

— INLEIDING

1.1 VOORPROEFJE: DE METHODE VAN THURSTONE

Psychofysica

Al lang voordat de psychologie in 1879¹ tot officiële wetenschap verheven werd, waren er diverse methoden ontwikkeld om de intensiteit te bepalen van psychische gewaarwordingen die door fysieke stimuli teweeg worden gebracht. In deze *psychofysische methoden* (de term stamt uit het in 1860 verschenen boek van Fechner: *Elemente der Psychophysik*²) gaat het erom het functionele verband te beschrijven tussen de fysieke eigenschappen van stimuli (bijvoorbeeld gewicht, toonhoogte, geluidssterkte) en de intensiteit waarmee die stimuli beleefd worden. De functie die het verband tussen de fysieke en psychische grootheden laat zien wordt de psychometrische kromme of psychofysische functie genoemd. Is deze functie eenmaal vastgesteld, dan kan men voor alle mogelijke stimuli de bijbehorende psychische intensiteiten uitrekenen. Op die manier krijgt men waarden die de betreffende stimuli op een psychologische *schaal*³ afbeelden: zogenaamde (psychologische) *schaalwaarden*. Deze schaalwaarden zijn dus *kwantificaties* van de psychische gewaarwordingen.

1 In dat jaar stichtte Wilhelm Wundt in Leipzig het eerste officiële psychologische laboratorium ter wereld.

2 In Fechners eigen woorden: 'Unter Psychophysik verstehe ich ... eine exacte Lehre von den Beziehungen zwischen Leib und Seele' (1860, Deel 1, p. v).

3 In Van Dales *Groot Woordenboek der Nederlandse Taal* (1992) wordt het woord schaal omschreven als '... maatstaf om de uitkomst van metingen ... af te lezen: (bijvoorbeeld) *de schaal van Fahrenheit*' en als 'reeks van getallen of verhoudingen voor het vaststellen van prijzen, lonen of andere bedragen; - reeks van gelijkmatig opklimmende getallen die de grondslag van een talstelsel vormen: *de getallen van 1-10 vormen de schaal van het tientallig stelsel*'.

Een beroemd voorbeeld van een psychofysische functie is de wet van Weber-Fechner. Deze ziet er als volgt uit

$$I = c \cdot \log \frac{S}{S_0} \quad [1.1]$$

en luidt in woorden dat I , de intensiteit van de sensatie, evenredig is met de logaritme van S/S_0 . Daarbij is S de fysieke waarde van de aangeboden stimulus en is S_0 de drempelwaarde die bij het desbetreffende soort stimuli hoort. De drempelwaarde is de waarde die een stimulus moet hebben om in vijftig procent van de gevallen een sensatie op te wekken, dat wil zeggen: bij de helft van de proefpersonen of bij de helft van het aantal keren dat de stimulus aan één proefpersoon wordt aangeboden. De evenredigheidsfactor c is eveneens afhankelijk van het type stimuli.

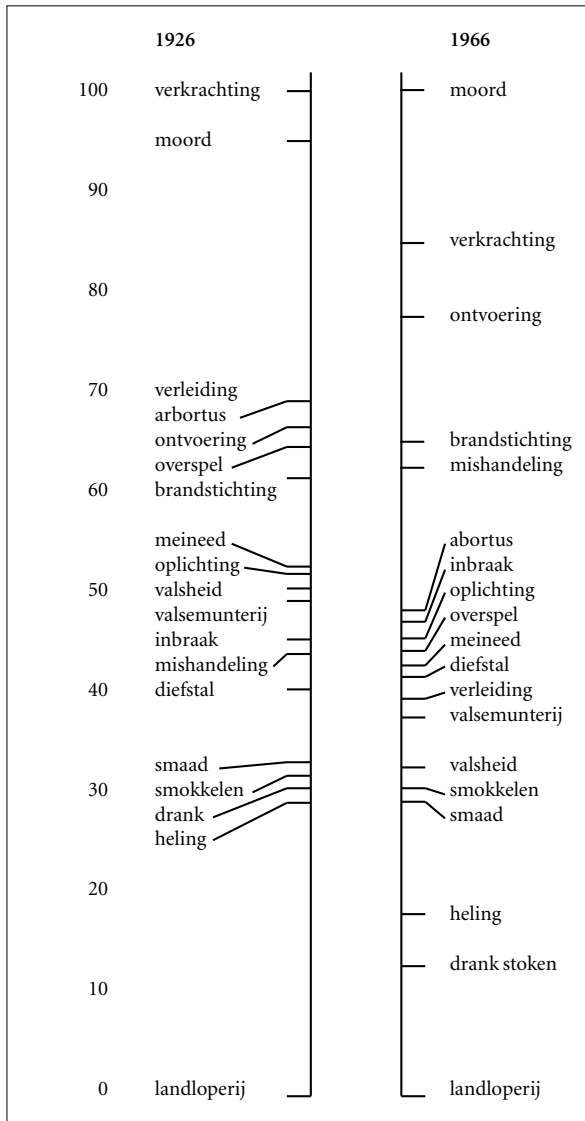
In de psychofysica is een veelheid van vernuftige methoden bedacht om de drempelwaarden van stimuli en de vorm van psychofysische krommen te bepalen (zie bijvoorbeeld Guilford, 1954; Torgerson, 1958; Gescheider, 1985). In al deze methoden komt het erop neer dat men proefpersonen vraagt een aantal stimuli met een of meer standaardstimuli te vergelijken. Deze aanpak werd door Thurstone uitgebreid tot de *methode van paarsgewijze vergelijkingen* waarin alle stimuli met alle andere vergeleken moeten worden. Door een bepaald statistisch model aan te nemen kon Thurstone uit de antwoorden van de proefpersonen psychologische schaalwaarden van de stimuli afleiden zonder dat er fysieke eigenschappen van de stimuli bekend hoefden te zijn. Bij deze methode wordt in de volgende paragraaf uitgebreid stilgestaan.

Thurstones methode

Zeventig jaar geleden publiceerde Thurstone (1927a) zijn *Law of comparative judgment*. Met behulp van deze wet kan men psychologische schaalwaarden van stimuli afleiden, als men geobserveerd heeft hoe vaak de ene stimulus een andere stimulus met betrekking tot een van tevoren vastgesteld criterium (bijvoorbeeld grootte, luidheid, toonhoogte, schoonheid, wenselijkheid) overtreft. Om de benodigde observaties te verkrijgen, laat men een aantal proefpersonen *alle* stimuli die men wil onderzoeken paarsgewijs met elkaar vergelijken. Het revolutionaire van Thurstones aanpak was nu dat men psychologische schaalwaarden kon afleiden zonder over fysieke intensiteiten van de stimuli te beschikken. Daarmee was het mogelijk geworden om allerlei stimuli, ook abstracte psychologische begrippen, met betrekking tot ieder willekeurig criterium te kwantificeren. Een gedetailleerde beschrijving van Thurstones *Law of comparative judgment* en de methode van paarsgewijze vergelijkingen wordt gegeven in Blok 1.1.

Een mooi voorbeeld van een toepassing van bovengenoemde methode is het onderzoek naar sociale waarden dat Thurstone (1927b) in 1926 uitvoerde met 266 studenten als proefpersonen. Aan deze proefpersonen werden alle 171 paren voorgelegd die gevormd konden worden uit 19 verschillende misdrijven:

abortus, overspel, brandstichting, mishandeling, illegaal drank stoken, inbraak, valsemunterij, oplichting, valsheid in geschrifte, moord, ontvoering, diefstal, smaad, meineed, verkrachting, heling, verleiding, smokkelen, landloperij. Van elk paar stimuli moesten de proefpersonen aangeven welke misdaad zij het ernstigst vonden. Toepassing van de *Law of comparative judgment* leidde tot schaalwaarden die de ernst van de misdrijven weergaven.



