

Kwaliteitseisen

Samenvatting

In dit hoofdstuk worden de kwaliteitseisen aan een toets besproken. Een toets moet valide, betrouwbaar, bruikbaar en transparant zijn.

4.1 Praktijk – 26

4.2 Reflectie – 26

4.3 Kwaliteitseisen voor toetsing – 26

4.3.1 Validiteit van een toets – 27

4.3.2 Betrouwbaarheid van een toets – 28

4.3.3 Bruikbaarheid van een toets – 29

4.3.4 Transparantie van de toets – 30

4.4 Portfolio – 30

4.1 Praktijk

Casus

In de kantine van de opleiding HRM zitten twee studenten, Lois en Lara, te praten over de studentenhandleiding van het vak Belastingrecht. Via Blackboard hebben ze de volgende tekst voor zich op de laptop:

De toetsing van de module vindt plaats op de volgende manier:

Je krijgt een individueel schriftelijk tentamen opgebouwd uit:

- *multiplechoicevragen (30 punten)*
- *casus Loonbelasting (20 punten)*
- *casus Inkomstenbelasting (40 punten)*
- *casus Vennootschapsbelasting (10 punten)*

De studiepunten krijg je als je voor het tentamen ten minste het cijfer 5,5 behaalt.

Datum toets: week 9 van periode 3

Herkansing: week 10 van periode 4

Lois en Lara bespreken de vóórs en tegens van een toets die bestaat zowel uit multiple-choicevragen als casusvragen. Lois, houdt niet zo heel erg van stampen en ziet het wel zitten dat zij met het beantwoorden van de casusvragen al 70 punten kan halen. Lara vindt juist de toepassing in de praktijk nog wel erg lastig. Zij had liever wat meer vragen gehad op basis van hard studeren. Ze besluiten in de voorbereiding van de toets wat meer samen te werken, dan zouden ze misschien allebei de toets kunnen halen.

■ Vragen

- Wat vind jij van het combineren van multiplechoicevragen en casusvragen in één toets?
- Welke voor- en nadelen van deze combinatie toetsen zie jij bij jouw opleiding?

4.2 Reflectie

Beantwoord de vragen in [tab. 4.1](#) met ja of nee om een indruk te krijgen van de kwaliteit van jouw toetsen.

4.3 Kwaliteitseisen voor toetsing

Vier belangrijke criteria voor de kwaliteit van toetsen zijn:

- validiteit
- betrouwbaarheid
- bruikbaarheid
- transparantie

Tabel 4.1 Test de kwaliteit van jouw toetsen.

Wat doe jij om de kwaliteit van de toets te verhogen?		ja	nee
1.	Ik maak toetsen waarbij de vragen de hele leerstof omvatten		
2.	Ik maak toetsen waarbij de verschillende onderwerpen evenwichtig over de toets zijn verdeeld		
3.	Ik maak toetsen waarbij de moeilijke vragen alleen goed worden beantwoord door de goede studenten		
4.	Mijn toetsen bevatten voldoende vragen om uit te sluiten dat het resultaat een toevalstreffer is		
5.	Het resultaat van de door mij beoordeelde toetsen is niet afhankelijk van mijn persoonlijkheid of humeur		
6.	De lay-out van de door mij gemaakte toetsen is overzichtelijk		
7.	In de toets is het voor de studenten duidelijk hoeveel punten ze per vraag kunnen verdienen		
8.	Ik communiceer over de methode van toetsing, de beoordeling en de deadlines vóór de start van de onderwijseenheid		
9.	Het is de studenten voorafgaand aan de lessen duidelijk aan welke eisen ze aan het eind van de onderwijseenheid moeten voldoen		
10.	Tijdens de lessen besteed ik aandacht aan de inhoud en wijze van toetsen om de studenten goed voor te bereiden		

4.3.1 Validiteit van een toets

➤ **Een valide toets meet wat de toets moet meten. De toets is afgestemd op het doel.**

Hierbij kan onderscheid worden gemaakt tussen begripsvaliditeit en inhoudsvaliditeit.

