

In dit laatste hoofdstuk staan we stil bij het construeren van schalen. In de voorgaande hoofdstukken bedoelden we met 'schaal' veelal een antwoordschaal (een onderzoekseenheid kan op een schaal van 1 tot en met 5 (of 7 of 9 of ...) aangeven in hoeverre hij het eens is met een stelling) of een indexscore (een samengestelde schaal waarbij verschillende concrete variabelen bij elkaar werden opgeteld). Hier zullen we stilstaan bij een gemiddelde schaal. Met *schaal* bedoelen we hier de combinatie van een aantal variabelen (items) waarop onderzoekseenheden een score krijgen op een abstract, complex kenmerk. Een abstract complex kenmerk wordt ook wel een *concept* of *latente variabele* genoemd. Dit zijn kenmerken die je niet direct kunt meten of observeren, zoals bijvoorbeeld met een vraag naar een 'mening' of 'gedrag'. Je kunt wel door middel van verschillende *manifeste variabelen*, variabelen die je wel direct kunt meten, proberen zo goed mogelijk in de buurt te komen van het meten van dit concept. Je moet dan echter wel weten of je de juiste manifeste variabelen hebt gebruikt, en of je geen systematische en/of toevallige fouten hebt gemaakt bij het meten van het construct. Hoe je dit kunt nagaan zal in dit hoofdstuk beschreven worden.

10.1 Validiteit van een meting

Wanneer we kenmerken van onderzoekseenheden gaan meten kan het zijn dat we daar fouten bij maken. Deze fouten hebben invloed op de *betrouwbaarheid* en de *validiteit* van onze metingen. De betrouwbaarheid van een meting hangt af van het aantal toevallige fouten dat gemaakt wordt, en zal besproken worden in paragrafen 10.2, 10.3.2 en 10.6. Wanneer er *systematische* fouten worden gemaakt, hebben we problemen met de *validiteit* van de meting. Bij validiteit gaat het om de inhoudelijke betekenis van een meting; meet je inderdaad wat je had willen meten?

Sommige concepten zijn moeilijker te meten dan andere. De variabele 'leeftijd' bijvoorbeeld is vrij simpel te meten. Je kunt aan iemand vragen: 'Hoe oud ben je?' (... jaar), of je kunt vragen naar iemands geboortejaar. Welke vraagstelling ook gekozen wordt, bij het meten van leeftijd zul je weinig systematische fouten maken: je meet wat je wilt meten, namelijk hoe oud iemand is.

Bij complexe of abstracte begrippen zijn de problemen met betrekking tot validiteit groter, omdat de manier waarop het concept gemeten wordt minder voor

de hand ligt of omdat de metingen op wezenlijk verschillende manieren kunnen worden uitgevoerd. Meestal worden voor het meten van dergelijke concepten meerdere metingen verricht die elk afzonderlijk slechts een aspect van het abstracte begrip meten. Deze metingen zijn in de sociale wetenschappen meestal uitspraken of vragen in een vragenlijst. Door middel van een *factoranalyse* kan worden nagegaan of een gemeenschappelijke dimensie ten grondslag ligt aan de vragen die gekozen zijn om een abstract begrip te meten. Als uit de factoranalyse blijkt dat de gebruikte vragen een gemeenschappelijke onderliggende dimensie hebben, is dat een indicatie voor de validiteit van de meting van het abstracte begrip. Met een factoranalyse kun je aantonen dat verschillende variabelen (aspecten van) hetzelfde begrip meten.

10.1.1 Latente en manifeste variabelen

Stel dat je in je onderzoek wilt meten hoe 'media-afhankelijk' je respondenten zijn. De meest simpele optie zou zijn je respondenten op een antwoordschaal van bijvoorbeeld 1 tot 7 te laten aangeven hoe afhankelijk ze van de media zijn. Helaas valt op deze simpele optie heel wat af te dingen. Media-afhankelijkheid is een abstract en complex begrip. Je respondenten weten niet dat jij met media-afhankelijkheid doelt op 'de mate waarin een individu (informatie uit) de media gebruikt om zijn of haar doelen te bereiken'. Als je respondenten direct naar hun media-afhankelijkheid vraagt, zullen ze zelf de betekenis van dit begrip invullen en deze invulling zal sterk variëren onder je respondenten. Op een schaal van 1 tot 7 zullen zij aangeven hoe media-afhankelijk zij zichzelf vinden of hoe media-afhankelijk zij willen zijn. De kans dat twee personen die in feite even media-afhankelijk zijn, op deze schaal hetzelfde antwoord geven, is erg klein. Kortom, met deze simpele vraag meet je niet wat je wilt meten: de meting is dus niet valide. Beter is na te gaan hoe media-afhankelijkheid zich bij een individu manifesteert en daar vragen of uitspraken over te formuleren. Als onderzoeker begin je eerst met een *theoretische definitie* van het verschijnsel. In dit geval zou de theoretische definitie van 'afhankelijkheid van media' zijn 'de mate waarin een individu (informatie uit) de media gebruikt om zijn of haar doelen te bereiken'. Daarmee is het begrip echter nog niet geoperationaliseerd, dat wil zeggen meetbaar gemaakt voor je onderzoek. De *operationele definitie* van media-afhankelijkheid zou kunnen zijn: 'de mate waarin mensen vinden dat ze zich met de media vermaken, ze meer te weten komen over normen en waarden in de maatschappij, ze zichzelf en de wereld om zich heen beter begrijpen, ze weten wat ze in allerlei situaties het best kunnen doen en ze zichzelf met anderen kunnen vermaken'. In deze operationele definitie is een aantal manifeste variabelen opgesomd waarvan je als onderzoeker vindt dat het aspecten van het begrip 'media-afhankelijkheid' zijn.

Je gaat vervolgens de manifeste uitingen van het latente verschijnsel 'media-afhankelijkheid' meten. Je zou de respondenten bijvoorbeeld de volgende zes uitspraken (zes direct te meten uitspraken, zes *manifeste variabelen*) kunnen

voorleggen waarbij ze op een schaal van 1 tot 7 kunnen aangeven in hoeverre ze het ermee eens zijn (1 = zeer mee oneens, 7 = zeer mee eens):

Door de media te gebruiken ...

1. ... kan ik mezelf goed vermaken.
2. ... kom ik meer te weten over normen en waarden in de maatschappij.
3. ... begrijp ik meer van de wereld om mij heen.
4. ... ga ik mijzelf beter begrijpen.
5. ... weet ik wat ik in allerlei situaties het best kan doen.
6. ... kan ik me met anderen vermaken.

Als in deze manifeste variabelen de latente variabele media-afhankelijkheid inderdaad tot uitdrukking komt, zullen respondenten die sterk media-afhankelijk zijn op deze stellingen hoog scoren en zullen respondenten die nauwelijks afhankelijk van de media zijn, lage scores hebben. De correlaties tussen deze zes variabelen zullen dus hoog moeten zijn als er een gemeenschappelijke dimensie (latente variabele) aan ten grondslag ligt.

Manifeste variabelen (ook wel: *items*) zijn dus kenmerken die direct gemeten kunnen worden en latente variabelen (ook wel: *concepten*) zijn abstracte kenmerken die niet direct gemeten kunnen worden, maar waarvoor door het meten van een aantal manifeste variabelen een indicatie kan worden gevonden.

Tabel 10.1 Latente en manifeste variabelen

	Manifest (uiterlijk kenmerk)	Latent (abstract, innerlijk kenmerk)
Enkelvoudig (één variabele)	Bijv. sekse Bijv. leeftijd	Niet mogelijk
Samengesteld / complex (meer variabelen)	Index Bijv. kijktijd gemeten door totaal aantal minuten televisiekijken per week samen te voegen met het aantal uur per week	Schaal Bijv. media-afhankelijkheid gemeten door zes verschillende items

Tabel 10.1 laat zien dat manifeste variabelen zowel enkelvoudig als samengesteld kunnen worden gemeten, en dat latente variabelen (die zijn samengesteld door manifeste variabelen) altijd complex zijn. In de volgende paragraaf laten we zien hoe je door middel van een factoranalyse vast kunt stellen of een latente variabele valide gemeten is.

Wanneer blijkt dat de manifeste variabelen inderdaad hetzelfde verschijnsel meten (we laten in paragraaf 10.3.1 zien hoe je dat kunt vaststellen), kun je concluderen dat je een valide meting hebt uitgevoerd.¹

10.2 Betrouwbaarheid van een meting

Naast een valide meting is het ook van belang dat je een betrouwbare meting hebt uitgevoerd. De *betrouwbaarheid* van een meting is de mate waarin die meting vrij is van toevallige fouten. Je wilt graag dat je meting valide en betrouwbaar is, dus vrij van zowel systematische als toevallige fouten. Nadat je hebt vastgesteld of een meting valide is (zie paragraaf 10.3.1) ga je daarom na of de meting betrouwbaar is (zie ook paragraaf 10.3.2).

We hadden al gezien dat als een verschijnsel of kenmerk iets gecompliceerder is, zoals het eerder genoemde kenmerk 'afhankelijkheid van de media', het vaak niet mogelijk is om dit kenmerk slechts met één variabele te meten. Doordat het verschijnsel of kenmerk verschillende aspecten omvat, zijn meer variabelen nodig om het gehele begrip te dekken. We verwachten dat de variabelen (de *items*) die media-afhankelijkheid moeten gaan meten sterk met elkaar samenhangen. Respondenten met een hoge media-afhankelijkheid scoren op alle items hoog en respondenten met een lage media-afhankelijkheid scoren op alle items laag. Als de correlaties tussen al deze items hoog zijn, kunnen we stellen dat de schaal om media-afhankelijkheid te meten *intern consistent* is. Het is echter onhandig om uitspraken te doen over de interne consistentie op basis van een groot aantal correlatiecoëfficiënten (in bovenstaand voorbeeld zou je zes correlatiecoëfficiënten krijgen). We gebruiken daarom *Cronbachs alfa*, waarmee we met één kengetal de mate van interne consistentie aangeven. Deze maat voor interne consistentie geeft aan hoe betrouwbaar je schaal is. Of het meetinstrument voor het meten van media-afhankelijkheid betrouwbaar is, stel je dus vast door na te gaan of respondenten consistent scoren op de afzonderlijke variabelen. Als dat het geval is, kun je een gemiddelde schaal maken die de mate van afhankelijkheid van de media aangeeft.