Figuur 1.1 Schaalwaarden uit 1926 en 1966 van negentien misdrijven (ontleend aan Coombs e.a., 1970, p. 47)

Hetzelfde onderzoek is veertig jaar later door Coombs (1967) herhaald met 311 studenten als proefpersonen. Figuur 1.1, waarin de resultaten van beide studies grafisch zijn weergegeven, laat enkele interessante maatschappelijke ontwikkelingen zien. Zo is in 1927 verkrachting erger dan moord en zijn abortus, overspel en verleiding even erg als brandstichting en ontvoering. In 1966 is er nogal wat veranderd: er is met name aan de bovenkant een veel duidelijkere gradatie ontstaan. De schaal loopt nu van moord via verkrachting, ontvoering, brandstichting en mishandeling naar een groep misdrijven die minder ernstig gevonden worden. Opgemerkt moet worden dat de gevonden schaalwaarden relatief zijn. We weten alleen dat men heling en landloperij als (veel) minder ernstig beoordeelde dan de andere misdrijven; we weten niet of men heling absoluut gezien als een ernstig misdrijf beschouwde of niet. Het zou interessant zijn om te onderzoeken of de gevonden schaalwaarden correleren met de (minimum of maximum) straffen voor deze misdrijven in het Wetboek van Strafrecht of met de gemiddelde straffen waartoe plegers van deze delicten worden veroordeeld.

Een belangrijk kenmerk van Thurstones schaalmethode is dat de onderzoeker van tevoren een criterium moet vaststellen waarop de stimuli met elkaar vergeleken worden. In Figuur 1.1 staan de ordeningen van de misdrijven in termen van ernst; men had echter ook de misdrijven kunnen laten beoordelen in termen van wreedheid, schadelijkheid, angstaanjagendheid of frequentie van voorkomen. We moeten ons realiseren dat de schaalwaarden van de delicten er bij verschillende beoordelingscriteria totaal anders uit zouden kunnen zien. Ook is het mogelijk dat het door een onderzoeker gekozen criterium niet door alle proefpersonen even relevant gevonden wordt.

Omdat we niet altijd van tevoren kunnen weten welke criteria voor welke stimuli en proefpersonen relevant zijn, zou het prettig zijn als we over een methode beschikten waarmee het mogelijk is schaalwaarden voor stimuli te vinden zonder de proefpersonen een specifiek vergelijkingscriterium aan te moeten bieden. Gelukkig bestaan er zulke methoden. Een voorbeeld daarvan zullen we in de volgende paragraaf behandelen.

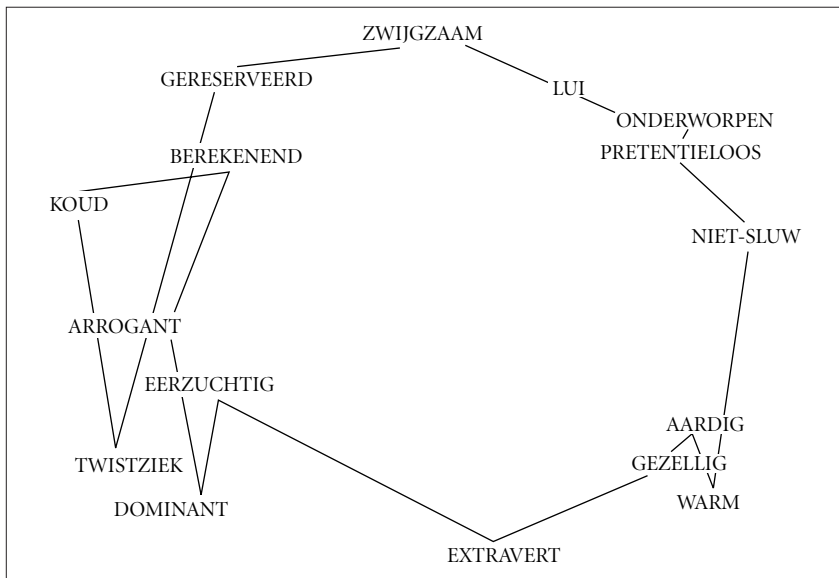
Schaalwaarden uit sorteergegevens

Met behulp van een methode die (waarschijnlijk) door Miller (1967) voor het eerst is toegepast, lieten Van der Kloot en Van Herk (1991) 25 proefpersonen een zogenaamde sorteertaak of (*card-*)*sorting task* verrichten. De stimuli bestonden uit vijftig adjectieven die persoonlijkheidseigenschappen (bijvoorbeeld lui, arrogant, warm) aanduidden. Deze adjectieven waren elk afzonderlijk op een systeemkaartje geschreven. Iedere proefpersoon kreeg een van tevoren geschud pakje kaartjes en hem of haar werd gevraagd stapeltjes te maken van adjectieven die naar zijn of haar mening 'bij elkaar hoorden'. Iedere proefpersoon mocht net zoveel stapeltjes maken als hij of zij wilde, en elk stapeltje mocht net zoveel woorden bevatten als de proefpersoon nodig vond. Na afloop werd genoteerd welke stimuli bij elkaar waren gelegd en werd voor ieder

paar stimuli geteld hoe vaak (dat wil zeggen: door hoeveel proefpersonen) zij samen in één groepje waren ingedeeld.

Het idee is nu dat twee stimuli ‘meer bij elkaar horen’, ‘dichter bij elkaar staan’, of ‘meer op elkaar lijken’ naarmate zij vaker samen in één stapeltje voorkomen. De frequentie van samengaan is dus een maat voor gelijkenis of nabijheid (*proximity*) van de adjectieven. Deze nabijheidsgegevens kunnen we gebruiken om de adjectieven in een grafiek af te beelden. In zo’n grafiek wordt ieder adjectief dan weergegeven door middel van een punt. Dat gebeurt zodanig dat de afstanden tussen de punten in de grafiek overeenkomen met de geobserveerde proximiteiten. Dat wil zeggen, als twee punten een grote proximity hebben, liggen ze dichter bij elkaar in de grafiek dan wanneer ze een kleine proximity hebben. Dit is het principe van de *meerdimensionale schaalmethoden*.

Van zestien (van de vijftig) stimuli zijn de desbetreffende gelijkenismaten op bovengenoemde manier geanalyseerd: voor ieder adjectief werden er twee afzonderlijke schaalwaarden verkregen. Deze schaalwaarden kunnen gebruikt worden als coördinaten van de adjectieven op de twee assen (de dimensies) van een tweedimensionale grafiek. De zestien geanalyseerde adjectieven zijn op deze manier weergegeven in Figuur 1.2. Omdat op grond van eerder onderzoek (Wiggins, 1979; Van der Kloot & Slooff, 1985) verwacht werd dat deze zestien adjectieven op een cirkel zouden liggen, zijn de adjectieven in de verwachte volgorde met elkaar verbonden. Geconcludeerd mag worden dat de gevonden structuur in redelijke mate met de verwachte volgorde overeenkomt. Alleen de negatieve eigenschappen koud, berekenend, arrogant en ambitieus liggen een beetje rommelig door elkaar.



Figuur 1.2 De tweedimensionale configuratie van zestien persoonlijkheidsadjectieven (ontleend aan Van der Kloot & Van Herk, 1991, p. 569)

Zoals gezegd, hebben we voor iedere stimulus hier twee schaalwaarden, namelijk een coördinaat op de horizontale en een coördinaat op de verticale as. Echter, wat deze dimensies inhoudelijk betekenen is niet zonder meer gegeven.