➤ **Bij begripsvaliditeit gaat het om de mate waarin de resultaten beschouwd worden als een maat voor het beoogde theoretische begrip of de bedoelde vaardigheid.**

De competenties die studenten moeten behalen, worden vertaald in termen van concreet gedrag. Dit gedrag wordt dan weer vastgelegd in de beoordelingscriteria. Dat is niet altijd gemakkelijk. Wil je een bepaalde kwaliteit zoals 'zelfverzekerdheid' of 'empathie' meten, dan moeten de vragen of opdrachten ook inderdaad deze kwaliteit meten.

➤ **Bij inhoudsvaliditeit gaat het om representativiteit van de toets. De verschillende onderwerpen komen in een goede verhouding terug in de toets.**

Veel toetsen gaan over leerstof met een breed domein. Een toets is representatief als deze alle aspecten van dit domein meet. Dit geldt ook voor toetsen in de vorm van een beroepsproduct of een vaardigheidstoets.

Voorbeeld

De basisvaardigheden voor medisch rekenen bestaan uit de volgende onderdelen: rekenvaardigheid, vaste en vloeibare medicatie, verdunningen en oplossingen en overig (infusie, transfusie, gassen en vochtbalans). Aan alle vier onderdelen worden twee lessen besteed. De toets bestaat totaal uit 20 vragen, waarvan 10 vragen betrekking hebben op rekenvaardigheid, 8 vragen betrekking hebben op vaste en vloeibare medicatie en 2 vragen betrekking hebben op infusie en transfusie. In dit geval is de toets niet representatief, omdat de vier onderdelen van medisch rekenen niet representatief verdeeld zijn. De inhoudsvaliditeit van dit meetinstrument is laag.

■ Toetsmatrijs

Een toetsmatrijs ondersteunt het borgen van de inhoudsvaliditeit van een toets. De samenstelling van de toetsvragen is uitgebalanceerd: er is een onderscheid tussen gemakkelijke en meer complexe vragen. De toetsmatrijs geeft weer welk type vragen en welke complexiteit in de toets opgenomen zijn (zie ► H. 7).

4.3.2 Betrouwbaarheid van een toets

➤ Een toets wordt betrouwbaar genoemd als de toets bij herhaalde afname onder dezelfde omstandigheden, steeds hetzelfde resultaat laat zien.

Een betrouwbare toets is:

- **Specifiek.** De opgaven zijn zo gesteld dat alleen studenten die de inhoud beheersen, ze goed kunnen beantwoorden of uitvoeren. De toets wordt alleen voldoende gemaakt door die studenten, die de leerstof of de vaardigheid echt beheersen.
- **Selectief** (differentiërend). De toets in zijn geheel moet onderscheid maken tussen de 'goede' en 'zwakke' studenten. Dit geldt ook voor de afzonderlijke opgaven. De moeilijke opgaven moeten vooral door de goede studenten goed gemaakt worden.
- **Van voldoende omvang.** De betrouwbaarheid van een toets wordt beïnvloed door het aantal opgaven dat een toets bevat. Bij een toets met weinig vragen is de invloed van elke afzonderlijke opgave veel groter dan bij een toets met veel opgaven. De lengte van de toets moet altijd afgestemd zijn op de beschikbare tijd voor de toets. De omvang van de toets moet voldoende zijn om uit te sluiten dat het resultaat een toevalstreffer is.
- **Objectief.** De vragen en opdrachten moeten zo eenduidig zijn dat gelijkwaardige antwoorden een gelijkwaardige beoordeling krijgen. De toets is consistent, de beoordeling is niet afhankelijk van de beoordelaar. De objectiviteit kan worden vergroot door gebruik te maken van een beoordelingsmodel of rubric (zie ► H. 9).

Voorbeeld

Meten is het vaststellen van een waarde voor een variabele. Stel: je wilt iemands lengte meten en je gebruikt daarvoor een weegschaal. Die toets is niet valide: het instrument meet niet wat je wilt meten.

Als je de lengte wilt meten met een elastisch meetlint, is de toets wel valide maar niet betrouwbaar; de uitkomsten zullen immers niet steeds hetzelfde resultaat laten zien.

