken?

Samenwerken en het gezamenlijk werken aan een product is een competentie die van veel afgestudeerden wordt gevraagd, het is daarom zinvol om daar tijdens de opleiding aandacht aan te besteden en studenten daar ook op te beoordelen. Binnen groepsopdrachten kunnen studenten van elkaar leren. Sommige studenten ervaren een sterkere motivatie om aan de slag te gaan; andere studenten zitten immers op hun bijdrage te wachten. Tegelijkertijd kunnen groepsopdrachten ook 'meeliftgedrag' in de hand werken. Daarom is het belangrijk om duidelijke afspraken te maken over ieders bijdrage en studenten te stimuleren elkaar feedback te geven op geleverd (of niet geleverd...) werk.

Meer weten?

- Over het beoordelen van groepsopdrachten: James, R., McInnis, C. & Devlin, M. (2002). *Assessing Learning in Australian Universities. Centre for the Study of Higher Education, Australian Universities Teaching Committee.*
- Over problemen bij en oplossingen voor het beoordelen van groepswork: Davies, W.M. (2009). *Groupwork as a form of assessment: common problems and recommended solutions. Higher Education, 58, 563-584.*

VUISTREGELS VOOR HET SAMENSTELLEN VAN TOETSEN

- Een toets kan als meetinstrument alleen bruikbare informatie opleveren als aan de volgende kwaliteitseisen wordt voldaan:
 - validiteit: de toets meet wat het beoogt te meten, is representatief voor datgene wat men met de toets wil nagaan;
 - betrouwbaarheid: de toets meet op verschillende momenten bij verschillende studenten en beoordelaars hetzelfde;
 - transparantie: de student weet wat er van hem of haar verwacht wordt;
 - bruikbaarheid: de toetsvorm past bij de leerstof, leerdoelen en groepsgrootte. De toets past binnen de financiële en/of materiële kaders, maar is ook binnen de beschikbare tijd af te nemen.

Streef naar een optimum van deze kwaliteitseisen zodat het meetinstrument daadwerkelijk een uitspraak kan doen over de vorderingen en stand van zaken in het leerproces van de student. Een goed meetinstrument zorgt ervoor dat studenten niet ten onrechte zakken of slagen, maar ook dat studenten een reëel beeld krijgen van hun leerproces.

- Zorg ervoor dat je tentamen (dat uit meerdere toetsen kan bestaan) aansluit op alle leerdoelen van het vak. Aan het eind van het vak wil je een uitspraak kunnen doen over het al dan niet behalen van alle leerdoelen.
- Kies een toetsvorm die aansluit op het niveau van de leerdoelen. Gebruik bijvoorbeeld geen meerkeuzevragen als studenten een standpunt moeten beargumenteren.
- Zorg voor een evenwichtige verhouding van studielast per toets en de mate waarin de toets meeweegt in het eindcijfer. Het staat bijvoorbeeld niet in verhouding om een werkstuk van 20 pagina's slechts 10% te laten meewegen in het eindcijfer.
- Formuleer heldere, eenduidige vragen of opdrachten. Vermijd ingewikkelde zinsconstructies, dubbele ontkenningen, onnodige tussenvoegsels, etc. (Zie ook Bijlage III)

3. Samenstellen van beoordelingsmodellen

Beoordelingsmodellen bestaan in verschillende vormen, van antwoordsleutel voor gesloten vragen tot rubrics voor bijvoorbeeld werkstukken of presentaties. Bij gesloten toetsvormen is het duidelijk wat de juiste antwoorden zijn. Antwoordsleutels zijn dan ook eenvoudig te maken. In dit hoofdstuk komen daarom alleen beoordelingsmodellen voor open toetsvormen aan orde, gevolgd door specifieke tips en aandachtspunten voor de beoordeling van:

- schriftelijke werkstukken (zoals essays en papers)
- eindwerkstukken (scripties)
- mondelinge presentaties
- groepsopdrachten

Ook komt aan bod hoe je, na het samenstellen van een beoordelingsmodel, de cesuur (= grens voldoende/onvoldoende) bepaalt.

3.1 Beoordelingsmodellen voor open toetsvormen

Wat is een beoordelingsmodel?

Een beoordelingsmodel is een model, een set richtlijnen voor het geven van een cijfer voor open toetsvormen. Waar gesloten toetsvormen eenvoudigweg een antwoordsleutel hanteren, hebben open toetsvormen een uitgebreider model voor het nakijken en beoordelen van de toets. Zo'n beoordelingsmodel kan verschillende vormen aannemen, van een korte beschrijving van het juiste antwoord tot een uitgebreid beoordelingsformulier (vaak een *rubric*).

Hoe stel je een beoordelingsmodel samen?

Een beoordelingsmodel bestaat uit drie onderdelen: een antwoordmodel (of *rubric*), een scoringsvoorschrift en beoordelaarsinstructies. Hieronder wordt uitgelegd hoe je deze drie onderdelen samenstelt.

1. Het *antwoordmodel* is een overzicht van juiste, deels juiste en foute antwoorden. Een antwoordmodel bevat het volgende:
 - Een beschrijving van het goede of ideale antwoord. Dit kunnen ook *beoordelingscriteria* zijn als de vraag geen duidelijk goed of fout antwoord heeft. Denk hierbij aan vragen waarbij studenten zelf met een mening op oplossing moeten komen.
 - Het maximum aantal punten per (deel)vraag.
 - Voorbeelden van foute antwoorden en / of veelgemaakte fouten. Het is handig deze vast in het antwoordmodel te noemen, zodat de beoordelaar weet welke fouten hij tegen kan komen.

Een *rubric* is een veel gebruikt type beoordelingsformulier voor het beoordelen van en feedback geven op open toetsvormen. Met een *rubric* operationaliseer je beoordelingscriteria naar een schaalindeling met duidelijke uitleg van ieder niveau (de zogenaamde 'prestatie-indicatoren'). Daarmee kun je aangeven of een student een criterium onvoldoende, voldoende of goed beheerst. De schaalindeling kan natuurlijk ook voor een cijfermatige beoordeling worden gebruikt. Het is wel van belang dat je, als je kwalitatieve scores uiteindelijk moet omzetten naar een cijfer, je vooraf bepaalt welk cijfer (of

cijferrange) je aan welke kwalitatief oordeel koppelt (bijvoorbeeld: ‘onvoldoende’ betekent ≤ 5 , ‘voldoende’ betekent een 6 of 7 en ‘goed’ betekent ≥ 8).

Tabel 2. Voorbeeld van een rubric voor een mondelinge presentatie (gedeelte).

Beoordelingscriterium	Onvoldoende	Voldoende	Goed
Inhoud en doel van de presentatie zijn afgestemd	De inhoud van de presentatie komt niet overeen met het doel en de informatie is niet volledig.	De inhoud van de presentatie komt overeen met het doel en bevat de benodigde informatie.	Inhoud en doel van de presentatie vormen een samenhangend geheel en de presentatie bevat de benodigde informatie.
Gebruik van bronnen	De student heeft geen bronnen geraadpleegd.	De student heeft enkele relevante bronnen geraadpleegd.	De student heeft meerdere relevante bronnen geraadpleegd.
Gebruik van voorbeelden	De student heeft geen voorbeelden gegeven.	De student heeft een of enkele voorbeeld(en) gegeven.	De student heeft verschillende adequate voorbeelden gegeven.

Vrij naar: Van Berkel et al, 2014

Het is belangrijk dat de prestatie-indicatoren zo concreet mogelijk omschreven zijn. De prestatie-indicatoren tonen namelijk de feedback die je aan de student geeft. Een rubric is dus niet alleen een middel om te beoordelen, maar ook om efficiënt en eenduidig feedback te geven aan studenten. Daarnaast kunnen studenten ook zichzelf of elkaar beoordelen met een rubric (*self en peer assessment*). Eenmaal ontwikkelde rubrics kun je in opeenvolgende jaren gebruiken voor dezelfde toets. Dat maakt de tijdsinvestering de moeite waard. Het is wel van belang de rubrics iedere keer bij te werken op basis van gebruikservaring. Het verdient de voorkeur een rubric samen met een of meer collega's te ontwikkelen.

Stappenplan voor het opstellen van een rubric⁸

1. *Reikwijdte bepalen.* Is de te beoordelen prestatie iets specifiek voor jouw vak, of meer een algemene vaardigheid? Vooral in het laatste geval is het handig om bij collega's na te vragen of zij rubrics hiervoor hebben ontwikkeld. Rubrics voor bijvoorbeeld mondelinge presentaties of werkstukken kunnen (deels) in verschillende disciplines toepasbaar zijn.