Om de dimensies te *interpreteren* moeten we kijken naar de ordening van de stimuli op de dimensies. Met name moeten we kijken naar welke stimuli extreme posities op de dimensies innemen. In de configuratie van Figuur 1.2 loopt de horizontale dimensie van koud en berekenend naar niet-sluw, terwijl de verticale dimensie van extravert naar zwijgzaam gaat. In de eerste dimensie lijkt het dus om de tegenstelling betrouwbaar-onbetrouwbaar te gaan; de tweede dimensie geeft de tegenstelling introvert-extravert weer.

Bovenstaande interpretatie is overigens niet de enig mogelijke. Als we de figuur een beetje zouden draaien (later zullen we zien dat dit geoorloofd is!) kunnen we een horizontale dimensie vinden die van gereserveerd naar warm loopt en een verticale as van dominant naar onderworpen. Deze twee dimensies, sociabiliteit en dominantie, zijn dimensies die vaker in dit soort onderzoek gevonden worden.

BLOK 1.1 THURSTONES LAW OF COMPARATIVE JUDGMENT

Thurstones uit 1927 daterende *Law of comparative judgment* is een verzameling van vijf modellen voor het bepalen van schaalwaarden uit data die verzameld zijn met de methode van paarsgewijze vergelijkingen (PGV). Zoals gezegd, bestaat de PGV-methode hieruit dat men de proefpersonen alle uit een verzameling stimuli te vormen paren aanbiedt met de opdracht voor elk paar aan te geven welke van de twee stimuli in een bepaald opzicht ‘geprefereerd’ wordt. Het kan hierbij letterlijk om voorkeur gaan, zoals bij de vraag ‘welke van deze twee groenten vind je het lekkerst?’ maar men kan ook andere vragen stellen, bijvoorbeeld ‘welke van deze twee misdrijven vind je het ergst?’ Formeel gezien geeft een proefpersoon in dit tweede voorbeeld aan welk delict de voorkeur heeft als er op de ernst van het misdrijf gelet moet worden. Daarom worden dit soort gegevens *voorkeursdata* genoemd.

Om Thurstones methode te kunnen toepassen, moeten we beschikken over voorkeursuitspraken over *alle* stimulusparen; bovendien moet er voor ieder stimuluspaar een *groot aantal* voorkeursuitspraken verzameld zijn. Dat kan door een en dezelfde proefpersoon ieder stimuluspaar een (groot) aantal keren te laten beoordelen, of door een (groot) aantal proefpersonen ieder stimuluspaar één keer te laten beoordelen. Het uitgangspunt is dat we per stimuluspaar N beoordelingen krijgen die als onderling verwisselbare *replicaties* beschouwd mogen worden. Voor ieder stimuluspaar (i, j) krijgen we dus een frequentie f_{ij} die aangeeft hoe vaak stimulus i boven j verkozen wordt. Delen we deze frequenties door N , dan krijgen we

proporties $p_{ij} = f_{ij}/N$ als schattingen van de populatiekansen dat stimulus i de voorkeur boven j krijgt. In het model van Thurstone heeft deze kans een relatie met de schaalwaarden van stimuli i en j op de dimensie waarop ze met elkaar vergeleken worden. Stel nu dat de stimuli i en j respectievelijk schaalwaarden u_i en u_j hebben. In Thurstones aanpak moeten de kansen p_{ij} opgevat worden als een functie van het *verschil tussen de schaalwaarden* u_i en u_j van de stimuli i en j . Voor die functie moet in het algemeen het volgende gelden:

- 1 als $u_i - u_j = 0$, dan is $p_{ij} = .50$;
- 2 als $u_i - u_j > 0$, dan is $p_{ij} > .50$;
- 3 als $u_i - u_j < 0$, dan is $p_{ij} < .50$.

Kortom, de relatie tussen p_{ij} en $(u_i - u_j)$ zal *monotoon stijgend* moeten zijn, met $p_{ij} = 1.0$ voor extreem grote waarden van $(u_i - u_j)$ en $p_{ij} = 0.0$ voor extreem kleine (hier: negatieve) waarden.

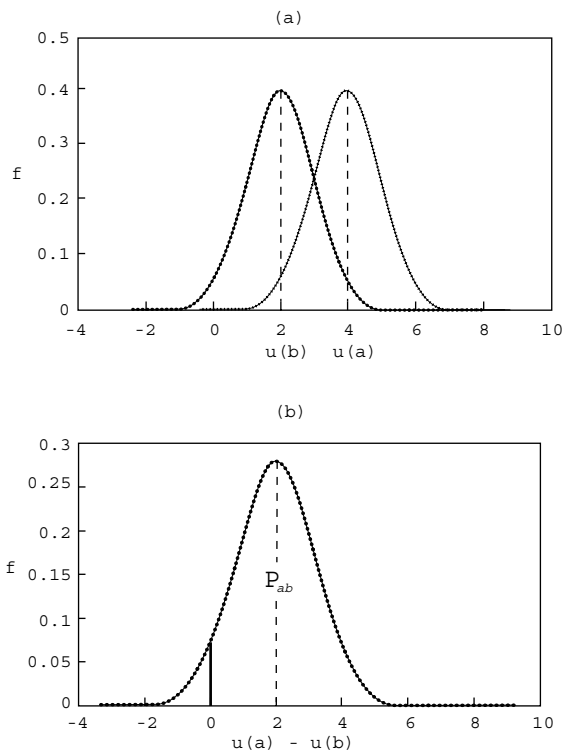
Thurstones redenering was nu als volgt. Bij iedere paarsgewijze vergelijking en iedere replicatie worden als het ware steeds twee ‘impliciete’ of ‘waargenomen’ schaalwaarden met elkaar vergeleken. Die impliciete schaalwaarden zijn niet steeds hetzelfde, maar fluctueren op toevallige wijze rond hun respectievelijke echte waarden u_i en u_j , waarbij zij een *normaalverdeling* volgen met respectievelijke standaarddeviaties s_i en s_j . Uit de statistiek is bekend dat het *waargenomen verschil* tussen de twee schaalwaarden dan eveneens volgens toeval varieert, namelijk volgens een normale verdeling met gemiddelde $(u_i - u_j)$ en standaarddeviatie $s_{ij} = (s_i^2 + s_j^2 - 2r_{ij}s_is_j)^{.5}$. In deze formule is r_{ij} de correlatie tussen de fluctuaties van de waargenomen schaalwaarden rond u_i en u_j . Berekenen we nu

$$z_{ij} = \frac{u_i - u_j}{\sqrt{s_i^2 + s_j^2 - 2r_{ij}s_is_j}} \quad [1.2]$$

dan geeft z_{ij} aan hoeveel standaarddeviaties s_{ij} de waarden u_i en u_j van elkaar verschillen. De kans p_{ij} is dan gelijk aan de cumulatieve proportie onder de normaalverdeling die bij de standardscore z_{ij} hoort. De score z_{ij} kan dus direct gevonden worden uit de cumulatieve-normaalverdelingstabel, die in ieder willekeurig statistiekboek staat.

