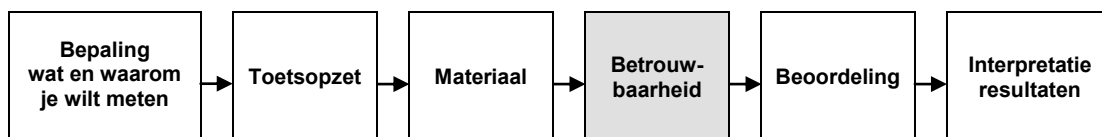


TOETSTIP 6 JANUARI '05



TIP 6: HOE KAN IK DE BETROUWBAARHEID VAN MIJN TOETS CONTROLEREN?

Bij het ontwikkelen, afnemen en beoordelen van toetsen moet aan twee basisprincipes worden voldaan: validiteit en betrouwbaarheid. Wanneer aan deze twee eisen is voldaan, is een toets in principe intern consistent. We frissen eerst nog eens op wat onder deze principes verstaan wordt. Daarna gaan we dieper in op de interne consistentie van toetsen. We reiken u een eenvoudige manier aan om na te gaan of uw toets een interne logica vertoont en wat u kunt doen indien dit niet het geval blijkt te zijn.

1. Wanneer is een toets valide en betrouwbaar?

Een toets is valide wanneer hij meet wat hij beoogt te meten.

Studenten krijgen bij een leestoets de volgende tekst m.b.t de toegangsprijzen voor de zoo van Antwerpen en moeten de bijhorende vragen oplossen.

TOEGANGSPRIJZEN ZOO VAN ANTWERPEN	
Individueel	
- kinderen tot 3 jaar:	gratis
- kinderen van 3 t.e.m. 11 jaar:	€ 9,90
- volwassenen vanaf 12 jaar:	€ 15,00
Groepen algemeen (vanaf 15 personen)	
- kinderen tot 3 jaar	gratis
- kinderen van 3 t.e.m. 11 jaar	€6,50
- volwassenen (vanaf 12 jaar)	€11,00
Vragen:	
- Je wil met je zoon van drie jaar en je dochter van dertien jaar naar de zoo. Hoeveel moet je betalen?	
.....	
- Je gaat met een groep van 25 volwassenen naar de zoo. Er is één meisje van 10 bij. Hoeveel moeten jullie samen betalen?	
.....	

Om de juiste antwoorden op de vragen te vinden, moeten de studenten niet enkel leesvaardig zijn, maar ook goed kunnen rekenen. Daarom kunnen we stellen dat de validiteit in deze toets niet zomaar gegarandeerd is: de toets meet niet wat hij beoogt te meten, namelijk leesvaardigheid, maar meet ook andere dingen. De resultaten op de toets hebben dus niet alleen te maken met leesvaardigheid, maar ook met de rekenvaardigheid van studenten.

Een toets heeft een hoge mate van betrouwbaarheid wanneer bij een hermeting van de toets de cursist dezelfde score zou behalen. Opdat een toets betrouwbaar zou zijn, spelen een aantal aspecten een rol:

- de meetfout moet zo klein mogelijk gehouden worden. Elke toets bevat in zekere mate een meetfout, omdat toevallige factoren die de prestaties van de cursisten beïnvloeden (bv. vermoeidheid bij de cursist, lawaaihinder in de klas, verkeerd begrijpen van de instructies, ...) nooit volledig onder controle gehouden kunnen worden. Door het minimaliseren van deze externe of toevalsfactoren, kan de betrouwbaarheid vergroot worden.
- twee of meerdere beoordelaars moeten los van elkaar een vrijwel zelfde score geven aan de prestaties van de cursist. Dat betekent bijvoorbeeld dat vermoeidheid bij de beoordelaar of voorkennis met betrekking tot het taalvaardigheidsniveau van de cursist geen rol spelen bij het beoordelen van de prestaties.

2. Wanneer is een toets intern consistent?

Een valide en betrouwbare toets vertoont in principe een interne consistentie, een soort van interne logica. Dit betekent dat de moeilijkere items bij voorkeur door de 'betere' studenten worden goed gedaan, en bij voorkeur niet door de 'zwakkere' studenten.

Omgekeerd worden de gemakkelijke items bij voorkeur niet fout gedaan door de 'betere studenten'. De 'betere' studenten in de toets, zijn de studenten met een hogere totaalscore. De 'zwakkere' zijn diegenen met een lagere totaalscore. We gaan er dus vanuit dat er een logisch verband is tussen de totaalscore van de kandidaten en de itemscores afzonderlijk.

Hoe kunnen we nu nagaan of dit verband wel degelijk in de toets zit en of er geen items zijn die dat verband tegenspreken en dus afbreuk doen aan de interne consistentie van de toets?

