

Toets- en itemanalyse

[Inleiding](#)

[Theorie](#)

[Drie niveau's van kwaliteitsaspecten](#)

[Kenmerken van betrouwbaarheid](#)

[Checklist](#)

[Doen](#)

[Bronnen](#)

Inleiding

De meeste toetsen in het hoger onderwijs hebben een kwalificatie functie: De docent bepaalt op basis van de toetsresultaten welke studenten de lesstof in voldoende mate beheersen en welke niet. Het is essentieel dat de toetsuitslag een objectieve en rechtvaardige uitspraak doet over het niveau van kennis en vaardigheden van de student.

Theorie

Drie niveau's van kwaliteitsaspecten

Van Berkel en Bax (2006)¹ beschrijven drie niveaus waarop de kwaliteitsaspecten moeten worden geadresseerd:

- **De toetsvraag.** Bij de constructie van toetsvragen moet worden voldaan aan inhoudelijke en vormtechnische eisen (Teelen,2004). Dit wordt ondervangen door een controle van de toetsvragen voorafgaand aan de toets.
- **De toets** dient valide en betrouwbaar te zijn. Een toets is valide als de vragen relevant en representatief zijn en als de toets de student in staat stelt om te laten zien welke kennis en vaardigheden hij of zij beheerst. Een toets is betrouwbaar als het toetsresultaat de daadwerkelijke kennis en vaardigheid van de student weerspiegelt.Het gebruik van een goede toetsmatrijs biedt hiertoe houvast bij het samenstellen van een toets.
- **De toetsresultaten.** Na afloop van de toets wordt een psychometrische analyse uitgevoerd. Hierbij wordt aan de hand van de indicatoren moeilijkheidsgraad (P-waarde), onderscheidingsvermogen (Rit -waarde) en betrouwbaarheid (Cronbachs α) gesignaleerd worden welke vragen nader onderzocht moeten worden.

Kenmerken van betrouwbaarheid

Kenmerken die de betrouwbaarheid van een toets bepalen

- **Objectiviteit:** mate waarin toetsuitslag onafhankelijk is van de beoordeling/beoordelaar of de toetsomstandigheden.
- **Specificiteit:** Alleen studenten die de leerstof beheersen kunnen de vraag beantwoorden.
- **Efficiëntie:** De toets ziet er verzorgd en overzichtelijk uit (lay-out) en de studenten krijgen voldoende tijd om de toets te maken.
- **Transparantie:** Het is voor de studenten duidelijk welke lesstof wordt getoetst. In een toetsinstructie bij de toets wordt aangegeven uit hoeveel vragen de toets bestaat, hoeveel tijd zij voor het maken van de toets krijgen en hoe de score van de antwoorden wordt gedaan, bijvoorbeeld of in de score wordt gecorrigeerd voor de raadkans

¹ Berkel, H.J.M. van & Bax, A.E (2006) Toetsen in Hoger Onderwijs. Bohn Stafleu Van Lochem

Checklist

In onderstaande tabel zijn de verschillende indicatoren weergegeven, die bij een psychometrische analyse worden berekend. De tabel geeft de range van de waarden aan en de actie die naar aanleiding van het berekende resultaat dient te worden ondernomen.

Indicator	Range	Resultaat	Reparatie
P-waarde	van 0 tot 1	Kleiner of gelijk aan de raadkans	keurt het toetssysteem het juiste antwoord goed? Nee, vraag/antwoord aanpassen Ja, vraag verwijderen uit toets Vervolgens de analyse opnieuw uitvoeren
		Hoger dan de raadkans, maar beduidend lager dan de wenselijke waarde	Indien Rit >0: vraag handhaven Indien Rit <0: vraag verwijderen en analyse opnieuw uitvoeren
		Is (bijna) gelijk aan 1	Vraag handhaven
Rit –waarde	van -1 tot 1	Kleiner of gelijk aan 0	keurt het toetssysteem het juiste antwoord goed? Nee, vraag/antwoord aanpassen Ja, vraag verwijderen uit toets Vervolgens de analyse opnieuw uitvoeren
Cronbachs α	van 0 tot 1	Lager dan 0,70*	Toetsuitslag uitsluitend als formatief beschouwen
		Tussen 0,70 en 0,80	Indien de toets kan worden gecompenseerd met andere toetsen, de uitkomst is niet verontrustend

* In de praktijk wordt een waarde vanaf 0,60 als valide betrouwbare toets beschouwd, omdat er soms doelbewust moeilijke vragen worden opgenomen om een goede en excellente studenten te kunnen onderscheiden. Dit heeft een negatieve impact op de waarde van Cronbachs α .

Doen

Bekijk de resultaten van uw itemanalyse en identificeer welke vragen niet voldoen aan de bovengenoemde ranges. Onderzoek waarom de betreffende items niet voldoen.

Bronnen

Dit stuk is integraal overgenomen uit '*Ontwikkelingen in de Delftse Toetspraktijk*', TU Delft , 2011 M.A. Keijzer-de Ruijter