Thurstones model is grafisch weergegeven in Figuur 1.3a en 1.3b, waarin de horizontale lijn het desbetreffende beoordelingscriterium vertegenwoordigt. De stimuli A en B hebben respectievelijk de gemiddelde schaalwaarden u_a en u_b op dit continuüm; hun waargenomen schaalwaarden fluctueren voor verschillende proefpersonen en voor verschillende aanbiedingen van de stimuli rond deze gemiddelden. Deze fluctuaties zijn beide normaal verdeeld met standaarddeviaties gelijk aan 1 (zie Figuur 1.3a). Als

op verschillende momenten verschillende personen een waarneming uit de ene verdeling met een waarneming uit de andere vergelijken, dan zijn de verschillen tussen de waargenomen schaalwaarden normaal verdeeld met standaarddeviatie $(1 + 1)^{.5} = 1.4$ en gemiddelde $(u_a - u_b)$. Deze verdeling is afgebeeld in Figuur 1.3b. We zien dat de kans p_{ab} dat het waargenomen verschil $u_a - u_b$ groter dan nul is, gelijk is aan de oppervlakte van het gebied rechts van de waarde nul gedeeld door de totale oppervlakte onder de normaalverdeling met gemiddelde $(u_a - u_b)$. Als we zouden weten hoeveel standaarddeviaties het gemiddelde $(u_a - u_b)$ van nul afligt, dan zouden we p_{ab} kunnen aflezen uit de tabel van de cumulatieve-normalverdeling. Omgekeerd: als we de waarde van p_{ab} kennen, dan kunnen we uit diezelfde tabel bepalen hoeveel standaarddeviaties $(u_a - u_b)$ boven nul ligt. Dit laatste is het principe waar Thurstones methode op berust. De frequenties f_{ij} worden omgezet in kansen p_{ij} , die gebruikt worden om schattingen (in standaardscores) van de verschillen $u_i - u_j$ te verkrijgen. Uit deze schattingen kunnen ten slotte schattingen voor de afzonderlijke schaalwaarden u_i en u_j worden afgeleid.



Figuur 1.3 Normaalverdelingen van (a) de waargenomen schaalwaarden van twee stimuli en (b) de waargenomen verschillen rond hun gemiddelden

Thurstone heeft verschillende speciale gevallen van Formule 1.2 geformuleerd. Het eenvoudigste is *Case V* waarin wordt aangenomen dat $s_i = s_j = s$ en dat $r_{ij} = 0$. Onder deze aannamen is

$$z_{ij} = \frac{u_i - u_j}{\sqrt{2s^2}} = c(u_i - u_j) \quad [1.3]$$

zodat de bij p_{ij} behorende standaardscore z_{ij} evenredig is met $(u_i - u_j)$; c is hierbij een multiplicatieve constante, een schaafactor die verder niet belangrijk is.

Op basis van dit model kunnen we de proporties p_{ij} omzetten in standaardscores z_{ij} . Noteren we deze in een tabel met m rijen en m kolommen (met indices $i = 1, \dots, m$ en $j = 1, \dots, m$; m is het aantal stimuli) dan is idealiter $\sum_j z_{ij}/m = \sum_j c(u_i - u_j)/m = cu_i - c\sum_j u_j/m$. Elk rij-gemiddelde van deze tabel is dus een schatting van de bijbehorende u -waarde. Voor deze schattingen geldt dat de som van de geschatte u -waarden gelijk aan nul is. Omdat de waarde nul voor het gemiddelde niet meer dan een arbitraire keuze is, hebben we nu schaalwaarden op *intervalniveau*.

Als rekenvoorbeeld van Thurstones Case V-methode presenteren we hier de PGV-gegevens uit een studie over het belang van vijf soorten bijdragen die onderzoekers kunnen leveren aan (publicaties van) wetenschappelijk onderzoek (Van der Kloot & Willemsen, 1991). Deze bijdragen waren: bedenken van het onderzoek, leiding eraan geven, dataverzameling, data-analyse en schrijven van de publicatie. Na paarsgewijze vergelijking van deze bijdragen beoordeelden de proefpersonen (zelf onderzoekers!) welke bijdragen zij belangrijk genoeg vonden om recht te geven op (co)auteurschap van een publicatie. Formeel gezien vergelijken de proefpersonen iedere bijdrage dus ook met de auteurschapsgrens (AG), de drempelwaarde die overschreden moet worden om (co)auteur te mogen zijn. De betreffende f_{ij} -, p_{ij} - en z_{ij} -waarden staan in de matrices **F**, **P** en **Z** die zijn weergegeven in de Tabellen 1.1, 1.2 en 1.3.

Tabel 1.1 De matrix **F** met de frequenties waarmee de rij-stimuli verkozen werden boven de kolomstimuli

	BE	LE	DV	DA	SC	AG
bedenken (BE)	–	21	31.5	29.5	5	31
leidinggeven (LE)	19	–	31	24.5	3.5	27.5
dataverzameling (DV)	8.5	9	–	15	2.5	11.5
data-analyse (DA)	10.5	15.5	25	–	1.5	25.5
schrijven (SC)	35	36.5	37.5	38.5	–	40
auteurschap (AG)	9	12.5	28.5	14.5	0	–

De waarden eindigend op .5 zijn het gevolg van het feit dat sommige proefpersonen geen keuze hadden gemaakt. Deze proefpersonen zijn voor de ene helft toegekend aan f_{ij} , voor de andere helft aan f_{ji} .

Tabel 1.2 De matrix **P** met de proporties waarmee de rij-stimuli verkozen werden boven de kolomstimuli

	BE	LE	DV	DA	SC	AG
bedenken (BE)	–	.525	.788	.738	.125	.775
leidinggeven (LE)	.475	–	.775	.613	.088	.688
dataverzameling (DV)	.213	.225	–	.375	.063	.288
data-analyse (DA)	.263	.388	.625	–	.038	.638
schrijven (SC)	.875	.913	.938	.963	–	.988
auteurschap (AG)	.225	.313	.713	.363	.012	–

De vetgedrukte proporties .988 (= 39.5/40) en .012 (= .5/40) zijn ingevuld in plaats van de waargenomen proporties 1.00 en 0. Dit is gedaan omdat anders de bijbehorende z-waarden $+\infty$ en $-\infty$ zouden worden.

Tabel 1.3 De matrix **Z** met de standaardscores voor de verschillen tussen de rij- en kolomstimuli

	BE	LE	DV	DA	SC	AG	$z_i/6$
bedenken (BE)	0	.063	.800	.636	-1.15	.755	.18
leidinggeven (LE)	-.063	0	.755	.287	-1.36	.489	.02
dataverzameling (DV)	-.800	-.755	0	-.321	-1.53	-.561	-.66
data-analyse (DA)	-.636	-.287	.321	0	-1.78	.352	-.34
schrijven (SC)	1.15	1.36	1.53	1.78	0	2.25	1.34
auteurschap (AG)	-.755	-.489	.561	-.352	-2.25	0	-.55

Uit de schaalwaarden in de meest rechtse kolom van Tabel 1.3 mogen we concluderen dat de belangrijkste bijdrage aan een wetenschappelijke publicatie het schrijven is, gevolgd door bedenken, leiding geven, data-analyse en dataverzameling. Alleen dataverzameling heeft een schaalwaarde die onder de auteurschapsgrens ligt. Iemand die alleen de dataverzameling verzorgt, zou volgens deze schaalwaarden dus geen co-auteur van een publicatie mogen zijn.