10.3 Schaalconstructie

Met schaalconstructie bedoelen we het maken (construeren) van een nieuwe schaal op basis van bestaande variabelen. Zoals we al besproken hebben, is die nieuwe schaal een latente variabele (een construct, een abstract begrip) die niet in één keer is gemeten, maar op basis van een aantal variabelen (manifeste variabelen, items) is samengesteld. Een voorwaarde voor schaalconstructie is dat alle items op dezelfde manier gemeten zijn (dat wil zeggen: alle vragen moeten met dezelfde antwoordschaal beantwoord zijn)², en zij moeten allemaal minimaal een ordinaal meetniveau hebben. De uiteindelijke schaal, waar de gemiddelde scores van de respondenten op de verschillende items worden weergegeven, heeft altijd een interval meetniveau (het is immers een gemiddelde schaal).

Schaalconstructie heeft drie doelen. Ten eerste willen we een *valide meting* van een concept door meting van verschillende items. We willen de verschillende

aspecten van een latente variabele meten door middel van verschillende variabelen (items) en we willen nagaan of deze items inderdaad in voldoende mate hetzelfde meten.

Ten tweede wil je een *preciezer meting*. Wanneer je scores op verschillende items met elkaar combineert tot één schaal, krijg je meer genuanceerde verschillen tussen de scores van de onderzoekseenheden. Bij het maken van een schaal is het ook mogelijk om ordinale items, bijvoorbeeld variabelen met ordinale *Likertschalen*, samen te voegen. Een Likertschaal is een veelgebruikte antwoordschaal in de sociale wetenschappen om houdingen en meningen te meten, waarbij respondenten kunnen antwoorden op een schaal van het ene uiterste tot het andere uiterste, bijvoorbeeld van 1 = nooit tot en met 5 = zeer vaak, of van 1 = mee oneens tot en met 5 = mee eens. Veelal worden vijfpuntsschalen gebruikt, maar ook zeven- en negenpuntsschalen kunnen een Likertschaal zijn. Formeel zijn dit ordinale schalen. Maar bij een ordinale schaal van 1 (mee oneens) tot 5 (mee eens) *lijken* de intervallen tussen de getallen min of meer gelijk. Daarom kunnen we met dergelijke ordinale variabelen toch rekenen (zoals optellen, aftrekken en gemiddeldes berekenen). Wanneer je nu een gemiddelde schaal maakt, is de schaal niet langer discreet maar continu, en zijn ook scores mogelijk van 2,3 of 4,1 enzovoorts.

Tot slot heeft schaalconstructie tot doel *data te reduceren*. Als je de zes items die media-afhankelijkheid meten samenvoegt, hoef je in je onderzoek niet zes verschillende analyses uit te voeren om te kijken naar bijvoorbeeld de verschillen tussen mannen en vrouwen en hun media-afhankelijkheid (uitgaande van ons voorbeeld waar we zes vragen hadden gesteld om de mate van media-afhankelijkheid te meten), maar slechts één (namelijk één analyse met geslacht en media-afhankelijkheid als variabelen).

Bij het maken van een schaal worden altijd de volgende stappen gevolgd:

1. factoranalyse (vaststellen van de validiteit van de schaal);
2. betrouwbaarheidsanalyse (vaststellen van de interne consistentie van de schaal);
3. maken van de schaal;
4. beschrijven van de schaal.

10.3.1 Factoranalyse

We gaan kijken of we een schaal kunnen construeren voor de latente variabele 'smartphoneverslaving'. In een onderzoek is gekeken naar de mate waarin studenten aan een universiteit verslaafd zijn aan hun smartphone.³

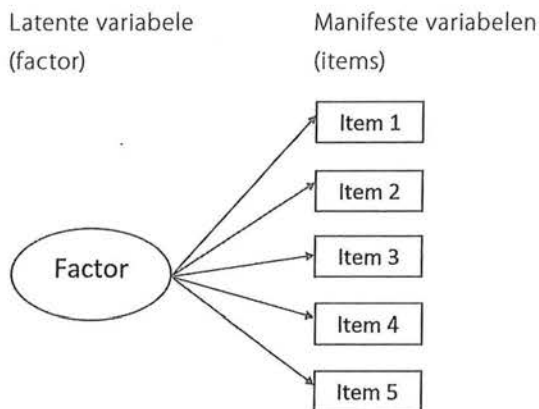
Om smartphoneverslaving te meten is een enquête afgenomen onder studenten waarbij op een zevenpunts Likertschaal variërend van 1 = geheel mee oneens tot en met 7 = geheel mee eens geantwoord kon worden:

1. Ik raak geïrriteerd wanneer mensen mij storen terwijl ik mijn smartphone gebruik.

2. Ik check voortdurend mijn smartphone om te kijken of ik niets heb gemist op sociale media.
3. Ik neem mijn telefoon mee naar het toilet, zelfs als ik heel nodig moet.
4. Er is niets leuker dan bezig te zijn met mijn smartphone.
5. Ik zou niet meer zonder smartphone kunnen.

Of dit vermoeden klopt, en of we inderdaad met deze vijf variabelen één latente variabele meten, zal moeten blijken uit de resultaten van de factoranalyse. Op basis van de gemeenschappelijke samenhang (correlaties) tussen een groep variabelen wordt een *factor* gevormd, de latente variabele. Die factor representeert wat de groep variabelen gemeenschappelijk heeft, een onderliggende dimensie. De onderliggende dimensie bij onze vijf vragen zou dan 'smartphoneverslaving' zijn.

Schematisch ziet het er als volgt uit:



Figuur 10.1 Latente en manifeste variabelen

In figuur 10.1 is te zien dat de pijlen vanuit de factor naar de items wijzen. Het idee van een factoranalyse is namelijk dat *omdat* de latente eigenschap aanwezig is, er op een bepaalde manier op de verschillende items wordt gescoord, en niet andersom. Kijken we bijvoorbeeld naar de eerste vraag die gesteld is om smartphoneverslaving te meten, dan stellen we bij een factoranalyse dat *omdat* een respondent verslaafd is aan zijn of haar smartphone, diegene hoog zal scoren op de vraag 'ik raak geïrriteerd wanneer mensen mij storen terwijl ik mijn smartphone gebruik'. Het is niet zo dat 'irritatie bij storen' de oorzaak is van smartphoneverslaving.

Als we de scores van de respondenten op de voornoemde vijf variabelen bij elkaar op zouden tellen, zou de resulterende variabele aangeven in welke mate de studenten verslaafd zijn aan hun smartphone. Maar mogen we dat wel doen? Dragen alle vijf variabelen bij aan het meten van hetzelfde begrip? We kunnen voor de eerste impressie een correlatiematrix uitdraaien (tabel 10.2) om te kijken of er een samenhang is tussen de variabelen. We vermoeden immers dat

een hoge score op de ene variabele gepaard zal gaan met een hoge score op de andere variabele.

Tabel 10.2 Correlatiematrix voor de zes variabelen m.b.t. smartphoneverslaving (SPSS-output)

		Correlations				
		v1 irritatie bij storen	v2 checken sociale media	v3 meenemen naar toilet	v4 niets leuker	v5 niet zonder kunnen
v1 irritatie bij storen	Pearson Correlation	1	,372**	,463**	,570**	,557**
	Sig. (2-tailed)		,000	,000	,000	,000
	N	239	239	239	239	239
v2 checken sociale media	Pearson Correlation	,372**	1	,584**	,549**	,618**
	Sig. (2-tailed)	,000		,000	,000	,000
	N	239	239	239	239	239
v3 meenemen naar toilet	Pearson Correlation	,463**	,584**	1	,565**	,580**
	Sig. (2-tailed)	,000	,000		,000	,000
	N	239	239	239	239	239
v4 niets leuker	Pearson Correlation	,570**	,549**	,565**	1	,714**
	Sig. (2-tailed)	,000	,000	,000		,000
	N	239	239	239	239	239
v5 niet zonder kunnen	Pearson Correlation	,557**	,618**	,580**	,714**	1
	Sig. (2-tailed)	,000	,000	,000	,000	
	N	239	239	239	239	239

** . Correlation is significant at the 0.01 level (2-tailed).

We zien dat er inderdaad een positieve samenhang is tussen alle items. Je zou kunnen redeneren dat studenten die de hoog scoren op 'ik neem mijn telefoon mee naar het toilet, zelfs als ik heel nodig moet' (v3) ook hoog scoren op 'er is niets leuker dan bezig te zijn met mijn smartphone' (v4).

Hoewel we aan de correlatiematrix een impressie kunnen ontleen van een construct, blijft het echter een overzicht van allemaal afzonderlijke bivariate correlaties. In een factoranalyse wordt de samenhang berekend van de afzonderlijke items met de onderliggende factor. Deze correlatie tussen een manifeste variabele (bijvoorbeeld 'Ik zou niet meer zonder smartphone kunnen') en een factor (hier: 'smartphoneverslaving') wordt de *factorlading* genoemd. Deze factorladingen variëren tussen -1 (perfecte negatieve samenhang met de factor) en $+1$ (perfecte positieve samenhang met de factor). Een factorlading van 0 betekent dat er geen relatie met de factor is. Er kan gesproken worden van een betekenisvolle samenhang wanneer de factorlading groter is dan $|0,45|$.⁴

Weer een andere naam voor factor, latente variabele of construct is *component*. In tabel 10.3 zie je de factorladingen die horen bij onze vijf items. Er wordt in dit geval één component gevonden waarbij alle items, alle manifeste variabelen, een factorlading hebben die groter is dan $0,45$. Dit wijst erop dat deze vijf items

inderdaad een valide meting zijn voor de latente variabele 'smartphoneverslaving'.⁵

Tabel 10.3 Componentenmatrix met factorladingen (SPSS-output)

