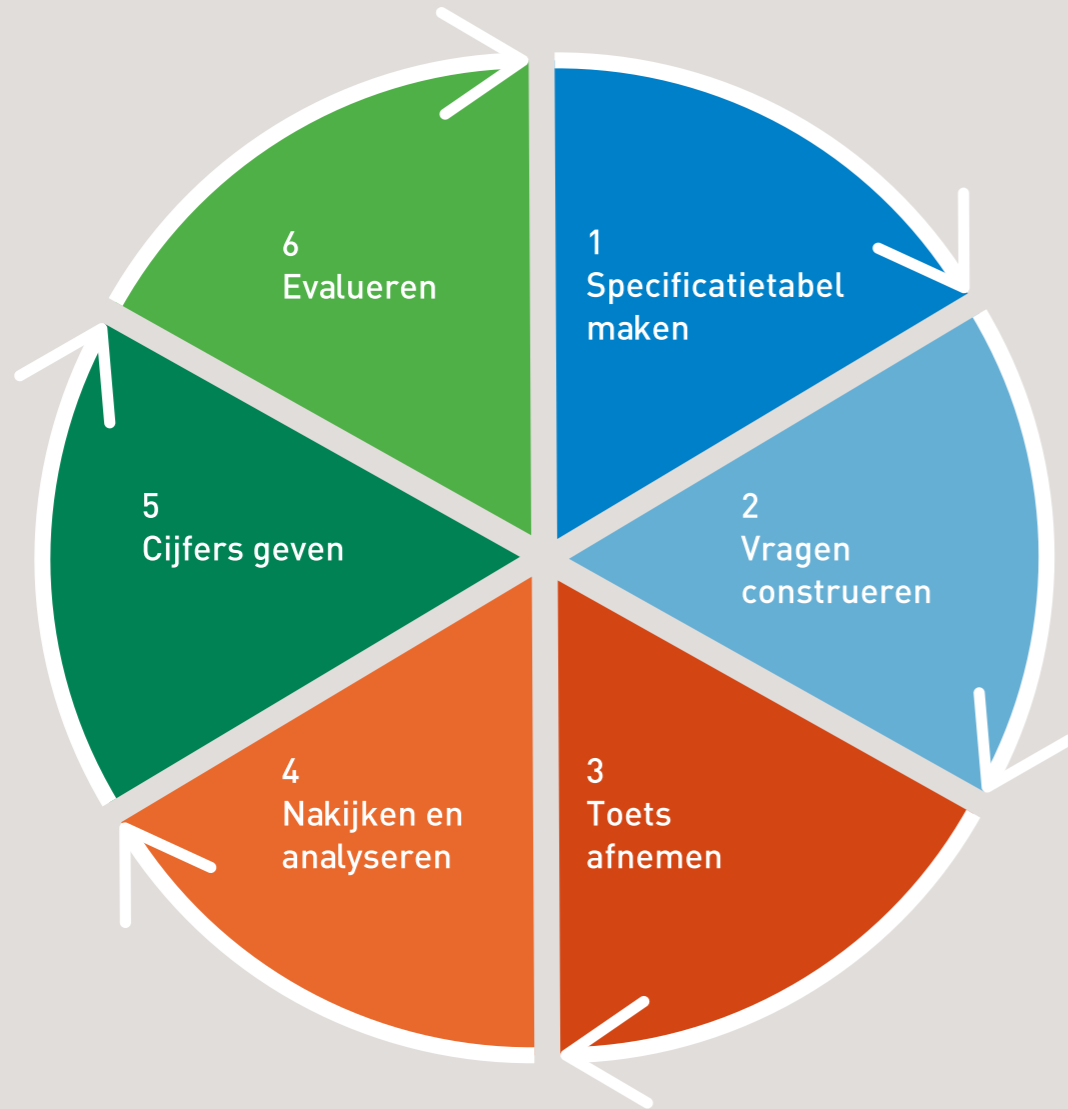


# Quickstart Toetsontwikkeling

Ontwerpen en analyseren van toetsen en toetsvragen



# Quickstart Toetsontwikkeling



## Waar gaat deze bijdrage over?

Steeds blijkt weer dat docenten en examinatoren (in het hoger onderwijs vallen deze twee functies meestal samen) behoefte hebben aan informatie die hen helpt bij het construeren van toetsen.