4.3.3 Bruikbaarheid van een toets

➤ **De bruikbaarheid van een toets gaat over de praktische uitvoerbaarheid ervan met betrekking tot de locatie, de mensen en de tijd.**

Voorbeeld

Een opleiding heeft een assessment ontwikkeld als afsluiting van een complete studiefase. Qua validiteit en betrouwbaarheid is aan alles gedacht. Bij de eerste afname blijkt dat het assessment wel erg tijdsintensief is, zowel voor de studenten als voor de docenten die assessor zijn. Mede door de grote studentenaantallen en het gebruik van externen als tweede assessor, is de toets ook lastig te organiseren. Ook blijkt dat de beoordelingscriteria zo ambitieus zijn uitgewerkt, dat bijna geen student slaagt. De opleiding besluit de toets flink aan te passen, zodat deze beter bruikbaar wordt.

De locatie voor een toets moet vooraf worden geregeld. Voor sommige toetsen moeten computers op de toetslocatie aanwezig zijn. De locatie moet groot genoeg zijn voor het aantal studenten dat de toets maakt. De tijd en locatie moeten worden doorgegeven aan de studenten en aan degenen die de toets afnemen (docenten of surveillanten). Het aantal vragen of opdrachten sluit aan bij de tijd die voor de toets staat. Voor een multiplechoice-toets van 30 vragen kun je rekenen op 30 minuten. De tijd die de studenten besteden aan de toets moet reëel zijn ten opzichte van de omvang van het te toetsen studieonderdeel.

Voorbeeld

De docenten Recht in een managementopleiding hebben de toets 'Inleiding Recht' voor de eerste periode grondig aangepakt. Ze zijn blij met de toets die representatief is voor de leerstof. Bij afname blijkt dat ze de tijd voor het beantwoorden van de vragen te rooskleurig hebben ingeschat. Geen enkele student is in staat om de toets helemaal af te ronden, zelfs de goede studenten niet. In overleg met de examencommissie wordt een beoordelingswijze afgesproken die recht doet aan de inspanning en getoonde beheersing van de leerstof door de studenten.

Aandachtspunten voor de kwaliteit van een schriftelijke toets:

- De toets is goed leesbaar, zowel de tekst als de afbeeldingen.
- De toets heeft een overzichtelijke lay-out:
 - de vragen zijn duidelijk van elkaar te onderscheiden;
 - de nummering is inzichtelijk;
 - verwijzingen naar een tekst of afbeelding zijn juist;
 - het onderscheid tussen vraag en vraagonderdelen of antwoordmogelijkheden (multiple choice) is duidelijk.
- De teksten zijn goed te begrijpen en afgestemd op het taalniveau van de studenten.

4.3.4 Transparantie van de toets

➤ **Transparantie is de mate waarin de toets heldere en eenduidige informatie geeft aan de student om de toets goed voor te kunnen bereiden en te maken.**

Vooraf, tijdens en na afname van de toets is het de student duidelijk hoe de toetsing plaatsvindt en hoe de toets wordt beoordeeld. Studenten moeten deze informatie goed kunnen vinden, bijvoorbeeld op de digitale leeromgeving. Daarnaast is het noodzakelijk dat deze informatie op tijd (aan het begin van de periode) beschikbaar is voor studenten.

Het is de studenten vooraf duidelijk:

- welke toetsvorm er wordt gebruikt;
- wat de toetsstof, de verplichte en de aanbevolen literatuur is;
- op welke leerdoelen de nadruk ligt;
- op welk niveau er moet worden geleerd;
- hoe de normering is;
- wat de consequenties van een onvoldoende zijn, met daarbij de beschrijving van de herkansingsregeling;
- hoeveel tijd ze voor de toets hebben;
- wat de deadline is waarop het werk ingeleverd moet worden (voor opdrachten);
- welke hulpmiddelen tijdens de toets gebruikt mogen worden, zoals rekenmachine, computer of boek.

4.4 Portfolio

➤ **Leerdoel**
Je kunt jouw toetsen beoordelen in het kader van de verschillende kwaliteitseisen.

1. Geef bij de toets van jouw module/vak aan wat je vindt van de toets op het gebied van de verschillende kwaliteitseisen. Is jouw toets valide, betrouwbaar, bruikbaar en transparant?

2. Bekijk de aandachtspunten om de kwaliteit van toetsen te verhogen. Welke punten zijn in jouw toets goed gerealiseerd en waar is nog winst te behalen?
3. Kijk terug naar jouw antwoorden op de reflectievragen in ► par. 4.2. Wat heb je geleerd door de tekst in dit hoofdstuk en het maken van de opdrachten? Wat kan er nog verbeterd worden met betrekking tot de kwaliteitseisen?