⁸ gebaseerd op Van den Bos, Burghout & Joosten-ten Brinke, 2014

2. *Koppelen aan leerdoelen.* Maak expliciet welke leerdoelen de rubric toetst.
3. *Beoordelingscriteria opstellen.* Welke kennis en/of vaardigheden wil je toetsen? Wil je ook de houding van de student beoordelen, bijvoorbeeld actieve bijdrage aan groepswork? Laat de criteria aansluiten op de leerdoelen en probeer tot een werkbaar aantal criteria te komen. Vijf tot acht blijkt in de praktijk haalbaar. Te veel en te gedetailleerde beoordelingscriteria leveren een schijnnaauwkeurigheid op, die afbreuk doet aan de waarde ervan voor feedback en beoordeling.
4. *Aantal niveaus bepalen.* Drie-, vier- of vijf niveaus zijn vaak goed bruikbaar; bij meer niveaus wordt het lastig om onderscheidende prestatie-indicatoren te bedenken.
5. *Prestatie-indicatoren schrijven.* Zorg dat de prestatie-indicatoren werkelijk van elkaar te onderscheiden zijn en hanteer daarbij meetbare, observeerbare formuleringen. Begin met het hoogste niveau: wat is de beste prestatie die een student kan tonen? Ga daarna verder met het laagste niveau: wat is het slechtste dat een student kan tonen? Vul dan de middelste niveaus in.
6. *Normering bepalen.* Je kunt verschillende gewichten toekennen aan verschillende beoordelingscriteria als er criteria zijn die je zwaarder mee wilt laten wegen dan andere. Als je verwacht dat studenten calculerend zijn en onvoldoende gescoorde criteria gaan compenseren met voldoende beoordeelde criteria, dan kun je als eis aangeven dat alle criteria voldoende moeten zijn. Je kunt eerder beoordeelde producten gebruiken als voorbeeld (ijkpunt) voor elk niveau. Nog een tip is om samen met collega's enkele proefbeoordelingen uit te voeren om de rubric uit te testen.

2. Een *scoringsvoorschrift* bevat in detail hoe punten moeten worden toegekend aan (deels) goede en (deels) foute antwoorden. Instructies die in een scoringsvoorschrift kunnen staan zijn:

- Rond scores af op hele / halve punten.
- Geen antwoord = ... punten.
- Correct = ... punten erbij, fout = ... punten aftrek.
- Laat fouten wel / niet doorwerken in vervolgvragen (meestal kiezen docenten ervoor fouten niet te laten doorwerken).
- Spel- en grammaticafouten = ... puntenaftrek.
- Scoor alleen de eerste twee antwoorden (als dit een antwoordrestrictie in de vraag was!).

3. *Beoordelaarsinstructies* zijn algemene richtlijnen voor de beoordelaar. Deze instructies moeten valkuilen bij het beoordelen tegengaan. Voorbeelden van beoordelaarsinstructies zijn:

- Kijk tentamens per vraag na, niet per student (het zogenaamde 'horizontale nakijken').
- Wissel de volgorde van de tentamens zo nu en dan.
- Geef pas cijfers als alle vragen zijn nagekeken.
- Voeg nieuw ontdekte fouten toe aan het antwoordmodel.

- Voeg nieuw ontdekte goede antwoorden toe aan het antwoordmodel (check ook of hier referenties voor zijn).
- Kijk eerst samen met een collega onafhankelijk van elkaar enkele open vragen na en toets daarna in hoeverre jullie overeenstemmen. Ga pas verder met nakijken als jullie beiden op één lijn zitten qua beoordelen.

Naast het hanteren van een beoordelingsmodel kun je, ter verhoging van de betrouwbaarheid van de beoordeling, de volgende maatregelen nemen:

- Beoordeel de antwoorden per vraag (in plaats van een gehele toets per student)
- Wijzig de volgorde van de antwoordbladen per vraag
- Kijk geanonimiseerd na
- Overweeg om meerdere beoordelaars in te zetten: hoe groter de gevolgen van de toetsuitslag voor student, hoe hoger de beoordelaarsbetrouwbaarheid dient te zijn. Indien er meerdere beoordelaars betrokken zijn:
 - Verdeel de vragen (niet de studenten) over de beoordelaars
 - Vraag collega's bij twijfelgevallen opnieuw te beoordelen

Waarom zou ik gebruik maken van een beoordelingsmodel?

Bij het beoordelen van open toetsvormen kan het lastig zijn om dit op *consistente* en dus betrouwbare wijze te doen. De volgende *beoordelaarseffecten* kunnen optreden⁹:

- *Normverschuiving*: de neiging van de beoordelaar zich in de strengheid of mildheid van de beoordeling aan te passen aan het gemiddelde prestatieniveau van de groep. Ook komt het voor dat tijdens het beoordelingsproces de examinerator steeds milder wordt, bijvoorbeeld door vermoeidheid.
- *Sequentie-effect*: de over- of onderwaardering van een antwoord onder invloed van de kwaliteit van voorafgaande antwoorden van andere studenten.
- *Halo-effect*: de kennis die de beoordelaar heeft over voorgaande prestaties van de student is mede van invloed op de beoordeling.
- *Signifisch effect*: het hanteren van verschillende criteria bij de beoordeling door meerdere beoordelaars.
- *Contaminatie-effect*: de vrijheid in de beoordeling wordt (bedoeld of onbedoeld) gebruikt voor andere doeleinden dan die van een onbevooroordeelde beoordeling. Hiervan kan bijvoorbeeld sprake zijn als de beoordelaars iets lagere scores geven om aan te geven dat hun discipline moeilijk is.
- *Restriction-of-range*: de persoonlijke neiging van de beoordelaar kan zijn om alle waarden in de beoordelingsschaal te gebruiken (scores van 1 tot 10), terwijl een ander waarden in een smallere range toekent (bijvoorbeeld scores van 5 tot 8).

Het werken met een goed beoordelingsmodel beperkt dergelijk nadelige effecten. De beoordelaarsinstructies gaan beoordelaarseffecten tegen en zorgen er zodoende voor dat je consistentere (betrouwbaarder) nakijkt. Een goed correctievoorschrift zorgt er ook voor dat je sneller kunt nakijken, wat de toets bruikbaar maakt.

⁹ Kallenberg et al (2014); Van Berkel et al (2014).

Daarnaast is het ook voor de student belangrijk te weten hoe een beoordeling tot stand komt. Het gebruik van antwoordmodellen, scoringsvoorschriften en beoordelaarsinstructies zorgt voor een hogere transparantie van de beoordeling, omdat je expliciet maakt wat en hoe je beoordeelt.

Als je geen beoordelingsmodel of/ -formulier opstelt, schakel dan een tweede beoordelaar in om de betrouwbaarheid en validiteit te verhogen, zeker bij opdrachten die je zelf als 'middelmatig' beoordeelt.

Meer weten?

- [Jensma, E. & Hubregtse, M. \(2014\). Vragen beoordelen: Richtlijnen voor het opstellen van theoretische toetsen en antwoordmodellen.](#) KCH Examens.
- Van Berkel, H., Bax, A. & Joosten-ten Brinke, D. (2014). *Toetsen in het hoger onderwijs.* Houten: Bohn Stafleu van Loghum.
- Over beoordelingsmodellen (met voorbeelden van beoordelingsmodellen voor taalvaardigheid): <http://www.toets.nl/uploads/Artikelen/Het-belang-van-een-goed-beoordelingsmodel.pdf>

3.1.1 Beoordeling van schriftelijke werkstukken (zoals essays en papers)

Bij het beoordelen van schriftelijke werkstukken is een zekere mate van subjectiviteit van de beoordelaar onvermijdelijk. De beoordeling wordt enigszins geobjectiveerd door te werken met een *gestandaardiseerd beoordelingsinstrument*, en door zo nodig een tweede beoordelaar in te schakelen.