Niet te veel theoretische achtergrondinformatie, maar concrete informatie 'hoe ze het moeten doen'. In die behoefte probeert deze bijdrage te voldoen. Het is een soort checklist met punten waar toetsconstructeurs aan moeten denken wanneer ze een toets gaan construeren. De theorie hierachter is zoveel mogelijk weggelaten.

Met nadruk moet worden gesteld dat de inhoud van deze bijdrage primair betrekking heeft op tentamens of tussentoetsen, niet meer en niet minder.

Het is niet onwaarschijnlijk dat de inhoud ook relevant kan zijn voor andere toets(situaties), maar die waren niet het uitgangspunt van de auteurs.

De bijdrage is chronologisch opgebouwd. De paragrafen volgen de verschillende stappen in het toetsconstructieproces zoals die zich ook in werkelijkheid zullen voordoen.

© Auteurs (v1-sept-2022)

Henk van Berkel, Universiteit Maastricht

Silvester Draaijer, Vrije Universiteit Amsterdam



# 1 | Maak een toetsmatrijs, kies een vraagtype en toetslengte

Een toetsmatrijs (specificatietabel) bestaat enerzijds uit een aantal belangrijke onderwerpen uit een cursus waar je vragen over wil stellen en anderzijds uit het cognitieve niveau waarop je wil richten met de vragen.

Onderwerpen	Weten – Reproductie Toetsvragen op het niveau van het reproduceren van kennis, doen een beroep op het geheugen (feitenkennis). Feiten dienen herkend te worden of feiten dienen in een verband te worden gebracht.	Toepassing Bij toepassingsvragen worden studenten voor een voor hen nieuw probleem gesteld. Het probleem wordt opgelost door verschillende kenniselementen met elkaar in verband te brengen of met een formule, wet of oplossingsprincipe.	Inzicht Inzichtvragen vereisen van studenten een hoge mate van reorganisatie van het probleem en zelfstandig denken. De vragen doen een beroep op intellectuele vaardigheden als interpretatie, analyse, synthese en evaluatie.
1			
2			
3			
Enz.			

## 1.1 Onderwerp en aantal vragen

- ☑ Vermeld de belangrijkste onderwerpen uit de cursus, meestal voldoet een aantal van rond de vijftien.
- ☑ Het aantal vragen is van belang voor de representativiteit (inhoudsvaliditeit) van de toets. Een juiste verdeling van de vragen over de cellen in de specificatietabel zal er voor zorgen dat de toets representatief is.
- ☑ Het aantal vragen is ook van belang voor de betrouwbaarheid van de toets (hoe meer vragen, hoe beter).
- ☑ Hoe belangrijker je het onderwerp vindt, des te meer vragen je er over moet stellen.
- ☑ Vermeld in de cellen het aantal vragen dat je wilt stellen over het betreffend onderwerp op het betreffend niveau.
- ☑ Zorg er voor dat over ieder onderwerp tenminste één vraag wordt gesteld.

Hoeveel vragen gesteld kunnen worden is afhankelijk van de vorm en de toetstijd.

## 1.2 Open vraag

Soort vraag	Antwoordtijd	Voorbeeldvraag
In- en aanvulvraag	½ minuut	Er wordt volgende week besloten of een van de meest ..... [haten] belastingen verdwijnt.  Wat is het onderdeel van het binnenoor waarin zich het orgaan van Corti bevindt? Dat is de/het .....
Kort antwoordvraag	één minuut	Wat deed Pilatus volgens het Nieuwe Testament met zijn handen nadat hij Christus ter dood had veroordeeld?
Lang antwoordvraag	Tien minuten	Wat waren de langetermijneffecten (noem er 3) van de Kruistochten? Beredeneer je antwoord.
Opstelvraag / Betoogvraag	25 minuten	Beredeneer de ontwikkelingen over de periode van tien jaar van de werkloosheid, als in één klap alle lonen worden gehalveerd. Ga uit van de situatie dat er verder geen andere wijzigingen in de conjunctuur optreden.
Berekeningsvraag	Afhankelijk van de complexiteit van het vraagstuk	

## 1.3 Gesloten vraag

Soort vraag	Antwoordtijd	Aantal vragen normaal gesproken nodig voor een betrouwbare toets
Juist/Onjuistvraag of twee alternatieven	50 seconden	80
Drie alternatieven	60 seconden	60
Vier alternatieven	75 seconden	40



Bedenk:

- Het is een hardnekkig misverstand dat met gesloten vragen uitsluitend feitenkennis is te bevragen.
- Het is een hardnekkig misverstand dat vragen over feitenkennis automatisch resulteren in hoge scores en vragen naar toepassing of inzicht in lage scores.
- Lang open toetsvragen en opstelvragen kosten je erg veel tijd om na te kijken. Het is bovendien bewezen dat bij het nakijken van open vragen allerlei dingen misgaan die ten koste gaan van de betrouwbaarheid (en dus de eerlijkheid) van de beoordeling.
- Gesloten vragen zijn doelmatiger dan open vragen als meer dan 50 studenten aan de toets deelnemen.
- Lang open toetsvragen kosten veel tijd om te beantwoorden waardoor er minder verschillende onderwerpen in de toetstijd bevroegd kunnen worden. Dit kan snel ten koste gaan van de representativiteit van de leerdoelen in de toets.



## 2 | Vragen construeren en controleren

### Inspiratiebronnen:

Inspiratiebronnen voor het construeren van vragen zijn:

- De eindtermen de problemen/taken in studieboeken, syllabi e.d.
- Hoofdpunten uit de literatuur en colleges.
- Het toekomstig beroep.
- De vaardigheden en practica tijdens de cursus.
- Discussies in de colleges, werkcolleges e.d..
- Vragen van reeds afgenomen toetsen (met name vragen die niet te moeilijk of makkelijk zijn en een goed onderscheid maken tussen de studenten die de stof wel en die de stof niet beheersen).
- Grafieken, tabellen, schema's en andere afbeeldingen van relevante kenmerken of processen kunnen dienen om meerdere vragen bij te ontwikkelen.

Als je de gangbare indeling, weten/reproductie – toepassing – inzicht wilt volgen, gebruik dan bijvoorbeeld bepaalde **werkwoorden** om vragen te ontwikkelen.

<b>Weten-Reproductie</b>	Benoemen, Beschrijven, Citeren, Definiëren, Extrapoleren, Identificeren, Interpretieren, Onderscheiden, Opsommen, Parafaseren, Samenvatten, Schatten, Selecteren, Uitleggen, Vertalen, Vertellen in eigen woorden
<b>Toepassing</b>	Berekenen, Demonstreren, Gebruiken, Maken, Ontwikkelen, Oplossen, Organiseren, Produceren, Relateren, Transfereren, Veranderen, Voorbereiden, Wijzigen
<b>Inzicht</b>	Bekritisieren, Categoriseren, Componeren, Concluderen, Contrasteren, Deduceren, Formuleren, Herschrijven, Illustreren, Interpretieren, Maken, Onderscheiden, Ondersteunen, Ontrafelen, Ontwerpen, Rechtvaardigen, Relateren, Samenvatten, Schetsen, Toelichten, Valideren, Verdedigen, Vergelijken, Verklaren, Waarderen

Ook onderstaande **startzinnen** geven goed houvast bij het formuleren van vragen, vooral vragen die kritisch denken of probleemoplossen beogen te meten.

<b>Weten:</b> Nagaan of de student kennis heeft van onderwerpen	Wat is de beste definitie voor ....? Wat is (niet) karakteristiek voor ....? Uit welke onderdelen bestaat het probleem? Wat is de geschiedenis van het probleem? Welke verschillende categorieën zijn er in het probleem?
<b>Kritisch denken (evalueren):</b> Nagaan of de student de eigenschappen van feiten, procedures, principes of theorieën kan gebruiken.	Wat is het meest effectief (gepast) voor ....? Wat is beter (slechter) ....? Wat is het meest effectief voor ....? Wat is de meest kritieke stap in aan procedure? Als je weet dat X waar is, wat is dan tevens waar over Y? Wat is (niet) nodig in een procedure? Wat is het belang van het probleem ?
<b>Kritisch denken (voorspellen):</b> Nagaan of de student op basis van feiten, procedures, principes of theorieën kan aangeven wat gevolgen, consequenties e.d. zijn.	Wat zou er gebeuren als ....? Als dit gebeurt, wat zou je doen? Op basis van wat ...., wat zou je doen? Gegeven ... wat is belangrijkste reden dat ....
<b>Probleem oplossen (gegeven een scenario):</b> Nagaan of de student op basis van gegeven problemen oplossingen kan geven of oplossingen kan beoordelen.	Wat is de aard van het probleem? Wat heb je nodig om dit probleem op te lossen? Wat is een mogelijke oplossing? Wat is de meeste effectieve (efficiënte) oplossing? Waarom is .... de meest effectieve (efficiënte) oplossing?