Een belangrijke vraag is nu of de verkregen schaalwaarden een goede beschrijving van de geobserveerde data geven, anders gezegd, of er een goede *fit* tussen de data en het model (Thurstone Case V) bestaat. Dat kunnen we op verschillende manieren proberen na te gaan. In de eerste plaats kunnen we voor de $m(m-1)/2$ verschillende paren stimuli de verschillen tussen de geschatte schaalwaarden berekenen. Dat levert ons dus schattingen van z_{ij} , die bij de geschatte schaalwaarden horen. Anders gezegd, we krijgen hierdoor ‘voorspellingen’ van z_{ij} , die met de geschatte parameters van het Case V-model corresponderen. Deze voorspellingen worden met het symbool \hat{z}_{ij} aangeduid. De geobserveerde z_{ij} -waarden kunnen we in een grafiek uitzetten tegen de voorspelde \hat{z}_{ij} . In het ideale geval zouden de punten van de grafiek dan op een rechte lijn moeten liggen. Een tweede mogelijkheid om de *fit* tussen data en model te bekijken is door uit de cumulatieve-normaalverdelingstabel de bij de \hat{z}_{ij} -waarden behorende voorspelde proporties \hat{p}_{ij} te bepalen. Ook de punten in de grafiek van p_{ij} tegen \hat{p}_{ij} zouden in het ideale geval een rechte lijn op moeten leveren. In de derde plaats kan men de voorspelde proporties vermenigvuldigen met N en de aldus voorspelde frequenties \hat{f}_{ij} vergelijken met de geobserveerde f_{ij} -waarden. Al deze methoden zijn natuurlijk equivalent, alleen zien afwijkingen van perfecte *fit* er in de ene methode getalmatig wat anders uit dan in de andere.

Hieronder staan de drie tabellen met \hat{Z} , \hat{P} en \hat{F} voor onze auteurschapsdata. De bijbehorende grafieken van Z tegen \hat{Z} , en P tegen \hat{P} , laten zien dat

de geobserveerde data slechts weinig afwijken van het model (immers: de punten liggen dicht bij de lijn $Y = X$).

Tabel 1.4 De matrix $\hat{\mathbf{Z}}$, met de uit de schaalwaarden terugberekende z-waarden

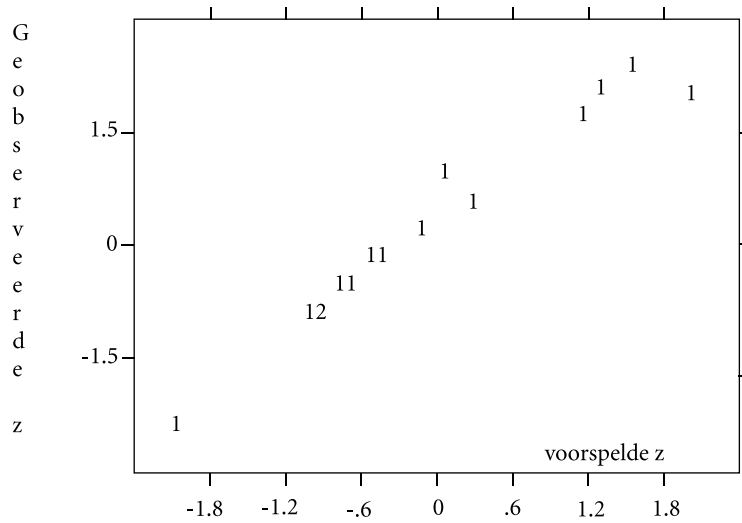
	BE	LE	DV	DA	SC	AG
bedenken (BE)	0	.16	.84	.52	-1.16	.73
leidinggeven (LE)	-.16	0	.68	.36	-1.32	.57
dataverzameling (DV)	-.84	-.68	0	-.32	-2.00	-.11
data-analyse (DA)	-.52	-.36	.32	0	-1.68	.21
schrijven (SC)	1.16	1.32	2.00	1.68	0	1.89
auteurschap (AG)	-.73	-.57	.11	-.21	-1.89	0

Tabel 1.5 De matrix $\hat{\mathbf{P}}$ met de uit de schaalwaarden terugberekende proporties

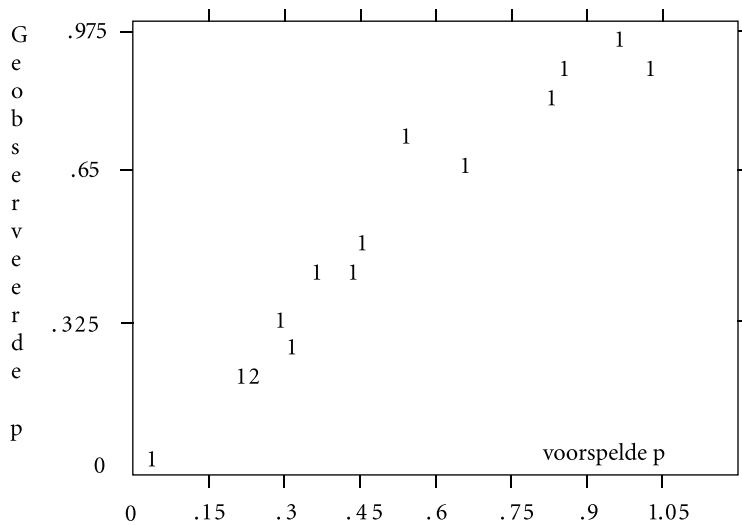
	BE	LE	DV	DA	SC	AG
bedenken (BE)	–	.564	.800	.698	.123	.767
leiding geven (LE)	.436	–	.752	.641	.094	.716
dataverzameling (DV)	.200	.248	–	.375	.023	.456
data-analyse (DA)	.302	.359	.625	–	.046	.583
schrijven (SC)	.877	.906	.977	.954	–	.971
auteurschap (AG)	.233	.284	.544	.417	.029	–

Tabel 1.6 De matrix $\hat{\mathbf{F}}$ met de uit de schaalwaarden terugberekende frequenties

	BE	LE	DV	DA	SC	AG
bedenken (BE)	–	22.6	32.0	27.9	4.9	30.7
leidinggeven (LE)	17.4	–	30.1	25.6	3.8	28.6
dataverzameling (DV)	8.0	9.9	–	15.0	.9	18.2
data-analyse (DA)	12.1	15.4	25	–	1.9	23.3
schrijven (SC)	35.1	36.2	39.1	38.1	–	38.8
auteurschap (AG)	9.3	11.4	21.8	16.7	1.2	–



Figuur 1.4 Oorspronkelijke z-waarden (verticale as) uitgezet tegen de uit de oplossing voorspelde waarden



Figuur 1.5 Oorspronkelijke proporties (verticale as) uitgezet tegen de uit de oplossing voorspelde proporties

De bovenstaande grafieken laten enerzijds zien dat er een grote mate van overeenkomst is tussen de voorspelde en de geobserveerde waarden van Z en P maar anderzijds dat er ook wel een paar afwijkingen zijn. Het zou dus handig zijn om via een statistische toets te kunnen beslissen of de geobserveerde data in overeenstemming zijn met het veronderstelde Case V-model. Zo'n toets is ontwikkeld door Mosteller (1951). De toetsingsgrootte is

$$\chi^2 = \frac{N}{821} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \left[\text{SIN}^{-1}\{\sqrt{p_{ij}}\} - \text{SIN}^{-1}\{\sqrt{\hat{p}_{ij}}\} \right]^2. \quad [1.4]$$

In deze formule is N het aantal proefpersonen; SIN^{-1} is de inverse sinus of arcsinus die uit sinuswaarden, die tussen 0 en 1 liggen, de waarden van de bijbehorende hoeken berekent. In deze formule zijn de arcsinuswaarden (de hoeken) uitgedrukt in graden (niet in radialen!). De dubbele sommatie

$$\sum_{i=1}^{m-1} \sum_{j=i+1}^m$$

geeft aan dat van een tabel met elementen i, j alleen de getallen bij elkaar worden opgeteld die rechtsboven de diagonaal van de tabel liggen. De sommatie begint met het tweede getal in de eerste rij en eindigt met het m -e getal in de $(m-1)$ -e rij. De toetsingsgrootte van Formule [1.4] is (onder de nulhypothese van geen verschil tussen p en \hat{p}) χ^2 -verdeeld met vrijheidsgraden

$$df = (m-1)(m-2)/2. \quad [1.5]$$

In ons voorbeeld is $\chi^2 = 8.77$ ($df = 10$; $.70 > p > .50$). Daaruit blijkt dat de geobserveerde proporties goed bij het Case V-model passen.