Tips voor de beoordeling van schriftelijke werkstukken¹⁰

- Schrijf van tevoren de beoordelingscriteria uit, zowel op gebied van inhoud als vorm. Maak zoveel mogelijk gebruik van gestandaardiseerde beoordelingsinstrumenten.
- Indien mogelijk, bediscussieer de beoordelingscriteria met collega's.
- Let erop dat het instrument en de weging van de beoordelingscriteria niet te veel het belang van de lagere cognitieve prestaties benadrukken (zoals vormaspecten van het werkstuk) ten koste van de hogere cognitieve prestaties (zoals kritisch evalueren of onderbouwd argumenteren).
- Goede beoordelingsinstrumenten bevatten *specifieke* criteria. Dat wil zeggen dat ieder criterium duidelijk anders is dan c.q. niet overlappend is met de andere criteria.
- Lees (een aantal van) de schriftelijke werkstukken eerst globaal door; dit kan een indruk geven over de te hanteren norm.
- Selecteer enkele 'modellen' van zeer goede, goede, voldoende en onvoldoende uitwerkingen, als basis voor het beoordelen van de rest van de schriftelijke werkstukken. Dit is vooral belangrijk als er sprake is van verschillende beoordelaars of wanneer de

¹⁰ Deels ontleend aan 'Onderwijstips' van het Expertisecentrum Hoger Onderwijs, Universiteit Antwerpen <https://www.uantwerpen.be/nl/faculteiten/antwerp-school-of-education/deelentiteiten/expertisecentrum-hoger-onderwijs/didactische-tips/onderwijstips/archief-onderwijstips/tip-43--een-schrijfe/>

beoordeling verspreid wordt over een ruimere tijdsperiode.

- Kijk indien mogelijk het werk na zonder lange onderbrekingen; dit bevordert de homogeniteit van de oordelen.
- Naast een wat meer *analytische beoordeling* aan de hand van een beoordelingsformulier of –schema kun je ook (met name als ervaren docent) een *globale beoordeling* uitvoeren op grond van de totale indruk die het werkstuk maakt. In de praktijk wordt niet zelden een in klad ingevuld beoordelingsformulier nog licht bijgesteld om het daaruit afgeleide eindoordeel meer in overeenstemming te brengen met het globale oordeel.
- Schrijf feedback aan studenten op, gerelateerd aan de beoordelingscriteria, om de nabespreking de vergemakkelijken en de beoordeling voor studenten inzichtelijker te maken.
- Overweeg eens om studenten een rol te geven in het beoordelen van werkstukken van medestudenten (*peer assessment*). Ook hiervoor geldt dat de beoordelingscriteria helder moeten zijn. Wanneer studenten in een groep werken aan een gezamenlijk product, kan peer assessment worden gebruikt om het groeps cijfer te corrigeren voor de bijdrage van de verschillende studenten aan het groepswork. Studenten kunnen elkaar beoordelen op medewerking bij groepswork, of ze kunnen elkaar inhoudelijke feedback geven. Van belang is wel om studenten te trainen/begeleiden in het geven van beoordelingen. Naast het trainen van studenten, kunnen (producten van) studenten ook telkens door twee of meerdere studenten beoordeeld worden, in combinatie met een beoordeling (of steekproef) door de docent.

3.1.2 Beoordeling van eindwerkstukken (scripties)

Het belang van het eindwerkstuk binnen het gehele toetsprogramma van de opleiding is zeer groot. Als afsluitend studieonderdeel worden tenslotte de meeste eindtermen tot op het eindniveau van de opleiding met het eindwerkstuk getoetst. Lees voor de regels en richtlijnen rond (het beoordelen van) eindwerkstukken ook de ‘Handleiding Examencommissies’ dat door de faculteit Geesteswetenschappen beschikbaar is gesteld.

Tips voor de beoordeling van eindwerkstukken

- Gebruik voor de beoordeling van eindwerkstukken, scripties of onderzoeksverslagen in de afsluitende fase van de opleiding ALTIJD het (facultaire) gestandaardiseerd beoordelingsformulier.
- In grote lijnen gelden voor de beoordeling van eindwerkstukken dezelfde kanttekeningen als voor andere schriftelijke werkstukken. Subjectiviteit van de beoordelaar is onvermijdelijk, maar voor ervaren beoordelaars die al veel eindwerkstukken hebben beoordeeld geldt dat de *globale beoordeling* op grond van de totale indruk die het werkstuk maakt een belangrijke raadgever is in het oordeel over de algehele prestatie van de student.
- Het beoordelen van het eindwerkstuk door een tweede lezer is ALTIJD voorgeschreven. De twee beoordelaars dienen bovendien hun beoordeling *aantoonbaar onafhankelijk* uit te voeren. Dat betekent in de praktijk dat je (bij voorkeur) in eerste instantie beoordeelt

zonder voorkennis van het oordeel van de andere beoordelaar. De beide onafhankelijke oordelen leiden vervolgens na onderling overleg tot een gezamenlijk eindoordeel plus bijbehorende feedback richting de student. Voor de herbeoordeling van eindwerken kunnen examencommissies, midtermcommissies en visitatiecommissies de gezamenlijke eindbeoordeling én de onafhankelijke onderliggende oordelen willen inzien.

- Als begeleider die rechtstreeks betrokken is geweest bij het onderzoek en/of het schrijfproces van de student beoordeel je het eindwerkstuk niet alleen op de kwaliteit van de eindversie, maar neem je het proces dat de student tot het eindproduct heeft gebracht in het oordeel mee. De tweede beoordelaar is doorgaans veel minder goed in staat om zich een oordeel te vormen over dat totstandkomingsproces. Dit verschil kan doorwerken in de beide initiële onafhankelijke oordelen, maar draagt bij aan de objectivering van de beoordeling en zal na onderling overleg doorgaans geen belemmering vormen voor het opstellen van een gezamenlijk eindoordeel.
- In het beoordelingsformulier dient bij voorkeur duidelijk onderscheid gemaakt te worden tussen beoordelingscriteria gericht op enerzijds bv. de uitvoering van onderzoek dat heeft geleid tot het eindwerkstuk of het schrijfproces zelf, en anderzijds op de kwaliteit van het definitieve eindwerkstuk.

3.1.3 Beoordeling van mondelinge presentaties

Een zekere mate van subjectiviteit en het optreden van beoordelareffecten spelen hoe dan ook een rol bij de beoordeling van mondelinge presentaties. Vergelijkbaar met de beoordeling van schriftelijke werkstukken beoordeel je mondelinge presentaties bij voorkeur dan ook door middel van een gestandaardiseerd beoordelingsinstrument.

Tips voor het beoordelen van mondelinge presentaties

- Met beoordelingsinstrumenten voor mondelinge presentaties moet je doorgaans zowel hogere cognitieve niveaus (bijvoorbeeld analyse van de leerstof, argumentatie, integratie, vergelijking en oplossing) als communicatieve en presentatievaardigheden van de student kunnen beoordelen. Maak in het instrument een duidelijk onderscheid tussen deze twee aspecten.
- Een rubric voor presentatievaardigheden kan bestaan uit drie inhoudelijke criteria (introdactie, structuur en conclusie), vijf criteria m.b.t. de presentatiestijl (oogcontact, stemgebruik, enthousiasme, interactie met het publiek en lichaamstaal), en een algemeen (eind-)oordeel.
- Een mondelinge presentatie kan nadelig beïnvloed worden door persoonskenmerken van de student als verlegenheid of plankenkoorts. Zo'n student moet ook beoordeeld worden en feedback krijgen op presentatievaardigheden, maar juist dan is het wel van belang om duidelijk onderscheid te maken in de verschillende aspecten van de beoordeling.
- Stem de weging van de verschillende beoordelingscriteria af op het belang van de eraan gelieerde leerdoelen.
- Informeer de studenten vooraf goed over de aard, inhoud en werkwijze van de

beoordeling.

- Werk indien mogelijk met twee beoordelaars, om beoordelingen te vergelijken en eventueel tot een gemiddelde te komen. Dit draagt bij aan de betrouwbaarheid van de beoordeling en aan de kwaliteit van de feedback.
- Overweeg om medestudenten een rol te geven in het ondersteunen van de beoordeling en het verrijken van de feedback naar de student.

3.1.4 Beoordeling van groepsopdrachten

Een zorg van veel studenten is dat hun persoonlijke prestatie niet eerlijk wordt beoordeeld bij beoordelingen van groepsopdrachten. Wees duidelijk naar studenten over wat het doel is van de groepsopdracht, wat en hoe beoordeeld wordt en overweeg het volgende:

- Beoordeel je het groepsproduct, het proces van het groepswerk of beide?
 - als je zowel proces als product beoordeelt, hoe zwaar laat je beide aspecten meewegen in het eindoordeel?
- Welke criteria hanteer je om de groepsopdracht te beoordelen?
 - en bepaal je die zelf, laat je het door studenten bepalen, of doe je het gezamenlijk?
- Wie beoordeelt het proces en/of product? Docent, student(en) of beide?
- Hoe geef je een cijfer: één gezamenlijk groeps cijfer, één groeps cijfer dat het gemiddelde is van de verschillende groepsleden, individuele cijfers of een combinatie van een groeps cijfer en een individueel cijfer (dit laatste verdient de voorkeur)?