#### Tips:

- Construeer minimaal 25% meer vragen dan dat je uiteindelijk in de toets wil opnemen. Je zal altijd vragen na de eerste ontwikkelronde niet goed achten.
- Denk vooruit, als je tevens al een hertentamen moet ontwikkelen of toetsen voor opeenvolgende jaren: ontwikkel direct meerdere varianten vragen van hetzelfde type.
- Maak gebruik van een vragenbankprogramma om je vragen in op te slaan.

## 2.1 Na de constructiefase

- ☑ Wijd minimaal één bespreking met anderen aan de conceptvragen (vier-ogen principe) waarin de inhoud, het antwoordmodel c.q. het correcte alternatief en de vorm ter discussie staan.

## 2.2 Antwoordmodellen

Een antwoordmodel is een hulpmiddel voor de nakijker van open vragen dat dient om de betrouwbaarheid van het nakijken te verhogen. Het bevat de volgende punten:

- ☑ Opsomming in trefwoorden van wenselijke antwoorden, en hun puntentoekening
- ☑ Procedure wanneer het antwoord niet voorkomt in de genoemde trefwoordenlijst
- ☑ Wat te doen met taalfouten?
- ☑ Bij lange antwoorden: inhoud, technische vormgeving en argumentatie en een puntentoekening op onderdelen.

## 2.3 Checklists

Hanteer de onderstaande checklist nadat je de vragen hebt geconstrueerd.

### Inhoud

- ☑ Is de gekozen toetsvorm (open of gesloten) de meest passende?
- ☑ Bestaat de toets uit een voldoende aantal vragen?
- ☑ Bevat de vraag slechts één duidelijk probleem?
- ☑ Bevat de vraag geen subjectieve uitspraken?
- ☑ Staat bij een citaat in de vraag de context?
- ☑ Bevat de vraag geen overbodige informatie?
- ☑ Heeft de vraag niet onnodig betrekking op een detail?
- ☑ Is de vraag geen strikvraag?
- ☑ Is bij een stellingvraag de stelling wel 100% juist of 100% onjuist?
- ☑ Bevat de vraag voldoende informatie voor beantwoording?

### Vorm

- ☑ Bevat de vraag geen vage aanduidingen?
- ☑ Is de vraag niet onnodig negatief geformuleerd?
- ☑ Is de vraag grammaticaal juist geformuleerd?
- ☑ Bevat de vraag geen ingewikkelde zinsconstructies?
- ☑ Kan de formulering van de vraag geen aanleiding geven tot misverstanden?
- ☑ Zijn ontkenningen als niet onderstreept of gecursiveerd?
- ☑ Bevat de vraag geen dubbele ontkenningen?
- ☑ Bevat de vraag geen woorden als altijd, nooit, meestal, zeker?
- ☑ Zijn de conventies wat betreft spelling, symboolgebruik, interpunctie, en dergelijke in acht genomen?
- ☑ Is de vraag opsplitsbaar in een gegeven en een vraaggedeelte?

### Proces

- ☑ Is er een toetsmatrijs gemaakt?
- ☑ Heeft er een bespreking met vakgenoten plaatsgevonden?

### Open vragen

- ☑ Geeft de vraag voldoende informatie over de gewenste lengte en vorm van het antwoord?
- ☑ Is duidelijk dat een antwoord verklaard/toegelicht moet worden?
- ☑ Is er een antwoordmodel gemaakt?

### Gesloten vragen

- ☑ Komen er geen overlappingen in de alternatieven voor (antwoorden die subsets van elkaar zijn veroorzaken veel verwarring)?
- ☑ Zijn alle alternatieven ongeveer even lang?
- ☑ Zijn de alternatieven in oplopende/alfabetische volgorde geplaatst?
- ☑ Zijn alle alternatieven plausibel?
- ☑ Is de vorm 'geen van bovenstaande' of 'alle bovenstaande' echt nodig?





### 3 | De toets samenstellen en afnemen

#### 3.1 Samenstellen

Een toets is meer dan een losse verzameling losse toetsvragen.

- ☑ Stel uit de goedgekeurde vragen de toets samen met behulp van de toetsmatrijs
- ☑ Cluster de vragen per onderwerp in de volgorde zoals die in het onderwijs of stof aan de orde zijn gekomen.