Component Matrix^a

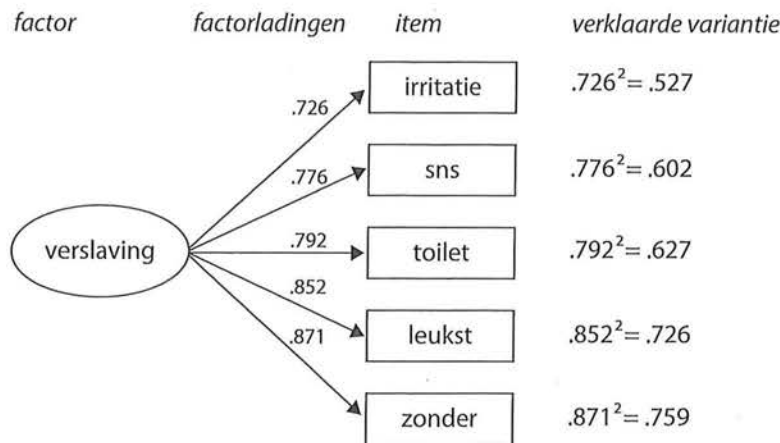
	Component
	1
v1 irritatie bij storen	,726
V2 checken sociale media	,776
v3 meenemen naar toilet	,792
v4 niets leuker	,852
v5 niet zonder kunnen	,871

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

Bij een factoranalyse worden coëfficiënten berekend, zodanig dat de factor de items zo goed mogelijk voorspelt. Je krijgt dus bij elke pijl vanuit de latente variabele een waarde (figuur 10.2), en die waarde geeft aan hoe goed de latente variabele de scores op dat item kan voorspellen. Die waarden zijn de factorladingen, en net als bij de gestandaardiseerde regressiecoëfficiënten en associatiematen liggen deze tussen -1 en $+1$ en hebben ze een richting. Je kunt deze factorladingen dus net zo interpreteren als de correlatiecoëfficiënt en de gestandaardiseerde regressiecoëfficiënt. De factorladingen geven de samenhang weer tussen de latente variabele en de manifeste items die we gemeten hebben. Een factorlading van 0 betekent dat er helemaal geen relatie is tussen de factor en het item, en een factorlading van $+1$ zou betekenen dat er een perfecte positieve samenhang zou zijn. Dat betekent dat een factorlading ook negatief kan zijn. Dat betekent niet dat de manifeste variabele dan niet bijdraagt aan een valide schaal. Wanneer de factorlading hoog genoeg is (dat wil zeggen hoger dan $|0,45|$) is het waarschijnlijk zo dat de manifeste variabele met een negatieve correlatie met de factor in dat geval 'omgekeerd' is gecodeerd. Wanneer we bijvoorbeeld de eerste vraag hadden geformuleerd als 'ik raak nooit geïrriteerd wanneer mensen mij storen terwijl ik mijn smartphone gebruik', zouden we nog steeds met die vraag 'smartphoneverslaving' meten, maar dan is juist een lage score op deze variabele een indicatie van de verslaving. In dat geval moet je dit item hercoderen (zie hoofdstuk 4) voordat je je schaal maakt.

Net als bij een regressieanalyse kunnen we ook een uitspraak doen over de proportie verklaarde variantie. Bij een factoranalyse kijken we dan hoe goed de variantie in de factor, de variantie in de afzonderlijke items verklaart.



Figuur 10.2 Factorladingen en verklaarde variantie per item voor 'smartphoneverslaving'

De verklaarde variantie kun je per item uitrekenen door het kwadraat te nemen van de factorlading. Uit figuur 10.2 kunnen we bijvoorbeeld opmaken dat de verslaving aan een smartphone sterk en positief samenhangt met de vraag of studenten zich geïrriteerd voelen als ze gestoord worden tijdens het telefoongebruik (*factorlading* = 0,73). De variantie in de factor (oftewel: de spreiding in de scores binnen de schaal voor smartphoneverslaving) verklaren voor 52,7% de variantie in hoe op het item van irritatie is gescoord. Zo kunnen we ook aflezen dat de variantie in smartphoneverslaving voor 75,9% de variantie in het antwoord 'ik zou niet meer zonder smartphone kunnen' verklaart.

Deze itemvarianties (de varianties per afzonderlijk item) worden niet door SPSS getoond, maar wel de *totale verklaarde variantie*. Dit is de gemiddelde verklaarde variantie van de afzonderlijke itemvarianties. In ons voorbeeld is de totale verklaarde variantie dus:

$$\text{totale verklaarde variantie} = \frac{0,527 + 0,602 + 0,627 + 0,726 + 0,759}{5} = \frac{3,241}{5} = 0,649$$

Dat wil zeggen dat de variantie in de factor voor 64,9% de variantie in de vijf items verklaart. Deze informatie vinden we ook in tabel 10.4, waar we onder andere in de laatste kolom (*Extraction Sums of Squared Loadings*) het percentage van 64,80 zien (de decimalen achter de komma wijken af wegens afrondingsverschillen). Dit betekent ook dat je een percentage informatieverlies hebt van 35,1% (100 – 64,90). De mate van smartphoneverslaving verklaart de variantie in de vijf items niet perfect. Ook andere factoren dan smartphoneverslaving zullen dus een rol spelen bij het beantwoorden van de vijf vragen. Er is geen standaardrichtlijn voor wat voldoende verklaarde variantie is en wat niet. Meestal beargumenteert de onderzoeker dat zelf bij zijn of haar resultaten.

Tabel 10.4 Totale verklaarde variantie en eigenwaarde in factoranalyse (SPSS-output)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,240	64,802	64,802	3,240	64,802	64,802
2	,657	13,142	77,944			
3	,452	9,039	86,984			
4	,375	7,503	94,486			
5	,276	5,514	100,000			

Extraction Method: Principal Component Analysis.

Als de verklaarde variantie laag is, meten we met onze items niet wat we willen meten en kunnen we geen valide schaal maken van de gebruikte items. Het kan ook zijn dat sommige items wel, en andere items niet bij de schaal horen (hier komen we op terug in de volgende paragraaf). Hoe meer de items hetzelfde meten, hoe hoger de *gemeenschappelijke variantie* zal zijn. Dit is de variantie die de latente variabele (component) gemeen heeft met de manifeste variabele. Wanneer alle items iets anders zouden meten dan in dit geval smartphoneverslaving, is de gemeenschappelijke variantie laag en kan er niet één factor of component worden gevormd. We willen dus dat de gemeenschappelijke variantie zo hoog mogelijk is, en de *specifieke variantie* en/of de *foutenvariantie* zo laag mogelijk zijn. De specifieke variantie is de variantie die alleen door die specifieke variabele gemeten wordt en die dus uniek is voor die variabele. Foutenvariantie is de variantie die geheel willekeurig is en die geen systematisch verband heeft met enige andere bron van variantie. Dit zou kunnen komen door meetfouten die worden veroorzaakt door onzorgvuldig invullen van een vragenlijst, gokken bij multiplechoicevragen enzovoort. Welk deel precies specifieke variantie en welk deel foutenvariantie is, is niet uit een factoranalyse af te leiden.

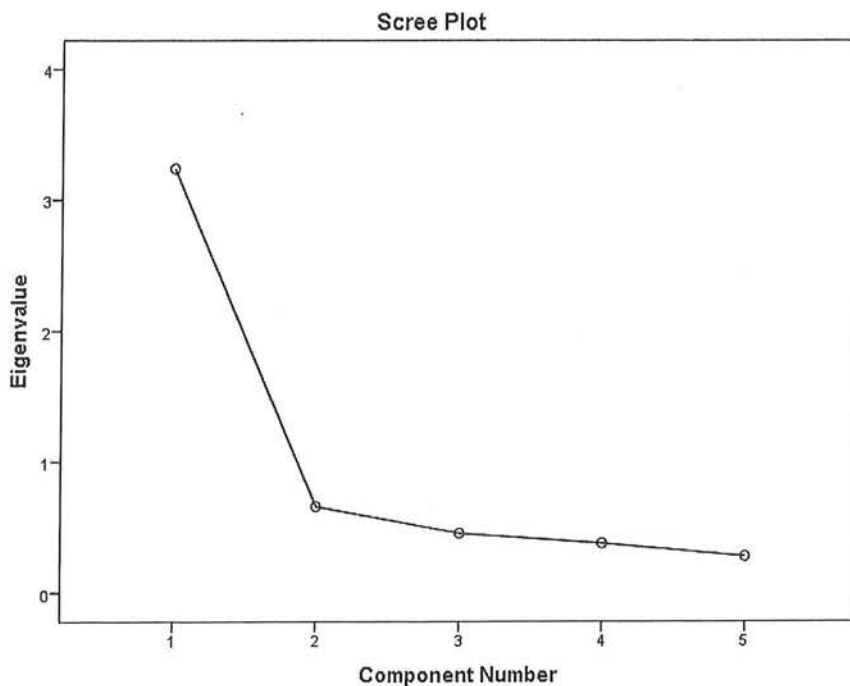
Hoe hoger de gemeenschappelijke variantie, hoe sterker de items onderling met elkaar en met de factor correleren, en hoe 'sterker' en informatiever onze factor is. Deze informativiteit wordt uitgedrukt in de *eigenwaarde* van de component. In tabel 10.4 zien we ook de *eigenwaarde* (in het Engels *Eigenvalue*) staan, in de eerste kolom. De eigenwaarde geeft aan hoe informatief de factor is. Wanneer je de items samenneemt tot één factor, is deze dan informatiever dan de losse items apart? Vaak besluiten onderzoekers om alleen factoren of componenten te gebruiken waarvoor de eigenwaarde groter is dan 1 (*criterium van Kaiser*). Dit is een arbitraire grens.

Ook de eigenwaarde wordt berekend aan de hand van de individuele itemvarianties. De *eigenwaarde* van een factor wordt berekend door de factorladingen (correlaties tussen variabelen en de factor) te kwadrateren en daarna te sommeren. Het is dus de teller van onze som bij het berekenen van de totale variantie, en

die is hier 3,241, zoals ook in tabel 10.4 te zien is in de eerste kolom onder *Initial Eigenvalues – Total*. De hoogst mogelijke eigenwaarde is in ons voorbeeld 5. We hebben immers vijf items gemeten, en wanneer deze precies hetzelfde zouden meten (zonder enige specifieke variantie of foutenvariantie) hebben we een perfecte schaal geconstrueerd zonder enig informatieverlies.

In tabel 10.4 zien we dat SPSS vijf componenten berekent als we vijf manifeste variabelen hebben. Alleen de eerste component heeft een eigenwaarde groter dan 1 en is de factor die we kunnen gebruiken. Dat komt goed uit want het was ook de bedoeling dat deze items één concept (namelijk smartphoneverslaving) zouden meten.