Tips voor het beoordelen van groepsopdrachten

- Geef studenten een rol in het vaststellen van criteria voor het beoordelen van het groepsproces. Dit geeft hen inzicht in waar het om gaat wat betreft het groepswerk en zorgt dat zij zich meer eigenaar voelen van het groepsproces.
- Verstrek beoordelingsinstrumenten vooraf aan studenten.
- Betrek de feedback en het oordeel van studenten in de eindbeoordeling van studenten.
- Beoordeel niet alleen eindproducten, maar ook tussenproducten zoals logboeken, notulen van overleggen of plannen van aanpak.
- Geef tussentijdse feedback, zodat studenten hun aanpak eventueel kunnen aanscherpen.
- Maak de beoordelingsprocedure inzichtelijk.
- Wanneer je studenten betreft in de beoordeling: zorg dat je als docent de regie en eindverantwoordelijkheid houdt. In principe kunnen studenten feedback geven of een advies uitbrengen, ze geven geen formeel cijfer.

Meer weten?

- Bax, A., Perrenet, J. & Berkel, H. van (2014). Toetsen van groepswerk. In: Berkel, H van, en Bax, A. (red); *Toetsen in het Hoger Onderwijs*. Houten: Bohn Stafleu van Loghum
- Over manieren om groepswerk te beoordelen: <http://tauu.uu.nl/wp-content/uploads/2015/01/Zeven-manieren-om-tot-een-individueel-cijfer-te-komen-bij-groepswerk.pdf>

3.2 Cesuurbepaling

Wat is de cesuur?

Voor alle toetsvormen geldt dat je moet bepalen wanneer studenten voldoende hebben gepresteerd. De cesuur is de zak/slaaggrens bij de beoordeling van een toets. Deze kan op drie manieren worden bepaald: absoluut, relatief of een mengvorm. In de praktijk worden vaak mengvormen toegepast.

Hoe bepaal je de cesuur?

- Bij de *absolute* methode houd je geen rekening met de score van overige studenten en houd je een vaste maat aan. Die maat is meestal een bepaald percentage van de vragen dat minimaal goed beantwoord moet zijn. Studenten die 75% van de vragen goed hebben, krijgen dan bijvoorbeeld een 7,5 als cijfer. Vaak wordt de cesuur ofwel de zak-/slaaggrens bepaald op 55% van de maximaal te behalen punten. Als in de toetsmatrijs een evenwichtige puntenverdeling over de leerdoelen is aangebracht, dan is dit ook een logische grens. Deze methode heeft als nadeel dat eventuele zwakheden in de toets zelf niet in de becijfering gecorrigeerd worden.
- De *relatieve* methode houdt in dat je het cijfer van een student bepaalt door deze te vergelijken met de andere deelnemers aan de toets. Het cijfer wat een student behaalt is dus afhankelijk van wat de andere toetsdeelnemers presteren. Als een student bijvoorbeeld 90 van de 100 te behalen punten heeft behaald op een toets, en de medestudenten hebben allen een lager aantal punten behaald, dan zal deze student een hoog cijfer krijgen. Van Berkel et al (2014) hanteert de metafoor van de ladder: de hoogst scorende student staat op de bovenste sport, de laagst scorende op de onderste. De rest van de studenten worden daartussen evenredig verdeeld. De relatieve methode wordt zelden in pure vorm toegepast omdat hij fraudegevoelig is (studenten kunnen afspraken maken over de voorbereiding op de toets) en bij kleinere aantallen geen betrouwbare uitslag geeft.
- In de praktijk wordt meestal een *mengvorm* gebruikt: er wordt een absolute norm gehanteerd, met de mogelijkheid van een correctie achteraf op grond van relatieve scores. Dat wil zeggen: als de toetsscores van een groep als geheel erg afwijken van wat gemiddeld met een dergelijke toets wordt behaald, kun je achteraf een correctie op de cijfers uitvoeren. Bijvoorbeeld: als meer dan 40% van de toetsdeelnemers een onvoldoende heeft behaald, en het gaat om een nieuw ontwikkelde toets waarvan de moeilijkheidsgraad nog niet helemaal duidelijk is, dan kun je overwegen de cesuur te verleggen.

Meer weten?

- Van Berkel, H., Bax, A. & Joosten-ten Brinke, D. (2014). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.
- Over het proces van cesuurbepaling (met aantal rekenvoorbeelden):
<http://www.toets.nl/uploads/Downloads%20van%20Bureau%20ICE/VO/Artikel%20toets%20wanneer%20geef%20ik%20een%205,5.pdf>

4. Afnemen van toetsen

Wat?

Bij het afnemen van toetsen is het belangrijk om fraude te voorkomen en herkennen. Onder fraude verstaan we alles wat maakt dat het resultaat van een toets niet meer direct te herleiden is tot het resultaat van een (groep) student(en); van spieken, plagiaat tot meeliften bij samenwerkingsopdrachten. Elke faculteit heeft 'Regels en Richtlijnen voor Examencommissies' vastgesteld, hierin staat onder andere aan welke eisen de toetsafname moet voldoen.

Hoe?

Tips om spieken te voorkomen:

- Expliciteer vooraf welke hulpmiddelen wel/niet zijn toegestaan (bijvoorbeeld wel/geen gebruik van woordenboeken). Gebruik van mobiele telefoon is altijd verboden.
- Zorg voor voldoende en actieve surveillanten tijdens een tentamen. Het aantal surveillanten wordt bepaald door het aantal aangemelde studenten en het aantal zalen waar het tentamen is ingedeeld. Overweeg om bij meer dan 50 studenten twee surveillanten in te zetten.
- Zorg ervoor dat studenten niet bij elkaar af kunnen kijken, door de onderlinge afstand tussen studenten te bewaken (minimaal twee meter links/rechts en voor/achter) en door eventueel met verschillende versies van toetsen te werken (wel zelfde vragen, maar andere nummering) en die om-en-om uit te delen.

Tips om [plagiaat te voorkomen](#):

- Informeer studenten over de wijze waarop ze met bronnen en bronvermeldingen moeten omgaan.
- Bespreek plagiaat met studenten; het wordt niet door elke student herkend.
- Maak het facultaire beleid met betrekking tot plagiaat bekend bij studenten.
- Maak opdrachten [minder gevoelig voor plagiaat](#), door bijvoorbeeld:
 - Specifieke en duidelijk omschreven opdrachten te geven;
 - Een lijst van onderwerpen voor studenten op te stellen;
 - Duidelijk te maken welke specifieke elementen het werk moet bevatten;
 - Studenten werkstukken gefaseerd te laten schrijven en inleveren. Op deze wijze heeft de docent meer zicht op de totstandkoming van het werk.

Tips voor het [herkennen van plagiaat](#):

- Soms duiden [kenmerken uit de tekst](#) op plagiaat, bijvoorbeeld:
 - verschillende citeer- en referentieconventies binnen een tekst;
 - compleet ontbreken van citaten en referenties;
 - opvallende opmaak, bijvoorbeeld met regels die halverwege afbreken;
 - gedateerde informatie.
- Maak gebruik van een tool voor het opsporen van plagiaat, zoals [Turnitin](#)

Tips voor voorkomen van meeliften:

- Vraag studenten een duidelijke planning met taakverdeling te maken. Laat notulen maken zodat duidelijk is wie verantwoordelijk is voor welk onderdeel en hoe het groepsproces is verlopen;

- Stel actieve begeleidingsvragen, afwisselend aan de groep en aan individuen;
- Als er een mondelinge presentatie moet worden gegeven, geeft dan pas op de dag zelf aan wie van de groep dat moet doen, zodat elke student zich op de presentatie voorbereid (en zich dus met de inhoud heeft beziggehouden);
- Bij eventuele problemen en als niet iedereen in de groep een gelijke bijdrage levert, duidelijk te maken bij wie studenten tijdig aan de bel kunnen trekken;
- Maak gebruik van een '[peer evaluation](#)' systeem, waarbij studenten elkaars prestaties beoordelen.

Waarom?

Je wilt er zeker van kunnen zijn dat het resultaat van een toets werkelijk het resultaat van die individuele student is, of bij samenwerkingsopdrachten, het resultaat van een groep studenten. Het informeren van studenten over fraude en het nemen van maatregelen om fraude op te sporen, draagt bij aan de betrouwbaarheid en transparantie van toetsing.

Meer weten?