#### 3.2 De afname

De toetsinstructie voor studenten dient informatie te bevatten over de volgende onderwerpen:

- ☑ Tijdsduur
- ☑ Gesloten vragen: invulinstructie (bijvoorbeeld: het streepje zetten of rondje aanvinken)

- ☑ Gewicht van de vragen (bij open vragen kan het gewicht verschillend zijn per vraag; bij gesloten vragen telt iedere vraag in beginsel even zwaar)
- ☑ Wijze van de zak/slaaggrensbepaling
- ☑ Bekendmaking van de uitslag: termijn en plaats
- ☑ Raadplegen van hulpbronnen: welke zijn wel/niet toegestaan?
- ☑ Verbod op het gebruik van mobiele telefoons en andere apparatuur
- ☑ Bij toetsen op papier: het aantal bladzijden en het aantal vragen in de toets; studenten zijn zelf verantwoordelijk voor eventuele misdrukken
- ☑ Verwijzing naar het fraudereglement
- ☑ Wat te doen bij vragen, bijvoorbeeld hand opsteken
- ☑ Regeling toiletbezoek
- ☑ Inleverinstructie
- ☑ Mogelijkheid tot het maken van opmerkingen



#### Bedenk:

Een toets met zowel meerkeuzevragen als open vragen kan voordelen hebben. Meerkeuzevragen zorgen dat in relatief weinig tijd studenten over een breed scala van onderwerpen kunnen worden bevraagd die tevens automatisch worden gescoord door het toetssysteem. De open vragen lenen zich goed voor het stellen van verdiepende vragen en voorkomen het meer oppervlakkig studeren dat sneller optreedt bij toetsen met uitsluitend meerkeuzevragen.

### 4 | De toets analyseren en nakijken

#### Open vragen

1. Gebruik het vooraf gemaakte antwoordmodel.
2. Kijk de toets 'dwars' na. Dat wil zeggen per vraag en niet per student. Maak gebruik van de anonimiseerfunctie als dat mogelijk is.
3. Verander zo nu en dan de nakijkvolgorde en gebruik zo mogelijk de anonimiseerfunctie van het nakijkprogramma.
4. In de praktijk zal het vaak niet mogelijk zijn om alle vragen door twee beoordelaars na te laten kijken, maar kijk in ieder geval de toets van studenten die rond de zak/slaaggrens scoren nog een keer na. Middel even tueel beide beoordelingen.
5. Rond de scores pas op het einde af.
6. Ook van een open vragentoets kun je een toets- en itemanalyse maken.

7. Verzamel het studentcommentaar, leg dit eventueel naast de toetsanalyse en bespreek dit met de vragenconstructeur.
8. Neem een beslissing over de antwoordmodellen: handhaven, wijzigen of aanvullen
9. Kijk, indien daar aanleiding toe is, de vragen opnieuw na met het vernieuwde antwoordmodel.

#### Gesloten vragen

1. In het digitale toetssysteem kun je direct de analyse zelf uitvoeren. Bij gebruik van afname met papieren antwoordformulieren die gescand kunnen worden zal de tentamenservice de analyse voor je uitvoeren.
2. Verzamel het studentcommentaar en leg dit naast de toets- en itemanalyse en bespreek dat met de vragenconstructeur.
4. Neem een beslissing over de vragen: handhaven, verwijderen of sleutelwijziging. Besef dat deze verandering over het algemeen slechts kleine wijzigingen in de uitslag opleveren.
5. Laat opnieuw een analyse uitvoeren.

## 5 | Cijfers geven

Er zijn vele methoden van cijfergeven die onder drie categorieën vallen: absoluut, relatief en de compromis methode.

Absoluut	Bij de absolute methode bepaal je vooraf de voldoende/onvoldoende grens (cesuurscore). Die grens wordt ingegeven door het uitgangspunt dat gecontroleerd moet worden of studenten voldoen aan de eisen die kunnen worden afgeleid uit de doelstellingen van het onderwijs.
Relatief	De relatieve methode gaat uit van het idee dat de toets afgestemd moet zijn op studenten die gerechtigd zijn tot het volgen van onderwijs – en in principe moeten zij in meerderheid in staat geacht worden om te slagen voor een toets. Omdat vooraf niet bekend is waartoe studenten bij het gegeven onderwijs en de gegeven toets in staat zijn, kan de norm niet vooraf worden vastgesteld. Het resultaat op de toets moet worden afgewacht om de norm te kunnen vaststellen.
Compromis	De compromismethode probeert de principiële verschillen tussen absoluut en relatief normeren te overbruggen en heeft in de onderwijspraktijk de voorkeur. Bij de compromismethode wordt in het algemeen uitgegaan van een absolute norm en wordt aangegeven onder welke omstandigheden van deze absolute norm zal worden afgeweken. Een van de compromis-methoden staat hieronder beschreven.