Een alternatief voor het gebruik van de grens van de waarde 1 voor de eigenwaarde is een *screeplot*. Ook daarmee kun je kijken naar de componenten die door een factoranalyse worden onderscheiden. De eigenwaarden van de componenten of factoren zijn hierin op volgorde gezet en vaak is duidelijk een ‘elleboog’ te zien in de lijn die deze eigenwaarden met elkaar verbindt (zie figuur 10.3). De onderzoeker beperkt zijn factoren dan tot het punt van de elleboog (de ‘knik’). Ook op basis van deze screeplot kunnen we concluderen dat er één component ten grondslag ligt aan de vijf manifeste items. Een screeplot is voornamelijk een handig grafische hulpmiddel wanneer er meer dan één component wordt onderscheiden, waar we in de volgende paragraaf bij stil zullen staan.



Figuur 10.3 Screeplot met één component

Een factoranalyse dient dus om te kijken in welke mate een aantal manifeste variabelen dezelfde latente variabele meten. Dit is een vorm van *constructvaliditeit*. Constructvaliditeit geeft aan in hoeverre het meetinstrument meet wat je wilt meten. Als de manifeste variabelen unidimensioneel zijn, dat wil zeggen

als er slechts één dimensie/factor aan ten grondslag ligt, kun je concluderen dat je variabelen één latente variabele meten.

Concluderend kunnen we aan de hand van de factoranalyse stellen dat onze vijf items een valide schaal vormen voor de latente variabele 'smartphoneverslaving'. Er werd namelijk één component gevormd met een eigenwaarde hoger dan 1. Bovendien was de totale verklaarde variantie voldoende hoog, en waren alle factorladingen hoger dan $|0,45|$.

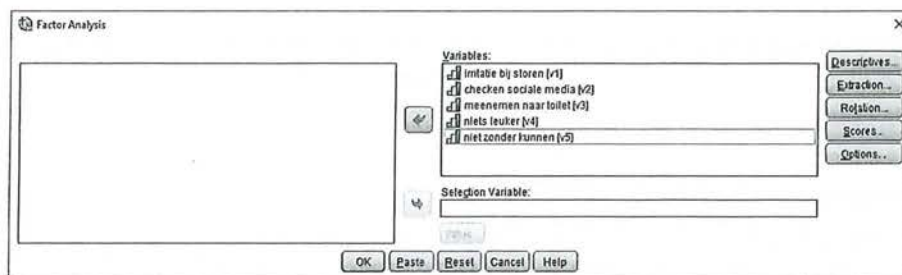


SPSS

Uitvoeren van een factoranalyse

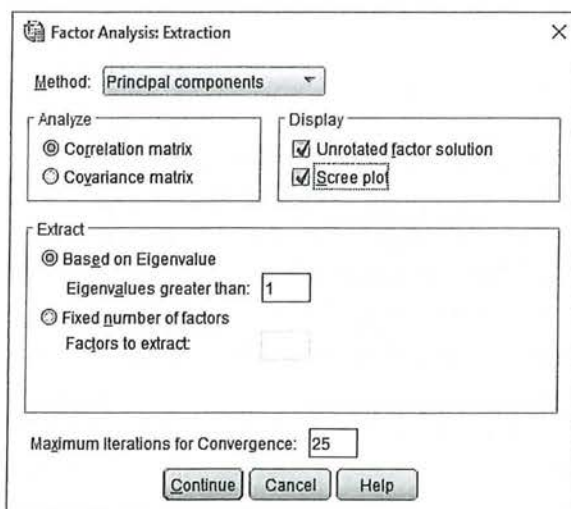
Wanneer we een factoranalyse in SPSS laten uitvoeren (via *Analyze* → *Dimension Reduction* → *Factor*), hebben we tal van opties om aan te geven welke informatie we willen.

In het vak *Variables* plaatsen we alle variabelen waar we een factoranalyse over willen uitdraaien. In figuur A is dat gedaan voor de vijf variabelen waarvan we denken dat ze een valide schaal voor smartphoneverslaving zullen meten.



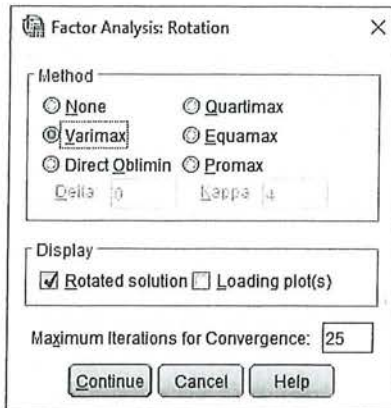
Figuur A Factor Analysis-venster

Wanneer je vervolgens op *Extraction* klikt, kun je onder *Display* aangeven dat je een screeplot wilt uitdraaien (figuur B).



Figuur B Extraction-venster: Scree plot aanvinken

Tot slot moet je bij Rotation (figuur C) aangeven dat je een *Varimaxrotatie* wilt laten uitvoeren. Het doel en nut daarvan wordt besproken in paragraaf 10.4.



Figuur C Rotation-venster: Varimax aanvinken

Kader 10.1

10.3.2 Betrouwbaarheidsanalyse: interne consistentie

Nadat we ons hebben verzekerd van de validiteit van onze meting, gaan we na of de schaal ook betrouwbaar is. Is het niet op basis van toeval dat de items met elkaar correleren en daardoor een valide schaal vormen? Zoals we in paragraaf 10.2 reeds aankondigden, wordt de betrouwbaarheid van de schaal, de interne consistentie van de schaal, gemeten aan de hand van Cronbachs alfa. SPSS geeft ons hiervoor twee tabellen, één met de waarde van alfa, en één waarin staat of we de mogelijkheid hebben om te schaal te verbeteren (dat wil zeggen: betrouwbaarder te maken).

Tabel 10.5 Betrouwbaarheidsanalyse door Cronbachs alfa (SPSS-output)

Reliability Statistics

Cronbach's Alpha	N of Items
,856	5

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
v1 irritatie bij storen	20,34	24,494	,577	,848
v2 checken sociale media	20,87	21,640	,650	,832
v3 meenemen naar toilet	21,15	20,002	,672	,831
v4 niets leuker	20,31	23,149	,738	,813
v5 niet zonder kunnen	20,40	22,250	,766	,803

De richtlijnen bij Cronbachs alfa zijn anders dan bij de interpretatie van factorladingen. We hanteren een ondergrens van 0,60 voordat we kunnen spreken van een redelijk betrouwbare schaal. Vanaf 0,80 vinden we de schaal betrouwbaar. In tabel 10.5 is in de tabel *Reliability Statistics* af te lezen dat onze vijf items een betrouwbare schaal kunnen vormen voor de latente variabele smartphonever-slaving: $\alpha = 0,86$. In de tabel daaronder, *Item-Total Statistics*, wordt per variabele aangegeven wat de betrouwbaarheid van de schaal zou zijn wanneer we dat item niet in onze schaal zouden opnemen (in de laatste kolom onder *Cronbach's Alpha if Item Deleted*). In dit geval kunnen we de schaal niet verbeteren: geen van de waarden in deze kolom is hoger dan de oorspronkelijke alfa van 0,856. We hoeven overigens de schaal ook niet te verbeteren, want de betrouwbaarheid is al hoog genoeg.

Mocht je een negatieve (maar voldoende hoge) factorlading bij het interpreteren van de factoranalyse over het hoofd hebben gezien, dan zul je in de kolom *Corrected Item-Total Correlation* dit nog kunnen nagaan door te kijken of hier een negatieve waarde in voorkomt. Is dat het geval, dan is het waarschijnlijk zo dat het item nog gehercodeerd moet worden.

Ook Cronbachs alfa kunnen we met de hand berekenen, en dat doen we aan de hand van een correlatiematrix. Aangezien alfa in de formule gebruikmaakt van correlaties op minimaal intervalniveau, dus van Pearsons correlaties, gebruiken wij deze ook, hoewel officieel het meetniveau van de variabelen ordinaal is. De informatie die we hiervoor gebruiken is de correlatiematrix zoals weergegeven in tabel 10.2.

Als verschillende variabelen hetzelfde verschijnsel (moeten) meten, is het te verwachten dat de correlatie tussen die variabelen hoog is. Zou één van de variabelen een lage of een qua richting tegengestelde correlatie met de andere variabelen vertonen, dan is deze variabele niet consistent. Die variabele wijkt dan af van het algemene patroon. Cronbachs alfa geeft een indicatie van de interne consistentie van de schaal die is samengesteld op basis van een aantal items. Bij de berekening van Cronbachs alfa maak je gebruik van de onderlinge correlaties. Het is een soort gemiddelde correlatie. Daarnaast is ook het aantal items van invloed op de waarde van Cronbachs alfa.

$$\alpha = \frac{k\bar{r}}{1 + (k-1)\bar{r}}$$

Formule voor Cronbachs alfa

In deze formule staat k voor het aantal items en \bar{r} voor het gemiddelde van de correlaties.

Met behulp van de correlaties tussen de vijf items die tezamen de schaal smartphoneverslaving vormen (tabel 10.2) kun je Cronbachs alfa berekenen. We hebben tien correlatiecoëfficiënten. De gemiddelde correlatie is de som van deze correlaties, gedeeld door het aantal correlaties:

$$\begin{aligned}\bar{r} &= \frac{0,372+0,463+0,570+0,557+0,584+0,549+0,618+0,565+0,580+0,714}{10} \\ &= \frac{5,572}{10} = 0,557\end{aligned}$$

We hebben vijf items, dus voor k vullen we 5 in.

$$\alpha = \frac{k\bar{r}}{1+(k-1)\bar{r}} = \frac{5*0,557}{1+(5-1)*0,557} = \frac{2,785}{3,228} = 0,863$$

De met de hand berekende alfa valt iets hoger uit dan de alfa zoals berekend door SPSS omdat wij met drie decimalen achter de komma rekenen en SPSS met meer decimalen. De interpretatie blijft hetzelfde: α is groter dan 0,80 en we hebben een goede betrouwbare schaal. De vijf variabelen vormen een goede schaal die intern consistent is. We kunnen nu ook ons databestand uitbreiden door een nieuwe variabele 'smartphoneverslaving' te maken.