- In de regels en richtlijnen van examencommissies staan richtlijnen over hoe te handelen in geval van fraude.
- Informatie over het voorkomen en herkennen van fraude:
<http://hum.leidenuniv.nl/medewerkers/onderwijs/onderwijsbeleid/plagiaat-docenten.html>
https://www.utwente.nl/ces/vop/onderwijs_links/toetsing_en_assesment_fraude_e/index.html#voorkomen_van_fraude_en_plagiaat
- Informatie over plagiaat: <http://hum.leidenuniv.nl/studenten/reglementen/plagiaat.html>

5. Evalueren en verbeteren van toetsen

Wat is een toetsevaluatie?

Een toetsevaluatie is een evaluatie van de kwaliteit van de toets. Tijdens het nakijken en beoordelen van toetsen merk je al welke vragen slecht werden gemaakt, of op welk beoordelingscriterium studenten slecht scoren. In feite ben je dan al bezig de toets (en het onderwijs) te evalueren. Hieronder staan wat meer structurele manieren om een toets te evalueren.

Hoe evalueer en verbeter ik mijn toetsen?

Er zijn verschillende manieren om een toets te evalueren en verbeteren:

- Maak een lijst van (door studenten) veelgemaakte fouten om je onderwijs aan te passen, nieuwe toetsvragen te bedenken, nieuwe beoordelingscriteria op te stellen of bestaande criteria aan te scherpen. Zodoende laat je studenten steeds meer werken aan zaken waar ze blijkbaar moeite mee hebben, dingen die ze moeten leren dus.
- Vraag studenten feedback op de toets, bijvoorbeeld door een korte (online) vragenlijst of door onderaan de toets(opdracht) een paar vragen te formuleren. Vraag studenten bijvoorbeeld welke vragen ze duidelijk of onduidelijk vonden, of de toets een goede afspiegeling was van het onderwijs, en vraag om tips ter verbetering.
- *Peer assessment* kun je ook als evaluatiemiddel inzetten: als studenten elkaar beoordelen met het beoordelingsmodel dat jij hebt opgesteld, test je de duidelijkheid en bruikbaarheid van het beoordelingsmodel. Zo kun je samen met studenten het beoordelingsmodel bijstellen, waardoor studenten 'mede-eigenaar' van de beoordeling worden.
- Naar aanleiding van de toetsresultaten van jouw hele studentpopulatie kun je bepalen wat het uiteindelijke *slagingspercentage* is en wat de *gemiddelde score van de 5% beste studenten*. Van de 5% beste studenten mag verwacht worden dat zij vrijwel alle vragen goed hebben kunnen beantwoorden. Als dit niet zo is, dan was de toets wellicht te moeilijk. In dat geval kun je besluiten de cesuur aan te passen aan het resultaat van de beste 5% (zie ook hoofdstuk 3.2)

Bij de verschillende toetsvragen (of toetsonderdelen) kun je letten op:

- Onderdelen die goed worden gemaakt zijn misschien te gemakkelijk. Je kunt dan overwegen die onderdelen volgend jaar minder zwaar te laten wegen, of op een hoger niveau te toetsen. Onderdelen die slecht worden gemaakt zijn misschien te moeilijk of onvoldoende uitgelegd tijdens het onderwijs. Het kan natuurlijk ook dat je die moeilijke onderdelen expres hebt toegevoegd om de beste studenten extra uitdaging te geven.
- Het onderscheidende vermogen van een vraag of onderdeel: als studenten met hoge cijfers op de toets slecht scoren op een bepaalde vraag of onderdeel in de toets, dan is er misschien iets mis met dat onderdeel.

Voor toetsen met open en gesloten vragen kan het ICLON een psychometrische toetsanalyse uitvoeren om de kwaliteit van de toets en de afzonderlijke toetsvragen te analyseren¹¹.

Meer weten?

- Over toetsanalyses: <http://media.leidenuniv.nl/legacy/toetsing-en-toetsanalyse.pdf>
- Over psychometrische toetsanalyse: Bijlage IV

¹¹ Zie Bijlage IV voor uitleg over psychometrische toetsanalyse

6. Praktijkvoorbeelden Faculteit Geesteswetenschappen

Voorbeeld van leerdoelen

De student die de cursus succesvol heeft afgerond is in staat om:

- een omschrijving van de epistemologie als wijsgerige en wetenschappelijke discipline te geven en te bediscussiëren;
- een toelichting van bovengenoemde kernbegrippen in de epistemologie te geven en te bediscussiëren;
- een omschrijving van bovengenoemde debatten in de epistemologie te geven;
- een omschrijving en beoordeling van standpunten en stromingen in het kader van deze debatten te geven;
- een eigen standpunt met betrekking tot deze debatten te kiezen en te motiveren;
- een omschrijving van bovengenoemde problemen in de hedendaagse epistemologie te geven en suggesties voor het oplossen daarvan te schetsen en te beoordelen.

Voorbeelden van toetsvragen

- Voorbeelden van open toetsvragen

VRAAG 10 (bij afbeelding kunstwerk)

10a. Noem de voor- en achternaam van de kunstenaar, het tijdvak en de stroming.

10b. Leg de naam van de stroming uit aan de hand van drie belangrijke kenmerken van dit kunstwerk.

10c. Verklaar de genoemde kenmerken in 10b op basis van de culturele context waarin deze stroming is ontstaan.

10d. Leg aan de hand van dit werk kort de problematiek met betrekking tot het definiëren van de termen 'abstract' en 'realistisch' uit, zoals besproken in het college met betrekking tot deze stroming.

10e. Kies drie benaderingen uit de mogelijkheden die Charles Harrison noemt om kunstwerken zoals dit object te interpreteren en vat kort samen wat die benaderingen inhouden.

- Voorbeeld van open toetsvraag met modelantwoord

VRAAG

Naar aanleiding van de lezing van Hans Bloemsma, *Carl Friedrich von Rumohr en de vroege Italianen*; Wat is de reden dat Carl Friedrich von Rumohr (1785-1843) de kunst van Cimabue en Duccio hoger waardeerde dan die van Giotto?

MODELANTWOORD

(Hans Bloemsma) Terwijl de kunst van Cimabue en Duccio volgens Rumohr gericht was op de

verheven verbeelding van heilige voorstellingen, was dit bij Giotto van ondergeschikt belang geworden. Daarentegen had hij zich voornamelijk gericht op de wereldse aspecten van zijn voorstellingen, waarbij hij de nadruk legde op het burleske en het pathetische. Giotto's voorstellingen voldeden volgens Rumohr daarmee niet aan het (Romantische) ideaal van de christelijke kunst, zoals verwoord door Friedrich Schlegel: "Het is de bestemming van de kunst om de religie te verheerlijken en haar geheimen nog mooier en duidelijker te maken." Met andere woorden, Giotto's voorstellingen zijn te aards, te weinig verheven en hebben een groot tekort aan religieus gevoel [vanzelfsprekend is het niet nodig om Schlegel te citeren voor het juiste antwoord!]

Voorbeeld van opdracht voor schriftelijk werkstuk

- Schrijfopdracht

In dit college wordt Tony Bennetts concept van het museum als 'exhibitionary complex' onder de loep genomen. Zijn invloedrijke artikel 'The Exhibitionary Complex' uit 1988 wordt gepositioneerd, becommentarieerd en geactualiseerd.

Lit.: Tony Bennett, 'The Exhibitionary Complex'. Als artikel eerst verschenen in *new formations*, Number 4, Spring 1988. Hele tekst is via internet te downloaden:

www.columbia.edu/itc/anthropology/.../bennett.pdf

Ook te vinden in: Greenberg, Reesa, Bruce W. Ferguson & Sandy Nairne (eds.), *Thinking about Exhibitions*, Routledge, 1996, pp.81-112.

OPDRACHT

Maak een kritische analyse van het werk van Tony Bennett aan de hand van een museum naar keuze.