Wanneer Mostellers toets zou aantonen dat data en model significant van elkaar afwijken, dan kan dat aan drie factoren liggen:

- (a) de z_{ij} -waarden zijn niet normaal verdeeld rond $u_i - u_p$
- (b) de standaarddeviaties s_i en s_j zijn niet aan elkaar gelijk, en
- (c) de stimuli variëren op meer dan één dimensie.

Om mogelijkheid (b) te toetsen kan men nu overgaan op een andere variant van het Thurstone-model, namelijk Case III waarin alleen wordt aangenomen dat $r_{ij} = 0$; s_i en s_j zijn dan vrije parameters die geschat moeten worden. Hoe de parameters van het Case III-model berekend moeten worden, is te vinden in onder andere Edwards (1957), Guilford (1954) en

Torgerson (1958). Is de Mosteller-test wel significant onder Case V maar niet significant onder Case III dan weten we dus dat voor sommige of alle stimuli er verschillende standaarddeviaties s_i geschat moeten worden. Is de toets ook onder Case III significant, dan zijn de stimuli waarschijnlijk niet eendimensionaal.

1.2 KENMERKEN VAN SCHAALTECHNIKEN

Terminologie

In de twee voorbeelden (de paarsgewijze vergelijking van misdaden en de sortering van persoonlijkheidsadjectieven) die hierboven besproken zijn, werden methoden gebruikt die onder de verzamelnaam schaaltechnieken vallen. Zoals we daarbij gezien hebben, zijn er schaaltechnieken die voor iedere stimulus één kwantificatie of schaalwaarde opleveren en technieken die elke stimulus twee of meer schaalwaarden toekennen. In het eerste geval spreekt men van eendimensionale schaaltechnieken, in het laatste geval van meerdimensionale schaaltechnieken. Het ontwerpen, bestuderen en toepassen van dit soort technieken wordt in het Engels *multidimensional scaling* genoemd. Deze activiteiten worden vaak kortweg aangeduid als *scaling* of als MDS. Omdat er voor de term *scaling* geen goede Nederlandse vertaling bestaat, zal in dit boek meestal de afkorting MDS gebruikt worden, met name in allerlei samengestelde woorden die met deze activiteiten te maken hebben. Ter illustratie: MDS betekent in dit boek dus vooral het toepassen van MDS-technieken in de analyse van gegevens. MDS-technieken berusten op MDS-modellen en worden toegepast met behulp van MDS-(computer)programma's waarin MDS-algoritmen zijn uitgewerkt. De gegevens moeten uiteraard voor MDS-analyse geschikt zijn. Hoe men zulke gegevens moet analyseren is een MDS-probleem; het resultaat van de analyse is een MDS-oplossing.

Zowel bij de eendimensionale als bij de meerdimensionale schaaltechnieken gaat het om data-analysemethoden die ons in staat stellen *schaalwaarden* te bepalen voor de *objecten* die we willen onderzoeken. Schaalwaarden zijn getallen op één of meer variabelen, criteria of dimensies waarop de objecten van elkaar verschillen. Enerzijds kunnen deze dimensies van tevoren door de onderzoekers zijn uitgekozen en expliciet aan de proefpersonen worden voorgelegd. Anderzijds kunnen ze impliciet in het (beoordelings)gedrag van de proefpersonen aanwezig zijn en later door de onderzoeker moeten worden opgespoord. Bij de eendimensionale technieken worden voor de objecten schaalwaarden berekend op één *schaal* of *dimensie*, terwijl er bij de meerdimensionale technieken schaalwaarden berekend worden op meerdere dimensies of schalen tegelijkertijd. Heeft men per object twee of meer schaalwaarden

berekend, dan kan men die opvatten als de coördinaten op de assen van een ruimte: we kunnen dan een ruimtelijke afbeelding van de objecten maken, of, anders gezegd, we hebben dan een *ruimtelijk model* waarin de objecten zijn afgebeeld. De relaties tussen de afbeeldingen van de objecten in zo'n model (bijvoorbeeld hun onderlinge afstanden) 'verklaren' of 'voorspellen' dan de geobserveerde relaties tussen de stimuli (bijvoorbeeld hoe vaak ze samen in één groepje zijn gesorteerd). Merk op dat het woord 'objecten' hier in zijn meest algemene betekenis gebruikt wordt: het kan afwisselend dingen, begrippen, stimuli, personen en variabelen betekenen.

Zoals hierboven is opgemerkt, is scaling een vorm van data-analyse; om schaalwaarden te (kunnen) bepalen zijn er data nodig en bovendien een aantal veronderstellingen over hoe die data tot stand zijn gekomen. Data kunnen zeer uiteenlopend van aard en vorm zijn. Ze kunnen verzameld worden door personen objecten te laten beoordelen, maar ze kunnen ook op heel andere manieren tot stand komen. Belangrijk is alleen dat de geobserveerde data de relevante relaties tussen de objecten weergeven die nodig zijn om ze in een ruimtelijk model af te beelden. Om voor MDS geschikt te zijn, moeten geobserveerde data opgevat of herleid kunnen worden tot zogenaamde *nabijheidsgegevens*⁴ (proximity data; proximities). Van zulke proximities wordt verondersteld dat ze corresponderen met *afstanden* tussen de objecten in een ruimtelijk model. De precieze relatie tussen de proximity data en de afstanden kan vele vormen aannemen. Soms zijn de proximities identiek aan afstanden, soms neemt men aan dat ze niet meer zijn dan een monotone (stijgende of dalende) functie van de afstanden. In dat geval zegt men dat de nabijheidsgegevens *ordinaal* zijn. Een van de problemen waar het bij MDS om gaat is dus het bepalen van de relatie tussen de data en ruimtelijke afstanden. Een tweede probleem is het herleiden van deze afstanden tot coördinaten in een meerdimensionaal assenstelsel. Daarbij zijn twee aspecten van belang: het aantal dimensies en de precieze relatie tussen de coördinaten en de afstanden. Deze relatie, de zogenaamde *afstandsfunctie*, kan vele vormen aannemen. De gebruikelijke *Euclidische* afstandsfunctie is er een van.

Het geheel van veronderstellingen met betrekking tot de onderlinge relaties van data, afstanden, dimensies en coördinaten wordt een *model* genoemd. We verwachten dat de ruimtelijke configuratie waarin de objecten worden afgebeeld een goede weergave is van de 'werkelijke' relaties tussen de stimuli. Bovendien verwachten we dat de afbeelding ons in staat stelt voorspellingen te doen over stimuli en relaties die (nog) niet geobserveerd zijn. Bijvoorbeeld: op grond van Thurstones en Coombs' schaalwaarden van misdrijven kunnen we voorspellen dat plegers van misdrijven met een hogere schaalwaarde strenger gestraft zullen worden. De configuratie van persoonlijkheidseigenschappen

4 In de rest van dit boek gebruiken we voor nabijheidsgegevens ook vaak de termen nabijheden, proximities en afstandgegevens. Twee Engelse termen voor dit soort gegevens zijn *similarities* en *dissimilarities*. Deze zullen we in het Nederlands vertalen met *gelijkenisgegevens* of kortweg *gelijkenissen*.

leidt tot de voorspelling dat zwijgzame mensen over het algemeen niet als warm beoordeeld zullen worden maar eerder berekenend worden gevonden.