10.3.3 Maken en beschrijven van de schaal

Nu we hebben vastgesteld dat we een valide en betrouwbare schaal kunnen construeren voor smartphoneverslaving, kunnen we de schaal daadwerkelijk gaan maken. Dit doen we door gebruik te maken van het commando 'MEAN' zoals beschreven in hoofdstuk 4 (paragraaf 4.3, en kader 4.2). Onze nieuwe schaal heeft als meetniveau interval: het is een gemiddelde schaal waarin scores op vijf items bij elkaar worden gevoegd en gedeeld door het aantal items. We kunnen voor het beschrijven van de schaal gebruikmaken van *Explore* (zie paragraaf 3.4 en kader 3.1).

Op deze schaal van 1 tot 7, waarbij hoe hoger gescoord wordt, hoe meer de onderzoekseenheden smartphoneverslaafd zijn, scoren studenten gemiddeld 5,15 met een standaarddeviatie van 1,16. Dat wil zeggen dat de studenten behoorlijk verslaafd zijn aan hun smartphone. Uit deze tabel kan ook worden opgemaakt dat de schaal niet te scheef verdeeld is (*Skewness* = -0,62).

Tabel 10.6 Explore om schaal te beschrijven (SPSS-output)

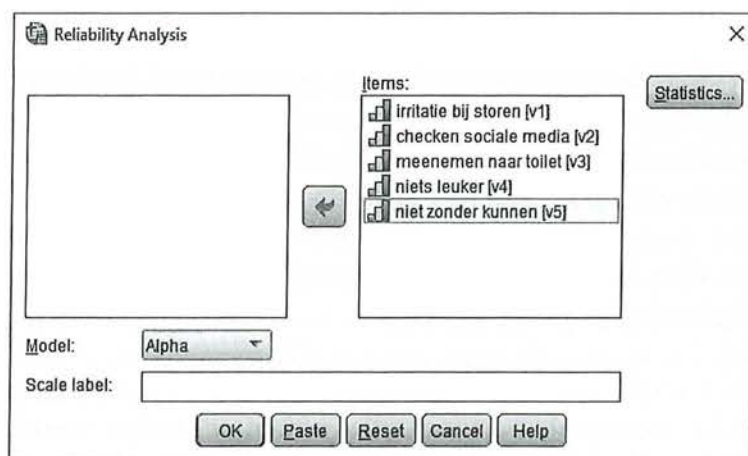
Descriptives			Statistics	Std. Error
smartphoneverslaving	Mean		5,1540	,07503
	95% Confidence Interval for Mean	Lower Bound	5,0062	
		Upper Bound	5,3018	
	5% Trimmed Mean		5,2055	
	Median		5,2000	
	Variance		1,346	
	Std. Deviation		1,16000	
	Minimum		1,20	
	Maximum		7,00	
	Range		5,80	
	Interquartile Range		1,80	
	Skewness		-,624	,157
	Kurtosis		,201	,314



SPSS

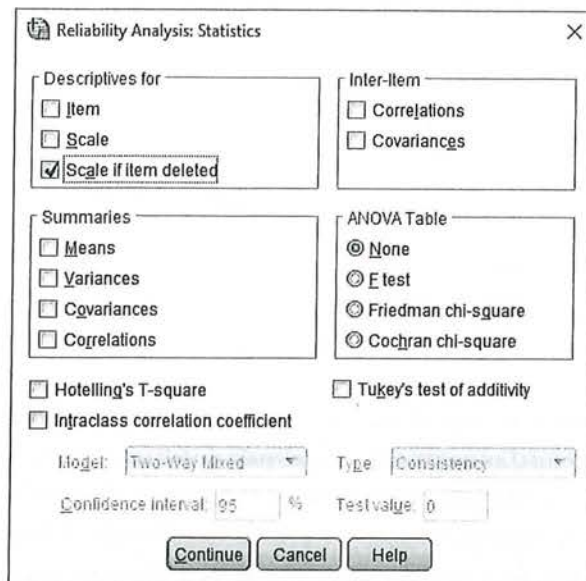
Berekenen van Cronbachs alfa

Het berekenen van Cronbachs alfa in SPSS gaat via *Analyze* → *Scale* → *Reliability Analysis*. In het vak *Items* vul je vervolgens de variabelen in die samen een schaal zouden moeten vormen (zie figuur A).



Figuur A Reliability Analysis-venster

Via *Statistics* kun je aan SPSS extra informatie vragen (zie figuur B), waarbij het aanvinken van 'Scale if Item deleted' de belangrijkste is.



Figuur B Statistics-venster

Kader 10.2

We hebben nu alle stappen van een schaalconstructie doorlopen en komen tot de volgende conclusie:

Uit een factoranalyse blijkt dat we één factor kunnen vormen voor de latente variabele 'smartphoneverslaving' met vijf manifeste items. Er is één component met een eigenwaarde hoger dan 1 ($EV = 3,42$) en de totale verklaarde variantie is 64,80%. Dit is een valide schaal waarbij geldt: hoe hoger wordt gescoord, hoe meer verslaafd de studenten zijn. De schaal is betrouwbaar ($\alpha = 0,86$). Op de schaal, die loopt van 1 tot en met 7, scoren studenten gemiddeld hoog: $M = 5,15$, $SD = 1,16$.

10.4 Meerdere factoren

Met een factoranalyse bereken je op basis van onderlinge correlaties een nieuwe variabele (een factor of component), die het best de onderliggende dimensie weergeeft. Het resultaat van een factoranalyse kan ook zijn dat er meer dan één dimensie ten grondslag ligt aan de variabelen die je gebruikt. We leggen dit uit aan de hand van een fictief voorbeeld.

Stel dat je wilt weten of leeftijd samenhangt met de vraag of mensen graag uitgaan. Je verwacht dat naarmate mensen ouder zijn, ze minder graag uitgaan. Je formuleert verschillende uitspraken over uitgaan, waarbij respondenten op een schaal van 1 tot 5 kunnen aangeven in welke mate ze het eens zijn met die

uitspraken (1 = zeer mee oneens, 5 = zeer mee eens). De volgende uitspraken leggen we ter beantwoording aan respondenten voor:

- Ik ga graag naar het café.
- Ik ga graag naar een popconcert.
- Ik ga graag naar een (dans)club.
- Ik ga graag naar een klassiek concert.
- Ik ga graag naar het theater.
- Ik ga graag naar het museum.

Als de zes stellingen over uitgaan, de manifeste variabelen, inderdaad 'het uitgaan', de latente variabele, meten, moeten ze voldoende informatie gemeenschappelijk hebben en elk apart iets toevoegen. Maar als je naar de stellingen kijkt, is het voorstelbaar dat die 'gemeenschappelijkheid' ontbreekt. Wanneer iemand bijvoorbeeld graag naar het theater gaat, graag een klassiek concert bezoekt, maar liever niet naar een café, popconcert of dansclub gaat, zou zijn score op de som van de zes items hetzelfde zijn als de score van iemand die juist graag naar café, popconcert en dansclub gaat en liever niet naar theater, concert of museum. Dat dit inderdaad het geval is, blijkt uit de correlatiematrix (tabel 10.7). Tussen de variabelen café, popconcert en club zien we sterke samenhangen en ook tussen de variabelen klassiek concert, theater en museum, terwijl de overige correlaties zwak tot nihil zijn. In deze correlatiematrix zijn duidelijk twee clusters te onderscheiden. Het eerste cluster wordt gevormd door de variabelen V1, V2 en V3 (café, popconcert en club) en het tweede cluster door de variabelen V4, V5 en V6 (klassiek concert, theater en museum).

Tabel 10.7 Correlatiematrix verschillende uitgaansbestemmingen

	V1	V2	V3	V4	V5	V6
V1 Café	1,00					
V2 Popconcert	0,680	1,00				
V3 Club	0,721	0,720	1,00			
V4 Klassiek concert	0,127	0,165	0,117	1,00		
V5 Theater	0,149	0,204	0,157	0,668	1,00	
V6 Museum	0,206	0,230	0,199	0,641	0,674	1,00

Je zou kunnen redeneren dat diegenen die hoog scoren op de variabelen klassiek concert, theater en museum hun vrije tijd liever aan 'kunst en cultuur' besteden en diegenen die hoog scoren op café, popconcert en club hun vrije tijd liever aan 'sociaal vermaak' besteden. De manifeste variabelen 'bezoek theater', 'museumbezoek' en 'bezoek klassieke concerten' zouden dan samen de latente variabele 'uitgaan: kunst en cultuur' meten. Café-, popconcert- en dansclubbezoek zijn dan de manifeste variabelen die de latente variabele 'uitgaan: sociaal vermaak' meten.

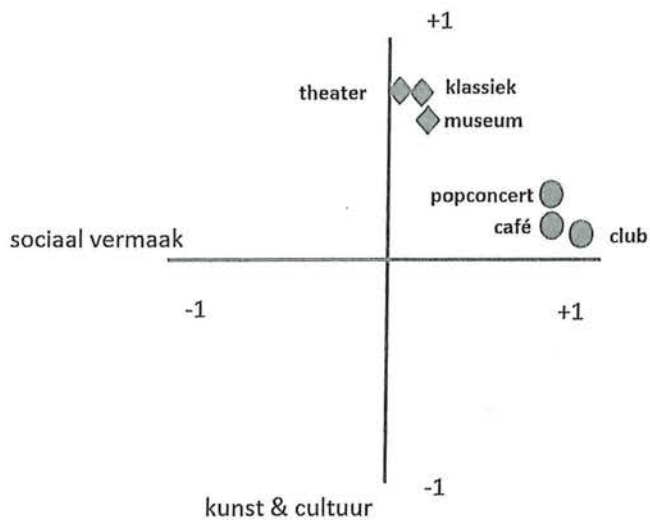
De namen van de factoren ('uitgaan: kunst en cultuur' en 'uitgaan: sociaal vermaak') hebben we in dit geval gekozen op basis van theoretische overwegingen. Een onderzoeker kan de naamgeving van de factoren ook pas naderhand bepalen op basis van de variabelen die het sterkst aan een factor bijdragen (de hoogste factorladingen hebben). Als onderzoekers een factoranalyse uitvoeren zonder dat ze van tevoren een idee hebben van mogelijke latente variabelen, en vervolgens op basis van de sterkst ladende manifeste variabelen een naam aan de gevonden factoren geven, spreken we van een *explorerende* factoranalyse. De naamgeving aan de factoren is afhankelijk van de keuze van de onderzoeker zelf. Bij een explorerende factoranalyse speelt de eigen interpretatie een grote rol bij de duiding van de factoren/latente variabelen. De onderzoeker gebruikt de variabelen die op een factor een hoge lading hebben om tot een interpretatie van de factor te komen. Als onderzoekers op theoretische gronden verwachten met een aantal manifeste variabelen bepaalde latente verschijnselen te meten, is sprake van een *confirmatieve* factoranalyse. De factoranalyse dient dan als bevestiging van de verwachtingen van de onderzoeker met betrekking tot de onderliggende structuur, dimensies van een verschijnsel. Met een confirmatieve factoranalyse hoopt de onderzoeker vooraf benoemde latente variabelen in de factoren te herkennen.