Voorbeelden van beoordelingsformulieren

- Voorbeelden van beoordelingsformulieren voor schrijfoopdrachten

Essay Score Form: English Linguistics	
student:	Grade
course:	
graded by:	
linguistic analysis [well-argued, detailed (observation, description), formal (using the appropriate terms and concepts), and precise analysis of linguistic data]	<input type="radio"/> excellent <input type="radio"/> good <input type="radio"/> acceptable <input type="radio"/> insufficient
structure [information presentation, logical build-up of text and argumentation]	<input type="radio"/> excellent <input type="radio"/> good <input type="radio"/> acceptable <input type="radio"/> insufficient
use of scholarly sources [informed, critical and creative use of relevant secondary material]	<input type="radio"/> excellent <input type="radio"/> good <input type="radio"/> acceptable <input type="radio"/> insufficient
language [correct and formal academic prose]	<input type="radio"/> excellent <input type="radio"/> good <input type="radio"/> acceptable <input type="radio"/> insufficient
presentation [use of stylesheet and layout]	<input type="radio"/> excellent <input type="radio"/> good <input type="radio"/> acceptable <input type="radio"/> insufficient
Comments	

Frequently Made Mistakes

Your tutor can use this list to indicate what common mistakes, if any, occur frequently in your essay.

linguistic analysis

insufficient empirical evidence for analytical claims	
overview of linguistic data and/or existing literature without offering a critical evaluation or analysis	
inaccurate or imprecise use of linguistic terminology	
imprecise and overgeneralized claims	

use of scholarly sources

secondary sources are of insufficient academic quality (<i>Wikipedia</i> is an especially egregious example)	
insufficient references or empirical evidence to support claims	
use of secondary sources not effectively integrated into the argument	

structure

the essay lacks a clear, precise thesis statement	
body paragraphs lack a clear topic sentence	
cohesion between sentences and/or paragraphs is unclear	

language

grammatical mistakes (e.g. dangling infinitives and participles: 'Having processed the questionnaire, the results were presented in a table'; subject/verb disagreement: 'the verbs in this sentence is ...'; incorrect verb tenses: 'In <i>The Canterbury Tales</i> , Chaucer is not using <i>do</i> -support'; etc.)	
spelling and punctuation mistakes (e.g. it's/its, belief/believe, there/their, traveled/travelled, etc.; comma splice, incorrect punctuation of relative clauses, etc.)	

presentation

inconsistent bibliographical references (no italics for titles of books and journals, no publisher or year of publication, no page range for articles, etc.)	
layout problems (no double-spacing, no tabs at beginning of paragraph, no page numbers, etc.)	

Essay Assessor

assignment: TITLE OF THE COURSE

personalia	
mark	
TOPIC, QUESTION, RESULT <ul style="list-style-type: none">· <i>relevance</i>· <i>validity</i>· <i>significance</i>· <i>originality</i>	
ARGUMENT <ul style="list-style-type: none">· <i>cogency</i>· <i>clarity & structure</i>· <i>originality</i>· <i>theoretical orientation</i>· <i>methodology</i>	
RESEARCH <ul style="list-style-type: none">· <i>sophistication</i>· <i>extent of the effort</i>	
LANGUAGE USAGE <ul style="list-style-type: none">· <i>correctness</i>· <i>style</i>	
PRESENTATION <ul style="list-style-type: none">· <i>scholarly (citation, bibliography etc)</i>· <i>general (formatting & neatness)</i>	
OTHER	

- Voorbeeld van beoordelingsformulier inzet BA-werkcollege

Beoordelingsformulier Bachelor Werkcollege (BA-WC) Universiteit Leiden – Opleiding Geschiedenis

Instructies voor de invulling van het formulier:

- het ingevulde formulier moeten worden ingeleverd bij het secretariaat.
- cijfers worden niet ingevoerd in USIS indien formulieren ontbreken.
- de student krijgt een ingevulde versie van het formulier.
- op het formulier moeten bij Schriftelijk werkstuk de beide velden Toelichting worden ingevuld.
- de leerdoelen uit de E-Gids kunnen door de docent worden toegevoegd aan het laatste vak.
- docenten moeten zelf wegingsfactoren toekennen aan onderdelen. Die wegingsfactoren moeten gelijk zijn aan de wegingsfactoren zoals die zijn aangeleverd bij de vakomschrijvingen in de e-gids. Wegingsfactoren moeten zo zijn gekozen dat er alleen een voldoende kan worden gehaald wanneer het werkstuk voldoende is.
- er mag als wegingsfactor 0 worden ingevuld.

BA werkcollege		naam docent	
Datum		handtekening docent	
Naam student		Studentnummer	
Eindcijfer			

Presentatie voor de collega-studenten (structuur, inhoud, vorm)	Participatie (aanwezigheid, deelname aan discussie, geven kritiek, samenwerking)	Opdracht(en) (planning, omgaan met kritiek, samenwerking, bibliografie)
<input type="checkbox"/> Uitstekend <input type="checkbox"/> Goed <input type="checkbox"/> Voldoende <input type="checkbox"/> Matig <input type="checkbox"/> Onvoldoende Cijfer: Weging:	<input type="checkbox"/> Uitstekend <input type="checkbox"/> Goed <input type="checkbox"/> Voldoende <input type="checkbox"/> Matig <input type="checkbox"/> Onvoldoende Cijfer: Weging:	<input type="checkbox"/> Uitstekend <input type="checkbox"/> Goed <input type="checkbox"/> Voldoende <input type="checkbox"/> Matig <input type="checkbox"/> Onvoldoende Cijfer: Weging:
Evt. toelichting:		
Schriftelijk werkstuk (inhoudelijke/analytische aspecten) - Probleemstelling & deelvragen, inbedding in historiografie, verantwoording van methode, gebruik theorie en/of hypothesen, analyse & interpretatie van onderzoeksgegevens. - Opbouw en consistentie van betoog, conclusie, kritische analyse, originaliteit.		weging:
Toelichting:		Oordeel
Schriftelijk werkstuk (vorm- en technische aspecten) - <i>Technische aspecten</i> : heuristiek, ordening van materiaal, achtergrondkennis en historisch besef. - <i>Vormaspecten</i> : taalbeheersing en stijl, alinea- en hoofdstukindeling, annotatie, illustraties, tabellen & grafieken.		weging:
Toelichting:		Oordeel

- Voorbeeld van beoordelingsformulier BA-scriptie

Beoordelingsformulier bachelor-Eindwerkstuk Universiteit Leiden –
 Faculteit der Geesteswetenschappen – Opleiding Geschiedenis

Opleiding	Geschiedenis	Afstudeerrichting	
Datum		EC	15
Naam student		Studentnummer	
Titel Eindwerkstuk			
Eerste beoordelaar		Tweede beoordelaar	
<p>1. Is het eindwerkstuk naar uw oordeel vrij van plagiaat? <input type="radio"/> Ja <input type="radio"/> Nee</p> <p>2. BA-Eindwerkstukken in het repository zijn in beginsel vertrouwelijk tenzij student en docent goede redenen zien om het werkstuk toch openbaar te laten zijn. Dan dit hierna aangeven. <input type="radio"/> Dit Eindwerkstuk mag via het repository direct openbaar gemaakt worden. <input type="radio"/> Dit werkstuk mag pas vanaf __-__-20__ via het repository openbaar gemaakt worden.</p>			
Criteria	Oordeel		
Kennis en inzicht (inhoud, relatie tot vakgebied)	Facultatief: weging: %		
Toelichting:	<input type="radio"/> zeer goed <input type="radio"/> goed <input type="radio"/> voldoende <input type="radio"/> onvoldoende		
Toepassen kennis en inzicht (methodologie)	Facultatief: weging: %		
Toelichting:			
Oordeelsvorming (interpretatie, argumentatie, conclusie)	Facultatief: weging: %		
Toelichting:			
Communicatie (schrijfvaardigheid, structuur)	Facultatief: weging: %		
Toelichting:			
Leervaardigheden (proces)	Facultatief: weging: %		
Toelichting:			
Formele vereisten			
Ca. 35 pagina's, max. 15.000 woorden (incl. voetnoten, excl. literatuurlijst)			
Toelichting:			
Opleidingsspecifieke criteria Geschiedenis			
1. Helder geformuleerde historische vraagstelling 2. Inbedding in historiografische en/of wetenschappelijke discussie 3. Verantwoording van de onderzoeksmethode			

4. Gebaseerd op kritisch onderzoek in primaire bronnen 5. Voldoende brede literatuurstudie 6. Goed gestructureerd betoog 7. Wetenschappelijke onderbouwde conclusie 8. Correct taalgebruik NB Alleen Eindwerkstukken die aan deze 8 criteria voldoen kunnen met een voldoende worden beoordeeld Toelichting:		
Samenvattend oordeel/toelichting		
Handtekening eerste beoordelaar	Handtekening tweede beoordelaar	Eindoordeel
NB Medewerkers OSZ		
Is het eindoordeel een 6- 6,5, voer het cijfer dan niet in naar USIS totdat er een handtekening is van een derde lezer. Deze wordt op verzoek van eerste lezer aangewezen door de ambtelijk secretaris van de Examencommissie.		
Naam derde lezer	Handtekening derde lezer	Oordeel derde lezer

Subcriteria beoordeling **bachelor-Eindwerkstuk** (aan te vullen met opleidingsspecifieke eisen gerelateerd aan de opleidingsspecifieke eindtermen)