3. Een analyse van de interne consistentie met de hand

Met de volgende methode kunt u handmatig de interne consistentie van een toets nagaan. U gaat als volgt te werk:

- U maakt een tabel.
- De kolomhoofden zijn de items die in de toets aan bod komen.
- De rijhoofden zijn de namen van de cursisten.
- U noteert de scores van de cursisten op alle items.
- U telt voor elke cursist zijn scores op alle items op. Dit geeft zijn totaalscore op de toets.
- U telt per item de scores van alle kandidaten samen. Dit geeft u een zicht op de moeilijkheid van een item. Een item met een lage totaalscore wordt door weinig cursisten juist opgelost en is dus een moeilijk item.

De tabel kan er dan als volgt uitzien:

Naam	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Totaal
A	1	1	1	0	1	1	1	1	7
B	1	0	0	0	0	1	1	0	3
C	1	1	1	1	1	0**	1	1	7
D	1	1	1	0	0	1	1	0	5
E	1	0	1	0	0	1	1	0	4
F	1	1	1	1	0	1	1	0	6
G	1	0	0	1*	0	0	0	0	2
H	1	1	0	0	0	1	1	0	4
	8	5	5	3	2	6	7	2	

Uw toets vertoont interne consistentie als de volgende principes van toepassing zijn:

1. De items die door kandidaten met een lage totaalscore juist zijn opgelost zijn over het algemeen die items die door veel kandidaten goed worden gedaan. Dit zijn de items met een hoge totaalscore en dus de gemakkelijke items.

2. Kandidaten met een hoge totaalscore hebben in het algemeen geen items fout die door veel kandidaten goed worden gedaan.

3. Moeilijke items - items met een lage totaalscore - worden in het algemeen door kandidaten met een hoge totaalscore juist opgelost.

Uit de scores van de kandidaten op deze toets vallen twee dingen op:

* Kandidaat G heeft een lage totaalscore, nl. 2. Toch heeft hij item 4 goed opgelost. Dit is raar, gezien het feit dat item 4 een moeilijk item is omdat het slechts door 3 van de 8 cursisten juist wordt opgelost. (principe 3 is hier niet geldig)

** Kandidaat C heeft een hoge totaalscore: 7 op 8. Toch heeft hij een gemakkelijk item (item 6) fout opgelost. (principe 2 is hier niet geldig)

Uit deze analyse kunnen we afleiden dat er bij twee items iets vreemds aan de hand is, iets dat ingaat tegen de interne consistentie. Omdat deze items telkens maar bij één student een vreemde score opleverden, zou het onmiddellijk verwijderen van deze items (niet meerekenen in de totaalscore) een te drastische beslissing zijn. Van belang is, naar de specifieke toetsbladen van de studenten te kijken, en trachten na te gaan waarom de student bij het item in kwestie een vreemde score kreeg. Dit kan te wijten zijn aan uiteenlopende factoren op verschillende vlakken die niet noodzakelijk met de kwaliteit van het item zelf te maken hebben:

- een fout bij de beoordeling;
- een fout op het toetsblad: de vraag was niet goed leesbaar op het blad van de kandidaat;
- een probleem bij de kandidaat zelf: concentratieverlies of juist gokken;
- ...

Wanneer bij een zelfde item echter in meer gevallen inbreuken tegen de interne consistentie kunnen worden vastgesteld, kan men ervan uitgaan dat er met het item zelf iets aan de hand is en dat kan leiden tot het verwijderen van het item in de toets. Het spreekt voor zich dat wanneer men meer studenten met de toets toetst, men een beter beeld krijgt van de interne consistentie van de toets en de eventuele inbreuken hierop. Hoe u kunt nagaan of er met het item zelf iets aan de hand is, komt in de volgende ToetsTip aan bod.

OPDRACHT

Voer zelf de analyse van de interne consistentie van een toets uit. U kunt de tabel en de interpretaties naar ons doorsturen voor feedback. Onze contactgegevens staan onder aan deze tip.

TOETSTIP 7 IN MAART '05

In de volgende ToetsTip geven we aan hoe u de interne consistentie van uw toets kunt vergroten. Wat kunt u doen indien u na een analyse merkt dat een item in uw toets herhaaldelijk ingaat tegen de interne consistentie.

Vragen, opmerkingen en suggesties

Opmerkingen over ToetsTip 6:

Suggesties voor volgende ToetsTips:

Inhoudelijke opmerkingen over evaluatie in het algemeen:

Praktische bemerkingen over de ToetsTip:

Voor vragen & reacties één adres:

het CNaVT-secretariaat

Certificaat Nederlands als Vreemde Taal
Blijde-Inkomststraat 7
B-3000 Leuven
België
0032 (0) 16 53 55 16

cnavt@arts.kuleuven.ac.be