In de correlatiematrix van tabel 10.7 konden we zelf al twee clusters onderscheiden. Als we in SPSS een factoranalyse uitvoeren, wordt dit bevestigd. Er zijn inderdaad twee factoren. In de zogenoemde factorladingmatrix (tabel 10.8) zien we voor elke variabele de correlaties – de factorladingen – met die twee latente variabelen.

Tabel 10.8 Factorladingen voor twee componenten

	Factor 1	Factor 2
Café	0,888	0,083
Popconcert	0,880	0,136
Club	0,907	0,076
Klassiek	0,093	0,887
Theater	0,049	0,877
Museum	0,150	0,864

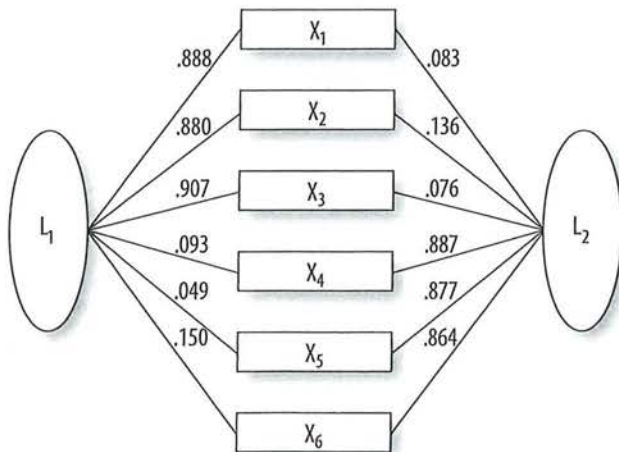
Als we deze informatie van de factorladingen grafisch zouden weergeven op een assenstelsel (figuur 10.4), wordt nogmaals duidelijk dat het hier om twee onderscheiden clusters gaat. Het idee van een factoranalyse is dat een item óf bij de ene factor, óf bij de andere factor wordt ingedeeld.



Figuur 10.4 Factorladingen van uitgaan items

De variabelen 'café', 'popconcert' en 'club' scoren allemaal hoog op factor 1, maar laag op factor 2. Bij de 'kunst en cultuur'-variabelen is dat juist omgekeerd. Hieruit is af te leiden dat 'café', 'popconcert' en 'club' gezamenlijk één factor, of latente variabele, vormen, namelijk 'uitgaan: sociaal vermaak', en 'klassiek concert', 'theater' en 'museum' de latente variabele 'uitgaan: kunst en cultuur' vormen.

Deze factorladingen staan bij de lijnen in de schematische voorstelling van de latente en manifeste variabelen van figuur 10.5.

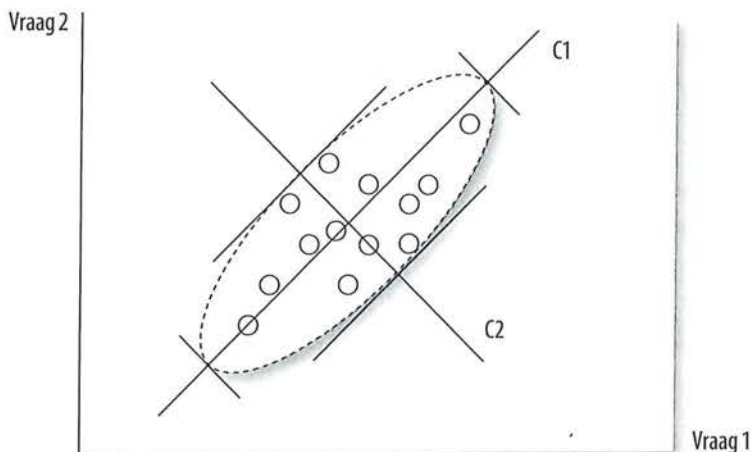


Figuur 10.5 Factorladingen schematisch

10.4.1 Het vinden van de factoren

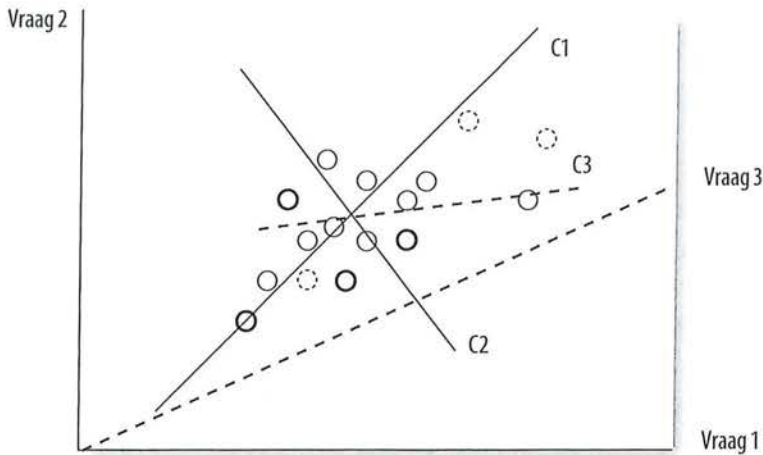
Er zijn verschillende methoden waarop onderliggende factoren uit de correlaties tussen de manifeste variabelen kunnen worden geëxtraheerd. *Principale componentenanalyse* (PCA) is een van deze methoden. Bij principale componentenanalyse worden de onderliggende dimensies die uit de samenhang tussen de variabelen worden geëxtraheerd, componenten¹ genoemd.

Om te verduidelijken hoe de onderliggende dimensies gevonden worden, starten we eerst met maar twee (manifeste) variabelen. In het spreidingsdiagram van figuur 10.6 zijn de scores van de respondenten op twee variabelen weergegeven. Aan de puntenwolk in dit spreidingsdiagram is te zien dat er samenhang is tussen de twee variabelen. Er kan een lijn worden getrokken die de variantie binnen de puntenwolk het best verklaart. Dit is de eerste component (C1). De tweede component (C2) wordt hier loodrecht op gezet. Door de tweede component loodrecht op de eerste te plaatsen is er geen samenhang tussen de twee componenten. Deze tweede lijn verklaart ook nog enige variantie in de puntenwolk, maar veel minder. Deze twee componenten zijn de onderliggende dimensies die de variantie in de puntenwolk verklaren, zij vertonen onderling geen samenhang.



Figuur 10.6 Spreidingsdiagram van twee variabelen

Als we hetzelfde doen voor drie manifeste variabelen, ontstaat een driedimensionale puntenwolk en vinden we een derde component die loodrecht op de eerste twee kan worden getekend. We zien dan een driedimensionaal assenstelsel en dat kunnen we ons nog wel voorstellen, maar bij vier manifeste variabelen is het visualiseren van de componenten al onmogelijk.



Figuur 10.7 Driedimensionaal spreidingsdiagram van drie variabelen

Bij de principale componentenanalyse worden op deze manier altijd evenveel componenten gevonden als er manifeeste variabelen zijn en deze componenten verklaren tezamen 100% van de variantie. Bij factoranalyse wordt in een puntenwolk gezocht naar een lijn, een factor, die zo veel mogelijk van de gemeenschappelijke variantie verklaart. Een volgende factor moet daarna zo veel mogelijk van de nog resterende gemeenschappelijke variantie verklaren.

10.4.2 Het interpreteren van de factoren en de noodzaak van rotatie

Bij de principale componentenanalyse zijn er evenveel componenten als manifeeste variabelen. Een onderzoeker wil het aantal variabelen graag zo ver mogelijk reduceren tot die componenten of factoren die er werkelijk toe doen. Hoe bepaal je nu hoeveel factoren er echt toe doen? Dat gebeurt op basis van de verklaarde variantie en de eigenwaarden (zie ook paragraaf 10.3.1). De eerste component is altijd de component die de meeste variantie verklaart, bij elke volgende is de verklaarde variantie lager en die is dus ook minder belangrijk. Naast een percentage verklaarde variantie wordt de informativiteit van een factor uitgedrukt in de eigenwaarde. Deze werd per component berekend door de factorladingen te kwadrateren en bij elkaar op te tellen. Deel je de uitkomst daarvan door het aantal items, dan krijg je de proportie verklaarde variantie per component.

Om te illustreren hoe je als onderzoeker probeert je factoren te interpreteren, gaan we verder met de zes stellingen die in paragraaf 10.1.1 zijn geformuleerd om de media-afhankelijkheid van respondenten te meten. Anders dan in het fictieve voorbeeld van de uitgaansvariabelen zijn in de correlatiematrix van de media-afhankelijkheidsvariabelen niet direct clusters te onderscheiden (tabel 10.9). De correlaties variëren tussen 0,02 en 0,58.

