

Methodologische verantwoording

Op deze plek wordt uitgelegd hoe de betrouwbaarheid van domeinen van stellingen wordt gemeten. Binnen de testtheorie wordt met item verwezen naar een stelling of vraag waar een respondent op kan reageren. Meerdere items samen meten een construct wat zich op een schaal bevindt.

In de meeste vragenlijsten van Kwaliteitscholen wordt gebruik gemaakt van een 5-punts Likertschaal met als mogelijke scores 1 t/m 5. Soms is er een schaal van 1 tot 10, soms een tekstveld. Er komen ook 0-antwoorden voor, als weet niet en niet van toepassing.

Bij de statistische verantwoording van vragenlijsten van Kwaliteitscholen wordt gebruik gemaakt van een aantal standaardmaten binnen sociaalwetenschappelijk onderzoek. Deze zullen niet voor elke lezer als vanzelfsprekend kunnen worden verwacht. Daarom wordt alle voorkomende terminologie hier kort toegelicht.

Basismaten

<u>Basismaten</u>	
Centrale tendentie	
Gemiddelde	Som van alle scores gedeeld door het aantal observaties
Mediaan	Middelste waarneming
Modus	Meest voorkomende waarneming
Spreiding	
Variantie	Gemiddelde gekwadrateerde afstand tot het gemiddelde
Standaarddeviatie	Gemiddelde afstand tot het gemiddelde (wortel van variantie)
Vorm van de verdeling	
Skewness	Geeft de scheefheid van een verdeling aan.
Kurtosis	Geeft de mate van piekvormigheid aan.

Normaliteit

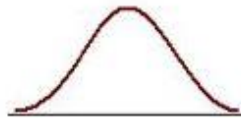
De verdeling van een variabele omschrijft de vorm welke de kansverdeling van de variabele aanneemt. De meest voorkomende is de zogeheten normaalverdeling. De normaal curve is een symmetrische verdeling met de typische Gass-Bell curve (zie onderstaand figuur). We kunnen meten wanneer er sprake is van een normaalverdeling, door te gebruik te maken van scheefheid (skewness) en de piekvormigheid (kurtosis). Wanneer deze 0 bedragen, is de variabele in kwestie normaal verdeeld. Veelal wordt als norm gebruikt dat beide waarden tussen -1 en +1 zijn wanneer een variabele in de populatie redelijk normaal verdeeld is, waar tussen de -0.5 en +0.5 sterk normaal verdeeld zou zijn. Hoewel er ook tests beschikbaar zijn om de normaliteit te formeel te testen (zoals de Shapiro Wilk test en Kolmogorov-Smirnov Test), zijn deze erg streng in hun beschouwing van normaliteit en daarom niet geschikt voor grotere datasets zoals deze (Field, 2009). Verder is de normaalverdeling van de items geen assumptie om de betrouwbaarheid te meten, dus is het benoemen van de verdeling van de variabelen slechts indicatief.

Skewness

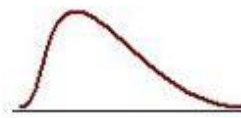
The coefficient of Skewness is a measure for the degree of symmetry in the variable distribution.



Negatively skewed distribution
or Skewed to the left
Skewness < 0



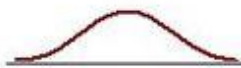
Normal distribution
Symmetrical
Skewness = 0



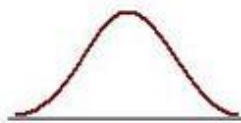
Positively skewed distribution
or Skewed to the right
Skewness > 0

Kurtosis

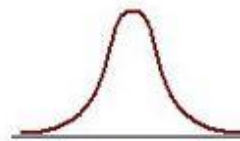
The coefficient of Kurtosis is a measure for the degree of peakedness/flatness in the variable distribution.



Platykurtic distribution
Low degree of peakedness
Kurtosis < 0



Normal distribution
Mesokurtic distribution
Kurtosis = 0



Leptokurtic distribution
High degree of peakedness
Kurtosis > 0

(bron: janzengroup.net)

Betrouwbaarheid

De betrouwbaarheid van een cluster van een aantal items kan worden gemeten aan de hand van verscheidene technieken. Deze berusten alle op een maat van gestandaardiseerde correlaties tussen een aantal variabelen. De meest gebruikte techniek binnen de sociale wetenschap is om de Cronbach's Alfa (Cronbach's α) te berekenen. De Cronbach's Alfa geeft een mate van homogeniteit van antwoorden op een vragenlijst weer: de interne consistentie. Cronbach's Alfa drukt de mate van betrouwbaarheid van een schaal uit in een proportie, wat te vertalen is naar een percentage.

De Commissie Testaangelegenheden Nederland (COTAN) beoordeelt psychodiagnostische instrumenten die in Nederland worden uitgegeven en bij de COTAN ter beoordeling aangeboden. Binnen het reguliere academisch onderzoek wordt meestal een minimumwaarde voor Cronbach's Alfa van .80 gesteld. Bij groepsonderzoek (bijvoorbeeld metingen van teamtevredenheid, klimaat in de klas of organisatiecultuur) dient de betrouwbaarheid van de schalen te liggen tussen .60 en .70 voor voldoende en gelijk aan of boven .70 voor goed (COTAN, 2010). Deze groepsnormen zullen in dit onderzoek worden aangehouden.

Naast de maat van Cronbach's Alfa voor de betrouwbaarheid, dient elke afzonderlijke vraag ook voldoende variantie te verklaren. De Item-restcorrelatie wordt door analyse software weergegeven als de Gecorrigeerde Item-Totaal Correlatie (item-test correlatie), meet de correlatie van elk item met de totaalscore op de schaal. Waardes lager dan .30 zijn problematisch voor de betrouwbaarheid van de totale schaal. De Gekwadraterde Multipele Correlatie (item-rest correlatie) is een maat waarmee wordt gemeten hoeveel variantie wordt verklaard in een vraag, wanneer deze op basis van de andere items wordt voorspeld. Het is aan te raden dat deze waarde zich boven .09 bevindt (Field, 2009).

Binnen de analyse van betrouwbaarheid wordt een aantal assumpties gemaakt. Ten eerste geldt de assumptie dat meer items betekent meer betrouwbaarheid. Hieruit valt af te leiden dat in situaties met een gelijke Cronbach's Alfa, de schaal met meer items de voorkeur geniet. Verder geldt voor de Cronbach's Alfa specifiek de assumptie van parallelle items en unidimensionaliteit. Parallelle items wil zeggen dat alle vragen een vergelijkbaar gemiddelde en standaarddeviatie hebben. Unidimensionaliteit verwijst naar het idee dat scores op een item allemaal in dezelfde antwoordrichting gecodeerd moeten zijn, zodat een hogere score op een bepaald item altijd een hogere score op de schaal weergeeft. Ten slotte moet een domein minstens 2 stellingen bevatten om de betrouwbaarheid te meten.