Definities van schaaltechnieken

Wanneer men (meerdimensionale) schaaltechnieken wil definiëren, dan kan men in meer of mindere mate de nadruk leggen op twee aspecten die bij scaling een rol spelen: (a) het toekennen of berekenen van schaalwaarden om iets over objecten te zeggen en (b) het construeren van een ruimtelijk model om de objecten in af te beelden. Legt men de nadruk op het eerste aspect, het toekennen van getallen, dan vat men scaling dus op als een vorm van *meten*. Immers, de meest gebruikte definitie van meten is 'het toekennen van getallen aan objecten volgens bepaalde regels' (vgl. Meerling, 1989). Bij scaling gebeurt dat zodanig dat de relevante relaties tussen de objecten optimaal worden weergegeven door de berekende, afgeleide of geschatte schaalwaarden. Deze opvatting van scaling is te herkennen in de definitie van Young (1984) waarin hij de term scaling gelijkstelt aan het meten van data van een laag (kwalitatief) meetniveau:

Scaling is the process that uses rules to assign numbers to attributes of things or events observed in circumstances assumed by the observer to be qualitative. Scaling produces measurements, and these measurements are called scale values.
(p. 55)

Vanuit dit gezichtspunt zijn schaaltechnieken met name geschikt om data die op ordinaal niveau gemeten zijn op hogere meetniveaus optimaal te kwantificeren. In dat geval ziet men schaaltechnieken vooral als methoden voor *optimal scaling*.

Legt men de nadruk op het tweede aspect, het ruimtelijk afbeelden, dan ziet men schaaltechnieken dus als een verzameling wiskundige (meetkundige, ruimtelijke) modellen die ervoor dienen om de onderlinge relaties tussen de objecten af te beelden. Deze verschillen in nadruk zijn terug te vinden in de diverse definities die er van scaling in de literatuur te vinden zijn. Zo stellen Carroll en Arabie (1980) dat

Broadly defined, multidimensional scaling comprises a family of geometric models for multidimensional representation of data and a corresponding set of methods for fitting such models to actual data. A much narrower definition would limit the term to spatial distance models for similarities, dissimilarities or other proximities data. (p. 608)

Zij leggen dus vooral de nadruk op het model en minder op de data, die (evenmin als bij Young) niet *per se* afstandsgegevens hoeven te zijn. MacCallum (1988) daarentegen, legt weer de nadruk op het modelkarakter van schaaltechnieken, die hij vooral reserveert voor de analyse van gelijkenisoordelen:

it is ... possible to identify a class of procedures which clearly fall in the category of mds and which have a common purpose. That purpose is to represent and provide a basis for understanding the structure inherent in certain types of data involving judgments about stimuli. ... the most common form of data to which mds is applied consists of judgments of perceived similarity or dissimilarity between stimuli. (p. 421)

Bovengenoemde definities van MDS lijken dus nogal van elkaar te verschillen. Wat betreft deze verschillen kunnen we het volgende opmerken.

- 1 Het eerste verschilpunt, MDS als model om structuur te beschrijven versus MDS als techniek om schaalwaarden te berekenen, is geen echte tegenstelling. Immers: om structuur te beschrijven is kwantificatie nodig; na kwantificatie laten de objecten hun onderliggende structuur zien. Aan de andere kant: om te kunnen kwantificeren is de structuur van een model nodig. Het gaat bij dit onderscheid dus meer om het doel van het onderzoek, om de vraag waar de belangstelling van de onderzoeker begint of ophoudt.
- 2 Er is geen enkele reden om MDS aan een inhoudelijk soort data te koppelen. MDS is niet uitsluitend geschikt voor subjectieve psychologische data en evenmin alleen maar voor gelijkenisoordelen. Er zijn vele soorten, zogenaamd objectieve gegevens (bijvoorbeeld echte afstanden, aantallen reizigers van de ene naar de andere stad, reistijden van de ene naar de andere bestemming, hoeveelheid import en export tussen landen, enzovoort) die bij uitstek geschikt zijn voor analyse met MDS-methoden. Evenmin is er een reden om de term schaaltechnieken uitsluitend te reserveren voor de analyse van gegevens die op een laag niveau (dat wil zeggen lager dan intervalniveau) zijn gemeten. Ook gegevens die op hoger niveau gemeten zijn (intervalniveau, rationiveau) kunnen uitstekend met behulp van MDS worden geanalyseerd. Wat MDS levert, is een *optimal scaling* van de data: de schaalwaarden leveren een optimale weergave van de relaties tussen de objecten. Optimaal betekent hier dan: in overeenstemming met het meetkarakter van de data, en dat meetkarakter kan lager of hoger zijn.

In overeenstemming met de twee bovenstaande opmerkingen kan de volgende definitie van scaling gegeven worden:

Scaling is het afleiden, bepalen, schatten of berekenen van schaalwaarden voor 'objecten' zodanig dat relevante relaties tussen die objecten optimaal worden weergegeven door de afgeleide schaalwaarden. In multidimensional scaling (MDS) worden er in principe twee of meer schaalwaarden per object afgeleid, waardoor elk object in een meerdimensionale (meestal Euclidische) ruimte afgebeeld wordt.

Hoewel er ook schrijvers zijn die een onderscheid maken tussen meten en scaling (zie Coombs, Dawes & Tversky, 1970), gaan we daar in dit hoofdstuk niet

verder op in. Wel is het van belang of men een schaalmethode puur *als techniek* gebruikt om schaalwaarden te vinden, of dat men de schaalmethode gebruikt als *criterium* om uit te maken of het veronderstelde model opgaat. In het laatste geval gaat men expliciet na of er voldoende overeenstemming is tussen de data en de uitkomsten om vol te kunnen houden dat het model juist is. In het eerste geval doet men dat niet. Gebruikt men een schaalmethode als criterium en blijkt het veronderstelde model houdbaar te zijn, dan kan men de gevonden schaalwaarden als zodanig gebruiken. De schaalwaarden vormen dan ‘een toegift van de theorie’ (Bezembinder, 1970). Is het model niet houdbaar gebleken, dan heeft het verder weinig zin om de schaalwaarden te gebruiken.

Expliciete en impliciete dimensies

Bij wijze van voorproefje werd in dit hoofdstuk een voorbeeld gegeven van de eindimensionale Thurstone-methode en van MDS op sorteergegevens. Daarbij kwamen twee belangrijke verschillen aan de orde. Het eerste verschil tussen beide methoden bestaat eruit dat de eerste één schaalwaarde voor ieder object oplevert, terwijl er in MDS twee of meer schaalwaarden per object berekend worden. Dat laatste hoeft overigens niet altijd op te gaan. Soms is een verzameling stimuli slechts eindimensionaal. In dat geval blijkt – ook bij toepassing van MDS – dat de stimuli op één rechte lijn zijn af te beelden. Een tweede, belangrijk, verschil tussen de Thurstone-methode en MDS is dat bij de eerstgenoemde methode het de onderzoeker is die de proefpersonen een criterium (grootte, ernst) voorschotelt waarop zij hun beoordelingen moeten maken. In het sorteringvoorbeeld werden de criteria voor gelijkenis aan de proefpersonen zelf overgelaten. Het doel van MDS is dan om door middel van schaalwaarden en de bijbehorende ruimtelijke afbeelding de onderliggende dimensies op te sporen waarin de stimulusobjecten variëren. Daarbij gaat het onder meer om het *aantal* dimensies en om de inhoudelijke *interpretatie* van die dimensies. Hierbij doet het er niet toe of de proefpersonen zich al dan niet bewust zijn van de criteria die zij gebruikten.