Tabel 10.9 Correlatiematrix media-afhankelijkheidsvariabelen (SPSS-output)

		Correlations					
		v1 ver- maken	v2 normen en waarden	v3 begrip wereld	v4 weten wat te doen	v5 mijzelf begrijpen	v6 anderen vermaken
v1 verma- ken	Pearson Correlation	1	,015	,122	,277**	,352**	,584**
	Sig. (2-tailed)		,818	,059	,000	,000	,000
	N	239	238	239	239	239	239
v2 normen en waar- den	Pearson Correlation	,015	1	,330**	,312**	,357**	,151*
	Sig. (2-tailed)	,818		,000	,000	,000	,019
	N	238	238	238	238	238	238
v3 begrip wereld	Pearson Correlation	,122	,330**	1	,399**	,403**	,240**
	Sig. (2-tailed)	,059	,000		,000	,000	,000
	N	239	238	239	239	239	239
v4 weten wat te doen	Pearson Correlation	,277**	,312**	,399**	1	,451**	,318**
	Sig. (2-tailed)	,000	,000	,000		,000	,000
	N	239	238	239	239	239	239
v5 mijzelf begrijpen	Pearson Correlation	,352**	,357**	,403**	,451**	1	,440**
	Sig. (2-tailed)	,000	,000	,000	,000		,000
	N	239	238	239	239	239	239
v6 anderen vermaken	Pearson Correlation	,584**	,151*	,240**	,318**	,440**	1
	Sig. (2-tailed)	,000	,019	,000	,000	,000	
	N	239	238	239	239	239	239

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

We gebruiken SPSS om door middel van een factoranalyse hier de onderliggende dimensies (factoren) uit te halen. In de output (tabel 10.10) vinden we een tabel met de geëxtraheerde componenten met de waarden van de eigenwaarden. De eigenwaarde van de eerste component is 2,63, deze component verklaart 43,8% van de variantie in de zes variabelen. De eigenwaarde van de tweede component is 1,23, deze component verklaart 20,5% van de variantie in de zes variabelen. De overige componenten hebben een eigenwaarde kleiner dan 1. Op basis daarvan kan de onderzoeker besluiten het onderzoek te beperken tot de eerste twee componenten. Gezamenlijk verklaren deze twee componenten 64,3% van de variantie. Daarmee accepteert de onderzoeker informatieverlies: de zes manifeste variabelen geven meer informatie over de variatie onder de respondenten dan de twee latente variabelen waartoe hij zich gaat beperken. Daar staat tegenover dat hij nu twee variabelen kan maken waarmee de mate van media-afhankelijkheid wordt gekwantificeerd.

Tabel 10.10 Verklaarde variantie bij factoranalyse met media-afhankelijkheidsvariabelen (SPSS-output)

Component	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,628	43,803	43,803	2,628	43,803	43,803	2,001	33,343	33,343
2	1,230	20,493	64,296	1,230	20,493	64,296	1,857	30,953	64,296
3	,674	11,227	75,524						
4	,582	9,700	85,224						
5	,494	8,233	93,458						
6	,393	6,542	100,000						

Extraction Method: Principal Component Analysis.

SPSS geeft ons ook een componentenmatrix (tabel 10.11) waarin de factorladingen op de twee componenten met een eigenwaarde groter dan 1 zijn gegeven. Deze tabel is te vergelijken met tabel 10.8, waar we de factorladingen op 'uitgaan' lieten zien. Op basis van de factorladingen in deze matrix kan de onderzoeker proberen de componenten te benoemen. Dat wordt in dit geval bemoeilijkt doordat een aantal variabelen op beide componenten hoog laden. Bij bijvoorbeeld de variabele 'mijzelf begrijpen' is de factorlading op de eerste component 0,786 en op de tweede component 0,062. Deze variabele 'hoort' dus duidelijk bij component 1. Maar bij de variabele 'normen en waarden' is de factorlading op de eerste component 0,507 en op de tweede component 0,587. Beide ladingen zijn in ieder geval hoger dan $|0,45|$, wat betekent dat de factorlading hoog genoeg is. Nu zouden we gewoon kunnen zeggen: 0,587 (factorlading van de variabele 'normen en waarden' op component 2) is hoger dan 0,507 (factorlading van de variabele 'normen en waarden' op component 1), maar zo eenvoudig werkt het niet.

Tabel 10.11 Componentenmatrix (SPSS-output)

	Component Matrix ^a	
	1	2
v1 vermaken	,598	-,660
V2 normen en waarden	,507	,587
v3 begrip wereld	,622	,422
v4 weten wat te doen	,711	,181
v5 mijzelf begrijpen	,786	,062
v6 anderen vermaken	,709	-,485

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

We kunnen ook dit probleem gelukkig door SPSS laten oplossen.

De twee componenten vormen een assenstelsel waar de zes variabelen mee samenhangen. Als we met dat assenstelsel gaan draaien (roteren), veranderen de factorladingen (de correlaties van de variabelen met de twee componenten). SPSS probeert een stand te vinden waarbij het aantal hoge factorladingen op een component wordt gemaximaliseerd en het aantal lage factorladingen wordt geminimaliseerd. Zo kan een beter onderscheid worden gemaakt tussen de twee componenten. Het resultaat vinden we in de geroteerde componentenmatrix (tabel 10.12). Het vinden van de optimale stand van het assenstelsel gebeurt volgens een 'trial and error'-procedure. Factorladingen worden steeds opnieuw berekend tot een optimale stand is bereikt (we noemen dit een iteratief proces). In dit geval vindt SPSS de optimale factorladingen na drie pogingen. Dit is onder de geroteerde componentenmatrix aangegeven ('Rotation converged in 3 iterations'). Er bestaan meerdere vormen van roteren. Wij gebruiken de methode *Varimaxrotatie*.

Tabel 10.12 Componentenmatrix na rotatie (SPSS-output)

Rotated Component Matrix^a

	Component	
	1	2
v1 vermaken	,002	,890
V2 normen en waarden	,770	-,097
v3 begrip wereld	,745	,104
v4 weten wat te doen	,649	,342
v5 mijzelf begrijpen	,626	,481
v6 anderen vermaken	,202	,835

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Nu is interpretatie van de componenten/factoren eenvoudiger. Vooral de tweede factor is direct duidelijk. De variabelen 'vermaken' en 'met anderen vermaken' hebben op deze component de hoogste factorladingen. Hier gaat het om afhankelijkheid van de media voor entertainment, voor het vermaak. De eerste factor zou surveillance kunnen worden genoemd, de afhankelijkheid van de media om te kunnen functioneren in je sociale omgeving. Dit interpreteren van de factoren blijft een subjectief proces en het is mogelijk dat een andere onderzoeker op basis van dezelfde componentenmatrix tot andere naamgevingen besluit. Na rotatie hebben we twee nieuwe componenten die niet meer overeenkomen met de initiële componenten en daardoor is ook de verklaarde variantie van de twee componenten veranderd. In tabel 10.10 zien we deze cijfers in de een-na-laatste kolom. Na rotatie verklaart component 1 (variabelen v2 tot en met v5) 33,3% van de variantie in de afzonderlijke items en component 2 (variabelen v1 en v6) 31,0%.

Nu we weten dat we twee valide schalen kunnen maken die beide een ander aspect van media-afhankelijkheid meten, moeten we ook tweemaal een betrouwbaarheidsanalyse uitvoeren, en tweemaal een schaal maken en beschrijven.

10.5 Gebruik en presentatie van de resultaten

Als je een factoranalyse in SPSS draait, kun je SPSS vragen de gevonden factoren als variabelen aan je databestand toe te voegen. Die 'factorvariabelen' kun je vervolgens gebruiken in je verdere analyses, maar je kunt ook besluiten de factoranalyse te gebruiken als rechtvaardiging om de variabelen die hoog op een factor laden, zelf op te tellen tot een nieuwe variabele. Welke keuze je ook maakt, het is wel noodzakelijk om op een correcte wijze verslag te doen van je factoranalyse. Naast de factoranalyse bespreek je de betrouwbaarheid van de schaal of schalen, het gemiddelde en de standaarddeviatie van de schalen, en de antwoordschaal om de waarden in perspectief te zetten. Op basis van de factoranalyse van de media-afhankelijkheidsvariabelen schrijf je bijvoorbeeld in je onderzoeksverslag:

'Media-afhankelijkheid' is gemeten door de respondenten te vragen in welke mate ze de media gebruiken om zichzelf te vermaken, meer te weten te komen over de normen en waarden in de maatschappij, meer begrip te krijgen van de wereld om hen heen, zichzelf beter te begrijpen, te weten wat ze in allerlei situaties het best kunnen doen en zich met anderen te vermaken. Uit de factoranalyse met Varimaxrotatie bleek dat er twee dimensies zijn te onderscheiden (eigenwaarde > 1), te weten 'surveillance' en 'entertainment'. Samen verklaren deze factoren 64,3% van de variantie. Beide schalen zijn redelijk betrouwbaar ('surveillance' $\alpha = 0,74$, 'entertainment' $\alpha = 0,71$) (tabel 10.13). Respondenten zijn redelijk afhankelijk van media om in hun sociale omgeving te functioneren ($M = 4,86$, $SD = 0,99$) en redelijk afhankelijk van de media voor entertainment ($M = 4,76$, $SD = 1,48$), beide gemeten op een schaal van 1 tot en met 7.

Tabel 10.13 Factoranalyse (met rotatie) media-afhankelijkheid

Door de media te gebruiken ...	Surveillance	Entertainment
– kan ik mezelf goed vermaken		0,89
– kom ik meer te weten over normen en waarden in de maatschappij	0,77	
– begrijp ik meer van de wereld om mij heen	0,75	
– ga ik mijzelf beter begrijpen	0,65	0,48
– weet ik wat ik in allerlei situaties het best kan doen	0,63	
– kan ik me met anderen vermaken		0,84
Eigenwaarde	2,00	1,86
Verklaarde variantie	33,3%	31,0%
Cronbachs alfa	0,74	0,71

Factorloadingen na varimaxrotatie; factorloadingen $< |0,45|$ zijn niet weergegeven.⁶

10.6 Overige vormen van betrouwbaarheid

In voorgaande paragrafen hebben we naar de interne consistentie gekeken om de betrouwbaarheid van een schaal te bepalen. We willen tot slot nog stilstaan bij andere vormen van betrouwbaarheid. Het uitrekenen van statistieken op basis van de variabelen in je databestand is alleen maar zinvol als je waarnemingen op een betrouwbare wijze zijn gemeten. De betrouwbaarheid van een meting is de mate waarin die meting vrij is van toevallige fouten. Dat betekent dat als je een meting nog een keer uitvoert op dezelfde manier bij dezelfde onderzoekseenheden, je dezelfde waarde moet vinden. Als dat niet het geval is, is je meting niet betrouwbaar.