### Open vragen

1. Begin met het bepalen van de grens voldoende/onvoldoende. Het is niet ongebruikelijk die te leggen bij 55% of 60% van de maximaal te behalen score.
2. Indien je vindt dat er teveel studenten zakken voor je toets, zijn (na overleg met de examencommissie) de volgende aanpassingen in de procedure mogelijk.
  - a. In plaats van 60% stel je de grens op 55% van de maximale score en/of
  - b. Je neemt niet de maximale score als uitgangspunt, maar het gemiddeld van de vijf hoogste scores.
3. Je hebt nu de ruwe score die een 0 oplevert en de ruwe score die een 10 oplevert: de overige cijfers bepaal je door de ruwe scores op te delen in 10 gelijke eenheden met daaraan gekoppeld een cijfer.

### Gesloten vragen

1. Begin met het bepalen van de grens voldoende/onvoldoende. Die leg je in principe bij 60% van de hoogste score en rekening houdend met de kansscore.

- De kansscore is het aantal geldige vragen in de toets gedeeld door het aantal alternatieven per vraag. Bijvoorbeeld de kansscore met 105 juist/onjuistvragen is  $105/2=52,5$ .
  - Als bij de genoemde toets de hoogste score 96% is, komt de voldoende/onvoldoende grens te liggen bij 60% van de score tussen  $52,5$  en  $96 = 78,6$ .
2. Studenten met de hoogste score (96) krijgen een 10, zij met de kansscore een 0. Met behulp van een lineaire regressie zijn de overige cijfers te berekenen.
  3. Indien je vindt dat er teveel studenten zakken voor je toets: zijn na overleg met de examencommissie de volgende aanpassingen in de procedure mogelijk:
    - a. In plaats van 60% stel je de grens op 55% van de maximale score en/of
    - b. Je neemt niet de maximale score als uitgangspunt, maar het gemiddeld van de vijf hoogste scores.



"Een faire toets biedt studenten die voldoende hebben gestudeerd een goede kans om te slagen"





## 6 | De toets evalueren

In het online rapport staan kwaliteitsindices die betrekking hebben op de toets als geheel en op de vragen afzonderlijk. Een bestudering van rapport levert informatie op om het volgend jaar in de toets verbeteringen aan te brengen:

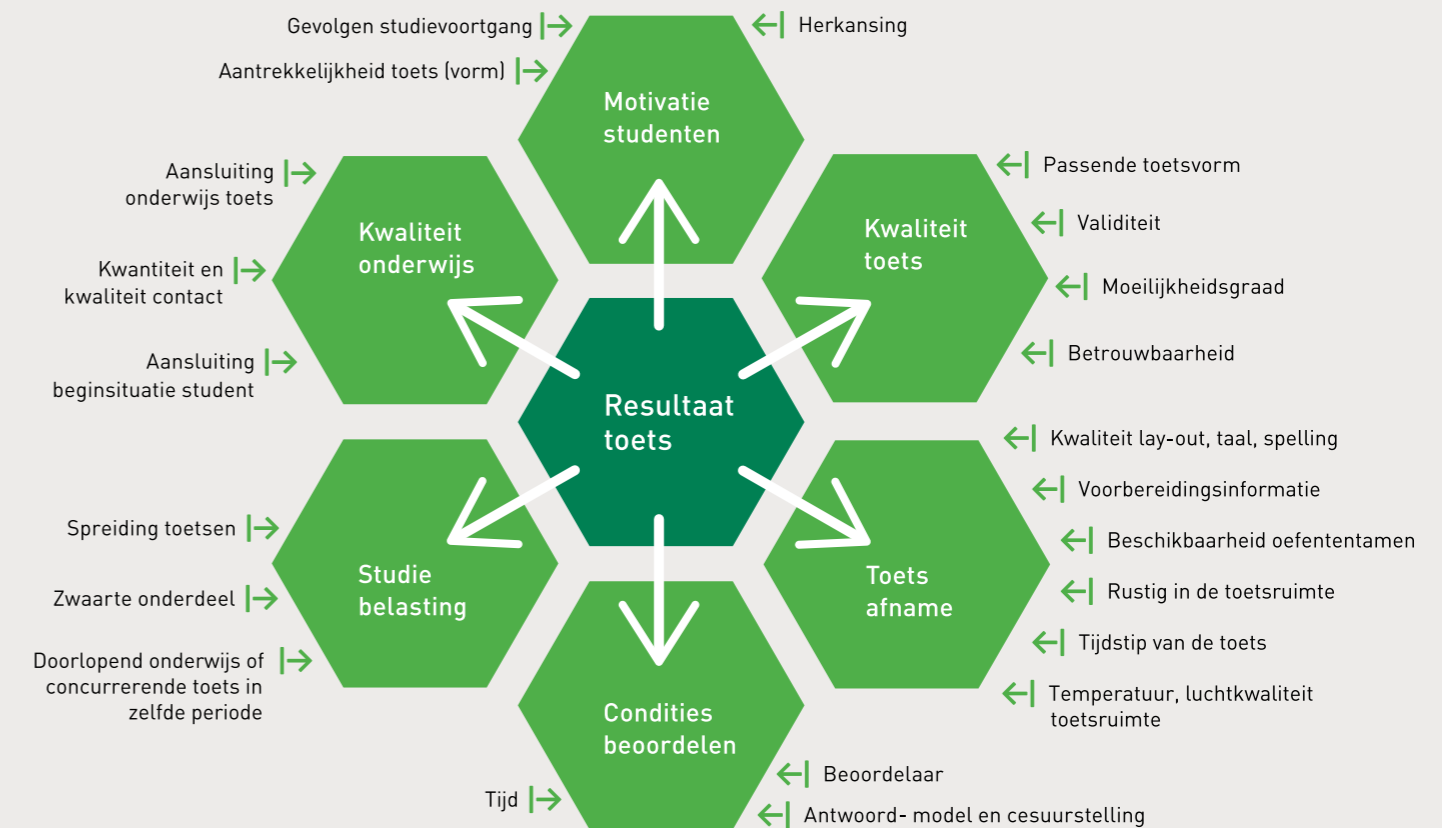
1. Kijk naar de betrouwbaarheid van de toets. Is deze lager dan 0,70, dan is aan te raden volgend jaar meer vragen in de toets op te nemen.

2. Bestudeer de zeer moeilijke en makkelijke vragen. Zoek oorzaken hiervoor zodat u de volgende keer dergelijke vragen kunt vermijden.
3. Bestudeer de psychometrisch zeer goede vragen (dat zijn vragen die een goed onderscheid maken tussen competente en niet competente studenten). Dat blijkt uit de zogenaamde Rit-waarde van de vraag. Dergelijke vragen zijn het waard om ook in de volgende toets op te nemen.

Het interpreteren van een toets- en itemanalyse vergt kennis en inzicht. Neem contact met de toetsservice van de VU als je aanvullend advies wenst.

Bedenk dat het resultaat op een toets door vele aspecten wordt beïnvloed.

Bestudeer onderstaande figuur waarin een aantal factoren genoemd worden die scores en betrouwbaarheid van de toets kunnen beïnvloeden in zowel positieve als negatieve richting.





Gemaakt in opdracht van VU Network for  
Teaching and Learning en SOZ Procesregie  
[vu.nl/ntl](https://vu.nl/ntl)



Meer informatie?  
[vu.nl/nl/medewerker/onderwijs-geven-aan-de-vu](https://vu.nl/nl/medewerker/onderwijs-geven-aan-de-vu)



Leer hoe het toetsproces op de VU is vorm-  
gegeven en wat verantwoordelijkheden zijn  
van verschillende verantwoordelijken:  
[vu.nl/nl/medewerker/toetsen-en-beoordelen](https://vu.nl/nl/medewerker/toetsen-en-beoordelen)



Leer hoe de VU toetsapplicatie TestVision werkt:  
[canvas.vu.nl/courses/31787](https://canvas.vu.nl/courses/31787)



Wil je meer tips over toetsing? Ga naar de  
didactische tip pagina's:  
[vu.nl/nl/medewerker/didactiek/tips-activerend-onderwijs-met-blended-learning](https://vu.nl/nl/medewerker/didactiek/tips-activerend-onderwijs-met-blended-learning)