Dat neemt niet weg dat men ook in MDS-onderzoek de proefpersonen expliciete criteria kan geven om de stimuli te beoordelen. Vaak wordt een aantal proefpersonen gevraagd een verzameling stimuli te beoordelen op zogenaamde *rating scales*. Door middel van cijfers moeten de proefpersonen aangeven in hoeverre een aantal eigenschappen van toepassing is op de stimuli. Voor iedere stimulus wordt dan een profiel van beoordelingen verzameld en die profielen kan men gebruiken om nabijheidsmaten voor de stimuli te berekenen (zie Hoofdstuk 4). De resulterende nabijheden kunnen vervolgens met behulp van MDS geanalyseerd worden.

Het nut van meerdimensionale schaaltechnieken

Aan het slot van dit hoofdstuk dringt zich misschien de vraag op waarom MDS-technieken nuttig, of zelfs ‘leuk’ zijn om toe te passen, anders gezegd, waarom men zo nodig schaalwaarden zou willen berekenen. Op deze vraag zijn minstens vier antwoorden te geven:

- 1 Schaalwaarden geven inzicht in de structuur van relaties tussen de objecten (over die structuur kan men van tevoren verwachtingen hebben of niet).
- 2 MDS leidt meestal tot een simpele, grafische weergave ('een plaatje zegt meer dan duizend woorden').
- 3 MDS stelt ons in staat na te gaan hoe bekende (zowel objectieve als subjectieve) kenmerken van de objecten een rol spelen in de uiteindelijke afstanden tussen de objecten (dat wil zeggen in de afstanden in het plaatje).
- 4 Schaalwaarden kunnen worden toegepast in vervolgonderzoek, of bij het nemen van beslissingen.

Bovenstaande antwoorden zullen we aan de hand van het volgende voorbeeld nader bekijken. Stel dat men personen oordelen laat geven over politieke uitspraken (bijvoorbeeld: 'de samenleving moet meer geld beschikbaar stellen voor de opvang van politieke vluchtelingen', 'er moeten meer gevangenisstraffen gebouwd worden', en 'de studietijd aan de universiteiten moet drie jaar worden'). Door middel van MDS-technieken is het nu mogelijk schaalwaarden voor die uitspraken (de objecten) te bepalen die, afgebeeld in een grafiek, laten zien welke uitspraken 'dicht bij elkaar liggen' en welke niet. We krijgen op die manier een beeld van de structuur van politieke uitspraken, de politieke ruimte. In de meeste gevallen beginnen we niet volstrekt naïef aan een dergelijk onderzoek. We weten van tevoren al het een en ander af van politieke uitspraken: we hebben een vermoeden welke uitspraken 'links' en welke 'rechts', welke conservatief en welke progressief, welke confessioneel en welke niet-confessioneel zijn. Op grond van zulke *a priori* kennis of vermoedens hebben we bepaalde verwachtingen of hypothesen over de structuur van politieke uitspraken. Het plaatje dat we bij de MDS-analyse vinden, geeft ons dus een beeld of en in hoeverre onze oorspronkelijke verwachtingen juist waren.

Soms zijn stimuli zodanig geselecteerd of geconstrueerd dat het bekend is welke variabelen zij vertegenwoordigen. Bijvoorbeeld: we kunnen politieke uitspraken construeren die variëren met betrekking tot *groepering in de samenleving* (werkgevers, werknemers, gepensioneerden, uitkeringstrekkers, studenten), *aandachtsgebied* (financiën, gezondheid, cultuur, rechten) en *inhoud* (positief, negatief)⁵. Al deze aspecten kunnen we met elkaar combineren en voor iedere combinatie kunnen we uitspraken verzinnen. Dit leidt bijvoorbeeld tot uitspraken als: 'werknemers moeten afzien van loonsverbetering', 'gepensioneerden moeten gestimuleerd worden om naar de opera te gaan', en 'studenten moeten meer bij het universitaire bestuur betrokken worden'. Hebben we nu gelijkenis-oordelen over de uitspraken, dan kunnen we die via MDS omzetten in afstanden. Het is dan interessant om na te gaan welke facetten van de stimuli al of niet bijdragen aan de uiteindelijke afstanden, en hoe de relatie tussen afstanden en facetten er precies uitziet.

5 Een dergelijke opzet heet een *facet design*. De term stamt van Guttman (zie Borg & Lingoes, 1987). Het construeren en analyseren van facet designs is onderwerp van de facet-theorie (zie Hoofdstuk 12).

Beschikken we eenmaal over schaalwaarden voor de politieke uitspraken, dan zijn die vervolgens te gebruiken om de politieke attitudes van personen te meten. Dat kan gebeuren door respondenten een vragenlijst aan te bieden waarin gevraagd wordt met welke politieke uitspraken zij het eens zijn en met welke niet. Op grond van de uitspraken waarmee iemand het eens is, kan men die persoon een score toekennen die zegt waar hij of zij in de politieke ruimte staat. Bijvoorbeeld door voor die persoon de gemiddelde schaalwaarde te berekenen van de uitspraken waar hij of zij het mee eens is. Deze scores kunnen dan weer in ander onderzoek gebruikt worden, dat erop gericht is correlaten van politieke attitude (bijvoorbeeld stemgedrag) te onderzoeken.

1.3 OPZET VAN DIT BOEK

In het tweede hoofdstuk van dit boek wordt een typisch, klassiek voorbeeld van MDS besproken: MDS als *cartografische techniek*. In het derde hoofdstuk wordt een indeling gegeven van de verschillende soorten en vormen van observaties en data: de zogenaamde Datatheorie. In dat hoofdstuk worden allerlei aspecten, waaronder de *meetkenmerken* van de data behandeld. In Hoofdstuk 4 volgt een overzicht van *afstandsmodellen* die in MDS van belang zijn, waarna in Hoofdstuk 5 wordt ingegaan op de meetkundige eigenschappen van MDS-oplossingen. Na deze achtergronden volgt Hoofdstuk 6, waarin het klassieke, niet-metrische MDS-algoritme behandeld wordt. In Hoofdstuk 7 wordt de interpretatie van MDS-oplossingen onder de loep genomen. Daarna worden in Hoofdstuk 8 en 9 twee zogenaamde gewogen MDS-modellen besproken. In de Hoofdstukken 6 tot en met 9 krijgt de lezer *en passant* een korte inleiding in het computerprogramma ALSCAL, dat gebruikt kan worden om de besproken MDS-problemen op te lossen. In Hoofdstuk 10 en 11 gaat het om de analyse van zogenaamde preferentiegegevens. Daarbij wordt vooral gebruikgemaakt van enkele computerprogramma's voor niet-lineaire multivariate analyse, met name HOMALS, PRINCALS en ANACOR. In het twaalfde hoofdstuk wordt de opzet van een MDS-onderzoek behandeld en worden verschillende vormen van data-verzameling beschreven. Het dertiende hoofdstuk behandelt enkele bijzondere toepassingen van MDS-technieken. Hoofdstuk 14, ten slotte, geeft een beknopt historisch overzicht van MDS waarbij verwezen wordt naar belangrijke literatuur op dit gebied. Het boek wordt afgesloten door een appendix waarin een aantal kernbegrippen uit het matrixalgebra wordt behandeld.