Kennis en inzicht (inhoud, relatie tot vakgebied) Dublin-descriptor: Heeft aantoonbare kennis en inzicht van een vakgebied, waarbij wordt voortgebouwd op het niveau bereikt in het voortgezet onderwijs en dit wordt overtroffen; functioneert doorgaans opeen niveau waarop met ondersteuning van gespecialiseerde handboeken, enige aspecten voorkomen waarvoor kennis van de laatste ontwikkelingen in het vakgebied vereist is.
<i>O.a.</i> <ul style="list-style-type: none"> · <i>de onderzoeksvraag is gebaseerd op een probleemstelling die getuigt van inzicht in de centrale debatten en methodes van het vakgebied;</i> · <i>helderheid, relevantie en afbakening van de probleemstelling;</i> · <i>inkadering in bestaande literatuur;</i>
Toepassen kennis en inzicht (methodologie) Dublin-descriptor: Is in staat om zijn/haar kennis en inzicht op dusdanige wijze toe te passen, dat dit een professionele benadering van zijn/haar werk of beroep laat zien, en beschikt verder over competenties voor het opstellen en verdiepen van argumentaties en voor het oplossen van problemen op het vakgebied.
<i>O.a.</i> <ul style="list-style-type: none"> · <i>kritische analyse van het materiaal en van bronnen(kwaliteit van de analyse);</i> · <i>operationalisering en gebruik van begrippen;</i> · <i>gebruik adequate onderzoeksmethoden;</i> · <i>beschrijving en verantwoording gevolgde werkwijze.</i>
Oordeelsvorming (interpretatie, argumentatie, conclusie) Dublin-descriptor: Is in staat om relevante gegevens te verzamelen en interpreteren (meestal op het vakgebied) met het doel een oordeel te vormen dat mede gebaseerd is op het afwegen van relevante sociaalmaatschappelijke, wetenschappelijke of

ethische aspecten.
<p><i>O.a.</i></p> <ul style="list-style-type: none"> · <i>logisch en consistent betoog; conclusies volgen logisch en onderbouwd uit het gepresenteerde materiaal;</i> · <i>de mate waarin de onderzoeksvraag daadwerkelijk wordt beantwoord;</i> · <i>Indien van toepassing: maatschappelijke en ethische aspecten meegewogen in oordeelsvorming.</i>
<p>Communicatie (schrijfvaardigheid, structuur) Dublin-descriptor: Is in staat om informatie, ideeën en oplossingen over te brengen op publiek bestaande uit specialisten of niet-specialisten.</p>
<p><i>O.a.</i></p> <ul style="list-style-type: none"> · <i>taalgebruik (voer- c.q. doeltaal opleiding: niveau van taalvaardigheid, leesbaarheid, stijl, spelling, grammatica, gebruik en uitleg juiste vaktermen);</i> · <i>opbouw en opmaak van de scriptie (hoofdstuk- en paragraafindeling, inhoudsopgave; gebruik illustraties);</i> · <i>apparaat met inbegrip van annotatie (correct gebruik verwijzingsstelsel, volledigheid verwijzingen, literatuurlijst, etc.).</i>
<p>Leervaardigheden (proces) Dublin-descriptor: Bezit de leervaardigheden die noodzakelijk zijn om een vervolgstudie die een hoog niveau van autonomie veronderstelt aan te gaan.</p>
<p><i>O.a.</i></p> <ul style="list-style-type: none"> · <i>mate van zelfstandigheid;</i> · <i>planning en tijdsmanagement;</i> · <i>omgang met feedback begeleiders;</i> · <i>Indien van toepassing: Participatie in scriptiegroep..</i>
<p>Formele vereisten</p>
<ul style="list-style-type: none"> · <i>Ca. 35 pagina's, max. 15.000 woorden (incl. voetnoten, excl. literatuurlijst)</i>

Geraadpleegde bronnen

Literatuur

Berkel, H.J.M. van, A. Bax & D. Joosten-ten Brinke (red.) (2014). *Toetsen in het hoger onderwijs*. Houten: Bohn Stafleu van Loghum.

Biggs, J.B. (2003). *Teaching for quality learning at university*. Buckingham: Open University Press/Society for Research into Higher Education. (Second edition)

Breukers, H. et al. (2011). *Handboek toetsen en tentamens bij de Open Universiteit*. Heerlen: Open Universiteit.

Bruijns, V. en M. Kok (2013). *Leidraad toetsen en beoordelen*. Amsterdam: Hogeschool van Amsterdam.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15(3), 309-333. doi: 10.1207/S15324818AME1503_5

Kallenberg T., L. van der Grijspaarde, A. ter Braak & G. Baars (2014). *Leren (en) doceren in het hoger onderwijs*. Den Haag: Boom Lemma.

Universiteit van Amsterdam (2012). *Kader Toetsbeleid*. Amsterdam: UvA. (Universitaire Beleidsnotities).

Websites

Universiteit Utrecht, Faculty of Humanities (geraadpleegd juli 2015). *Toetsing - Modules*.
▶ <http://toetsing.hum.uu.nl/modules/>

Hogeschool van Amsterdam (geraadpleegd juli 2015). *Score – Toetsen en beoordelen*.
▶ <https://score.hva.nl/>

Bijlage I Opstellen van leerdoelen en passende toetsvormen

Leerdoelen bevatten het beoogde niveau qua kennis, vaardigheden en/of gedrag na afloop van een bijeenkomst, cursus of opleiding. Een leerdoel bestaat idealiter uit vier componenten:

1. Inhoud: welke inhoud moet de student zich eigen maken;
2. Gedrag: wat moeten studenten met de leerstof kunnen doen;
3. Voorwaarden: onder welke condities, met welke hulpmiddelen of bronnen moet de student het gedrag vertonen;
4. Norm: hoeveel moet een student doen, binnen welke tijd (vooral van belang bij het construeren van toetsen).

Specifieke en meetbare leerdoelen zijn SMART, dat wil zeggen:

- Specifiek: Het leerdoel is in termen van concreet gedrag beschreven
- Meetbaar: Dat gedrag is te beoordelen aan de hand van vooraf te stellen criteria
- Acceptabel: Het gedrag is voor studenten en docenten relevant en motiverend
- Realistisch: Het doel is haalbaar, en sluit aan op de zone van naaste ontwikkeling van de student
- Tijdgebonden: Het doel is binnen de gestelde tijd te bereiken

Tip voor het formuleren van leerdoelen:

- Neem geen leeractiviteit op in de leerdoelen (zoals ‘door groepswerk moeten studenten ertoe komen om...’), beschrijf alleen het eindresultaat
- Beschrijf alleen het hoogste beheersingsniveau (‘toepassen’ impliceert ‘kennen’), dus “studenten hanteren verschillende methoden voor X” en NIET “studenten kennen verschillende methoden voor X en passen deze toe”
- Gebruik bij voorkeur eenduidige actiewerkwoorden. Kies hiervoor actiewerkwoorden die aansluiten bij de activiteit die je van studenten verwacht (zie het schema hieronder)

Beheersings-niveau	Gedragskenmerk	Voorbeelden van actiewerkwoorden voor het formuleren van leerdoelen	Voorbeelden van passende toetsvormen
Onthouden	feiten-reproductie,	classificeren, herkennen, identificeren, in volgorde	Schriftelijk: gesloten (multiple choice) vragen en open

	herkenning, herinnering	plaatsen, lokaliseren, navertellen, noemen, onderkennen, opsommen, reproduceren, rubriceren, selecteren, uit elkaar houden, weergeven <i>(Bijvoorbeeld: "Herkent belangrijke gebeurtenissen in de geschiedenis van de VS")</i>	vragen
Begrijpen	begrip, interpretatie, logische reproductie	aanduiden, aangeven, formuleren, illustreren, karakteriseren, opdracht geven, schetsen, signaleren, typeren, met eigen woorden vertellen, vertegenwoordigen <i>(Bijvoorbeeld: "Parafraseert een gedeelte uit een gegeven stuk tekst")</i>	Schriftelijk: gesloten en open vragen, essayvragen, papers Mondeling: presentatie, ondervraging
Toepassen	elementen uit 'onthouden' en 'begrijpen' hanteren in nieuwe situaties; kiezen van de juiste wetten, regels, schema's, begrippen	aandeel leveren, behandelen, beschrijven, bewaken, bijdragen, definiëren, demonstreren, gebruiken, hanteren, oplossen, opstellen, een overzicht geven, procedure kiezen en volgen, rapporteren, uitleggen, voorspellen, vormgeven, vragen formuleren <i>(Bijvoorbeeld "Hanteert verschillende strategieën om relevante vakwetenschappelijke literatuur te selecteren")</i>	Schriftelijk: gesloten en open vragen, essayvragen, papers, werkstukken Mondeling: presentatie, assessment
Analyseren, synthetiseren	ordenen naar inhoud, vorm, functie e.d.; samenstellen van elementen tot een uniek en origineel	afkeuren, afleiden, afwegingen maken, alternatieven voorleggen, argumenteren, becommentariëren, belangen afwegen, concluderen, construeren, controleren, discussiëren, herformuleren, leiding geven, modelleren, motiveren, onderhandelen, ontwerpen, fouten opsporen, opbouwen, problemen oplossen,	Schriftelijk: gesloten en open vragen, essayvragen, papers, werkstukken, groepsopdracht Mondeling: presentatie, assessment, debat