De manier waarop je een kenmerk van je onderzoekseenheden meet, verdient daarom de nodige aandacht. Sommige kenmerken zijn gemakkelijk te meten, zoals geslacht en leeftijd. De kans op toevallige fouten is voor deze kenmerken niet groot, maar ook niet uitgesloten. Andere kenmerken zijn minder eenvoudig te meten, bijvoorbeeld een concept als 'vertrouwen in de media' of 'media-afhankelijkheid'. De kans op toevallige fouten wordt dan groter.

De betrouwbaarheid van metingen kunnen we met behulp van verschillende procedures en statistieken vaststellen. De te onderscheiden procedures leiden tot het vaststellen van verschillende aspecten van betrouwbaarheid, namelijk betrouwbaarheid in de zin van *stabiliteit*, *equivalentie* en *interne consistentie* die we al eerder in dit hoofdstuk hebben we besproken. Om na te gaan of de metingen *intern consistent* zijn gebruiken we een schaalanalyse. In deze paragraaf gaan we dieper in op de stabiliteit en equivalentie van metingen. Of een meting stabiel is, stel je vast door herhaalde metingen, de test-hertestprocedure.

Je herhaalt dan je meting op dezelfde manier. Als we eenzelfde verschijnsel op een andere manier kunnen meten (parallele tests), kunnen we nagaan of de resultaten van die metingen tot dezelfde (*equivalente*) resultaten leiden. Als herhaalde metingen of parallele metingen tot dezelfde resultaten leiden, is de meting betrouwbaar. Voor het testen van stabiliteit en equivalentie kunnen we voor een deel de in de vorige hoofdstukken besproken associatiematen gebruiken, maar daarnaast worden in dit hoofdstuk ook andere statistieken genoemd.

10.6.1 Stabiliteit en equivalentie

Als je metingen betrouwbaar zijn, zou je elke keer dat je dezelfde meting bij dezelfde onderzoekseenheden opnieuw uitvoert dezelfde waarden moeten vinden. Er is dan sprake van *stabiliteit*. Als het mogelijk is eenzelfde meting te herhalen, is dat een test-hertestbetrouwbaarheidsmeting. Wanneer je een vraag in een vragenlijst twee keer aan dezelfde respondenten voorlegt, krijg je in je databestand twee variabelen waartussen je de samenhang kunt berekenen. Aangezien beide metingen hetzelfde verschijnsel meten, zou het verband, de samenhang, zeer sterk/perfect moeten zijn. Om dat vast te stellen kun je de eerder besproken associatiematen gebruiken. De waarde die je dan vindt, is een indicatie voor de betrouwbaarheid van je metingen. Je interpreteert de waarden van de associatiematen dan wel anders dan wanneer we de samenhang tussen twee verschillende variabelen berekenen. Als je twee keer de respondenten naar hun inkomen vraagt en je daarna de correlatiecoëfficiënt berekent (omdat het hier om variabelen op rationiveau gaat), ben je niet erg tevreden met een r van 0,80. Als er helemaal geen toevallige fouten in je metingen zitten, zouden de twee variabelen identiek moeten zijn. Je verwacht dan dat r gelijk is aan 1.

Het herhalen van een meting is niet altijd mogelijk. Respondenten zullen aardig geïrriteerd raken als je vragen twee keer stelt. Bij het coderen van categorieën voor een inhoudsanalyse en bij observaties is dat gemakkelijker te verwezenlijken. Voor het vaststellen van de betrouwbaarheid van een meting kun je de meting door dezelfde codeur of observator laten herhalen (intracodeur- of intraobservatorbetrouwbaarheid) of een tweede codeur of observator inschakelen die dezelfde onderzoekseenheden codeert of observeert (intercodeur- of interobservatorbetrouwbaarheid). De eerste vorm waarborgt de stabiliteit van de meting, de tweede vorm de *equivalentie* van een meting. Equivalentie betekent dat bij dezelfde onderzoekseenheden op (minimaal) twee manieren hetzelfde verschijnsel wordt gemeten. Wanneer de resultaten (nagenoeg) identiek zijn, is er sprake van een betrouwbare meting. In een vragenlijst zou je dat kunnen doen door een parallele test, waarbij je op een andere manier dezelfde informatie bij dezelfde onderzoekseenheden meet. Dit kun je doen door bijvoorbeeld niet alleen de leeftijd in jaren te vragen, maar ook het geboortjaar. Een hoge correlatie tussen die twee variabelen duidt op betrouwbare metingen.

De meeste simpele manier voor het berekenen van de inter- of intracodeurbetrouwbaarheid bij een inhoudsanalyse is het berekenen van het *overeenstemmingspercentage*. Het overeenstemmingspercentage is een heel grove maat die aangeeft welk percentage van de coderingen door twee codeurs identiek zijn uitgevoerd. Je berekent dit overeenstemmingspercentage op basis van een kruistabel. In deze kruistabel worden de codering door codeur 1 en de codering door codeur 2 van eenzelfde variabele tegen elkaar afgezet.

In tabel 10.14 staan de coderingen van de gevolgen van klimaatverandering (1 = geen, 2 = droogte, 3 = overstromingen) die in nieuwsberichten genoemd zijn. Er zijn tien nieuwsberichten (de onderzoekseenheden) die door zowel codeur 1 als door codeur 2 zijn gecodeerd. De coderingen van codeur 1 staan in de kolommen en de coderingen van codeur 2 staan in de rijen van de kruistabel. In de diagonaal van deze tabel vinden we de berichten die van beide codeurs dezelfde waarde hebben gekregen. Van drie berichten vinden ze allebei dat er 'geen gevolgen' in worden genoemd. In drie berichten hebben ze allebei geconstateerd dat hierin 'de droogte' is vermeld en van twee berichten vinden ze beide dat 'de overstromingen' genoemd worden. Die $3 + 3 + 2 = 8$ berichten hebben ze identiek gecodeerd. Over de overige twee gecodeerde berichten verschillen ze van mening. Het overeenstemmingspercentage is $8 \div 10 = 80\%$. Met 80% overeenstemming zul je als onderzoeker over het algemeen niet zo heel erg tevreden zijn. Het betekent immers ook dat de ene codeur in 20% van de gevallen de operationalisatie van de begrippen anders heeft opgevat dan de andere codeur. Met een klein aantal onderzoekseenheden zoals in ons voorbeeld, hebben twee fouten een groter effect op het overeenstemmingspercentage dan wanneer je bijvoorbeeld op honderd artikelen twee keer geen overeenstemming bereikt. Deze betrouwbaarheidsstatistiek gebruik je dan ook meestal als proefcoderingen om je codeersysteem nog te kunnen verbeteren.

Tabel 10.14 Coderingen gevolgen klimaatverandering door twee codeurs:

Codeur 1 \ Codeur 2	(1) Geen gevolgen	(2) Droogte	(3) Overstromingen	Totaal
(1) Geen gevolgen	3	0	0	3
(2) Droogte	1	3	1	5
(3) Overstromingen	0	0	2	2
Totaal	4	3	3	10

Verschillende statistische maten die gebruikt worden bij inhoudsanalyse zijn gebaseerd op de mate van overeenstemming, zoals Scotts pi, Krippendorfs alfa en Cohens kappa. Deze maten houden ook nog rekening met de toevalskans. We zullen deze maten niet bespreken in dit boek. Associatiematen als Kendalls tau-b en Spearmans rho (ordinaal) en Pearsons productmomentcorrelatie r (interval, ratio) kunnen ook worden gebruikt voor het vaststellen van de stabiliteit of equivalentie van een meting.

10.7 Samenvatting

In dit hoofdstuk hebben we laten zien welke stappen genomen moeten worden in het construeren van een schaal. Daartoe ga je na of je valide en betrouwbare metingen hebt uitgevoerd. We hebben laten zien hoe je door middel van een factoranalyse latente variabelen kunt vinden die ten grondslag liggen aan de (manifeste) variabelen die je hebt gemeten. Je kunt factoranalyse gebruiken om iets te zeggen over de validiteit van je metingen.

Als de factoranalyse aantoont dat je manifeste variabelen unidimensioneel zijn en dus hetzelfde verschijnsel meten, rechtvaardigt dit het samenvoegen van de manifeste variabelen tot één schaalvariabele voor het te meten verschijnsel. Op die manier kan een groot aantal variabelen worden teruggebracht tot hanteerbare hoeveelheden (datareductie), waardoor de analyses voor je onderzoek vereenvoudigd worden.

Daarnaast moet de meting betrouwbaar zijn. Betrouwbaarheid is de mate waarin je waarnemingen vrij zijn van toevallige fouten. Daarbij onderscheid je drie aspecten: stabiliteit, equivalentie en interne consistentie.

Bij een test-hertest kun je op basis van associatiematen iets zeggen over de stabiliteit van je metingen. Een parallelle test geeft informatie over de equivalentie. De interne consistentie van een schaal waarvoor je meerdere items gebruikt, stel je vast door Cronbachs alfa uit te rekenen.



Ga naar de website om de opdrachten bij dit hoofdstuk te maken.

Noten

- 1 Er zijn veel soorten validiteit, die we niet allemaal zullen bespreken in dit boek. Bij het uitvoeren van een factoranalyse om na te gaan of je een valide meting hebt uitgevoerd, gaat het in de regel om *construct- of begripsvaliditeit*, waarbij je nagaat of de items die je hebt gemeten inderdaad een indicator zijn van datgene wat je beoogde te meten.
- 2 Er zijn manieren om schalen te maken met wegingen van scores bij variabelen die niet op dezelfde manier gemeten zijn. Deze zullen wij echter niet in dit boek behandelen.
- 3 Dit voorbeeld is gebaseerd op het volgende onderzoek: Kwon, M., Kim, D.J., Cho, H. & Yang, S. (2013). *The Smartphone Addiction Scale: Development and Validation of a Short Version for Adolescents*. PLoS ONE 8(12): e83558. doi:10.1371/journal.pone.0083558.
- 4 Net als bij de richtlijnen bij de interpretatie van een associatiemaat is dit geen universele regel.
- 5 Het onderscheiden van meerdere componenten wordt besproken in de volgende paragraaf.
- 6 Bij principale componentenanalyse zouden we consequent de term component moeten gebruiken, maar omdat de interpretatie van factoren en componenten identiek is, worden in dit hoofdstuk beide termen gebruikt.