	geheel	<p>ordenen, organiseren, overleggen, prioriteiten stellen, relateren, samenstellen, samenvatten, uitvoeren, tot stand brengen</p> <p><i>(Bijvoorbeeld "Bepaalt het standpunt van de schrijver van een essay wat betreft zijn of haar politieke visie")</i></p>	
Evalueren, creëren	<p>evalueren, beoordelen, toepassen buiten eigen discipline</p>	<p>adviseren, beoordelen, commentaar geven, kritisch doorlichten, evalueren, ondersteunen, oordelen, verdedigen, toetsen, zelfstandig optreden</p> <p><i>(Bijvoorbeeld: "Formuleert een hypothese om een waargenomen fenomeen te verklaren")</i></p>	<p>Schriftelijk: essay, paper, literatuurstudie, werkstuk, project, onderzoeksvoorstel</p> <p>Mondeling: advies, debat, presentatie</p>

Bijlage II Maken van een toetsmatrijs

1. Maak een tabel met in de kolommen: leerdoelen, kennisvragen, inzichtvragen, toepassingsvragen¹², punten (en eventueel toetsvormen)
2. Bepaal het gewicht van elk leerdoel (sommige leerdoelen zijn belangrijker dan andere)
3. Geef bij elk leerdoel aan hoeveel vragen je erover wilt stellen en hoeveel punten elke vraag waard is
4. Tel het aantal punten per niveau op en controleer of de verhouding overeenkomt met de wat je voor ogen had (misschien staan er in verhouding teveel kennisvragen in en te weinig toepassingsvragen)

Voorbeelden van soorten toetsmatrijzen:

Leerdoelen	Kennisvragen	Inzichtvragen	Toepassingsvragen	Punten
Leerdoel A	2 x 5 pt			10 pt
Leerdoel B		1 x 10 pt	1 x 15 pt	25 pt
Leerdoel C	2 x 5 pt	1 x 10 pt		20 pt
Leerdoel D		1 x 10 pt	1 x 15 pt	25 pt
Leerdoel E			1 x 20 pt	20 pt
Totaal	20 pt	30 pt	50 pt	100 pt

Inhoud	Cognities niveau waarop wordt getoetst						Praktische organisatie			
	Kennis	Inzicht	Toepassen	Analyse	Synthese	Evaluatie	Soort vraag	Aantal	Gewicht binnen de toets	Wie maakt de vraag?

¹² Je kunt ook voor andere niveaus kiezen, bijvoorbeeld: reproductie, begrip, toepassen, probleem oplossen, motorische vaardigheid, interactieve vaardigheid etc.

Bijlage III Checklist voor het samenstellen, beoordelen en redigeren van open vragen

Bron: CITO - Het construeren van open vragen

http://www.toetswijzer.nl/html/toetsenopschool/hoofdstuk7_herzien.pdf

Checklist voor het samenstellen, beoordelen en redigeren van open vragen

Taalgebruik

- Is de vraag grammaticaal juist geformuleerd?
- Bevat de vraag ingewikkelde zinsconstructies?
- Bevat de vraag een dubbele ontkenning?
- Bevat de vraag onnodige tussenvoegsels?
- Is de vraag onnodig negatief gesteld?
- Kan de formulering van de vraag aanleiding geven tot misverstanden?
- Kan de vraag door een verschuiving van de klemtoon een andere betekenis krijgen?

Informatie

- Bevat de vraag voldoende informatie voor beantwoording?
- Geeft de vraag voldoende informatie over de gewenste lengte en vorm van het antwoord?
- Is duidelijk dat een antwoord moet worden verklaard/toegelicht?
- Zijn informatie en probleemstelling duidelijk gescheiden?

Relevantie

- Kan de vraag met gebruik van andere dan de beoogde kennis worden beantwoord?
- Suggereert de vraag een probleem dat niet aan de orde is?
- Bevat de vraag onbedoelde hints voor de beantwoording?
- Is de moeilijkheidsgraad van de vraag respectievelijk van de gehele toets acceptabel?
- Wordt de moeilijkheidsgraad van de vraag onnodig verhoogd door irrelevante gegevens?

Context

- Is het gebruik van context (tekeningen, grafieken, teksten, afbeeldingen) functioneel?
- Is de context duidelijk en correct weergegeven?
- Bevat de context irrelevante informatie?

Presentatie

- Zijn de vragen en de vraagonderdelen duidelijk van elkaar te onderscheiden?
- Is de nummering van de vragen logisch en overzichtelijk?
- Zijn de conventies wat betreft spelling, symboolgebruik, interpunctie, e.d. in acht genomen?
- Zijn de verwijzingen in de vraag naar teksten, tekeningen, e.d. juist?

Bijlage IV Uitleg over psychometrische toetsanalyse

Met een psychometrische toetsanalyse kun je het volgende meten:

- De *betrouwbaarheid* van de toets, gemeten met Cronbach's alpha. Alpha ligt tussen 0 en 1. Bij 0 is er totaal geen verband tussen de scores op de verschillende toetsvragen. Je kunt je dan afvragen of de toets een consistent geheel is. Bij 1 is er een maximaal verband tussen de verschillende toetsvragen. In dat geval kun je je afvragen of de toets in feite niet slechts een vraag toetst. Het scoreverloop van elke toetsvraag is dan namelijk precies hetzelfde. In de praktijk is alpha echter nooit 0 of 1, maar ligt deze ergens tussen de 0,5 tot 0,9. Een waarde van rond de 0,8 wordt gezien als goed (Van Berkel, Bax & Joosten-ten Brinke, 2014).
- De *moeilijkheid* van elke toetsvraag, gemeten met de p-waarde. Deze ligt tussen 0 (alle studenten hadden 0 punten op de vraag) en 1 (alle studenten hadden het maximaal aantal punten op de vraag). Met andere woorden: hoe hoger p, hoe gemakkelijker de vraag.
- Het *onderscheidend vermogen* van de vraag, gemeten met de item-restcorrelatie (Rir). Dit is een correlatie tussen de scores op de vraag en de scores op alle andere vragen. Bij een positieve correlatie scoren studenten die de betreffende vraag goed hebben gemaakt ook goed op de andere vragen en scoren studenten die de vraag slecht hebben gemaakt ook slecht op de andere vragen. De vraag onderscheidt dus goed tussen goede en slechte studenten. Is de correlatie echter negatief, dan scoren studenten die de vraag goed hebben gemaakt juist slecht op de overige vragen en scoren studenten die de vraag slecht hebben gemaakt juist goed op de overige vragen. In dat geval is het onderscheidend vermogen van de vraag blijkbaar niet goed en kan er wat mis zijn met de vraag. Wat er mis is, dat moet jij als docent beoordelen. Het kan bijvoorbeeld dat een vraag voor meerdere interpretaties vatbaar is, of de antwoordsleutel van een gesloten vraag kan niet kloppen.
- Bij meerkeuzevragen kun je ook nagaan of veel studenten een bepaalde afleider hebben gekozen. Zo ja, dan kun je nagaan of dit alternatief ook als correct beschouwd kan worden. Is het antwoord hierop ook 'ja', dan kun je naast het goede antwoord ook deze 'aantrekkelijke afleider' goed rekenen.
- Vaak kiezen studenten bij meerkeuzevragen een of meerdere afleiders niet tot nauwelijks. Je kunt je dan afvragen of dit plausible afleiders zijn en of je in plaats van (bijvoorbeeld) een vierkeuzevraag niet feitelijk een tweekeuzevraag zou moeten gebruiken. Hoe minder plausible afleiders er zijn, hoe gemakkelijker het voor de student wordt de vraag correct te beantwoorden. Als een afleider niet plausibel blijkt te zijn, probeer dan een betere afleider te bedenken. Dit kan door tijdens het onderwijs goed op te letten welke fouten studenten vaak maken. Deze fouten kun je gebruiken als afleiders in meerkeuzevragen.