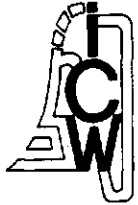


NN31545.1859

BIBLIOTHEEK
STARINGGEBOUW

ICW nota 1859 #
maart 1988



nota

— instituut voor cultuurtechniek en waterhuishouding, wageningen —

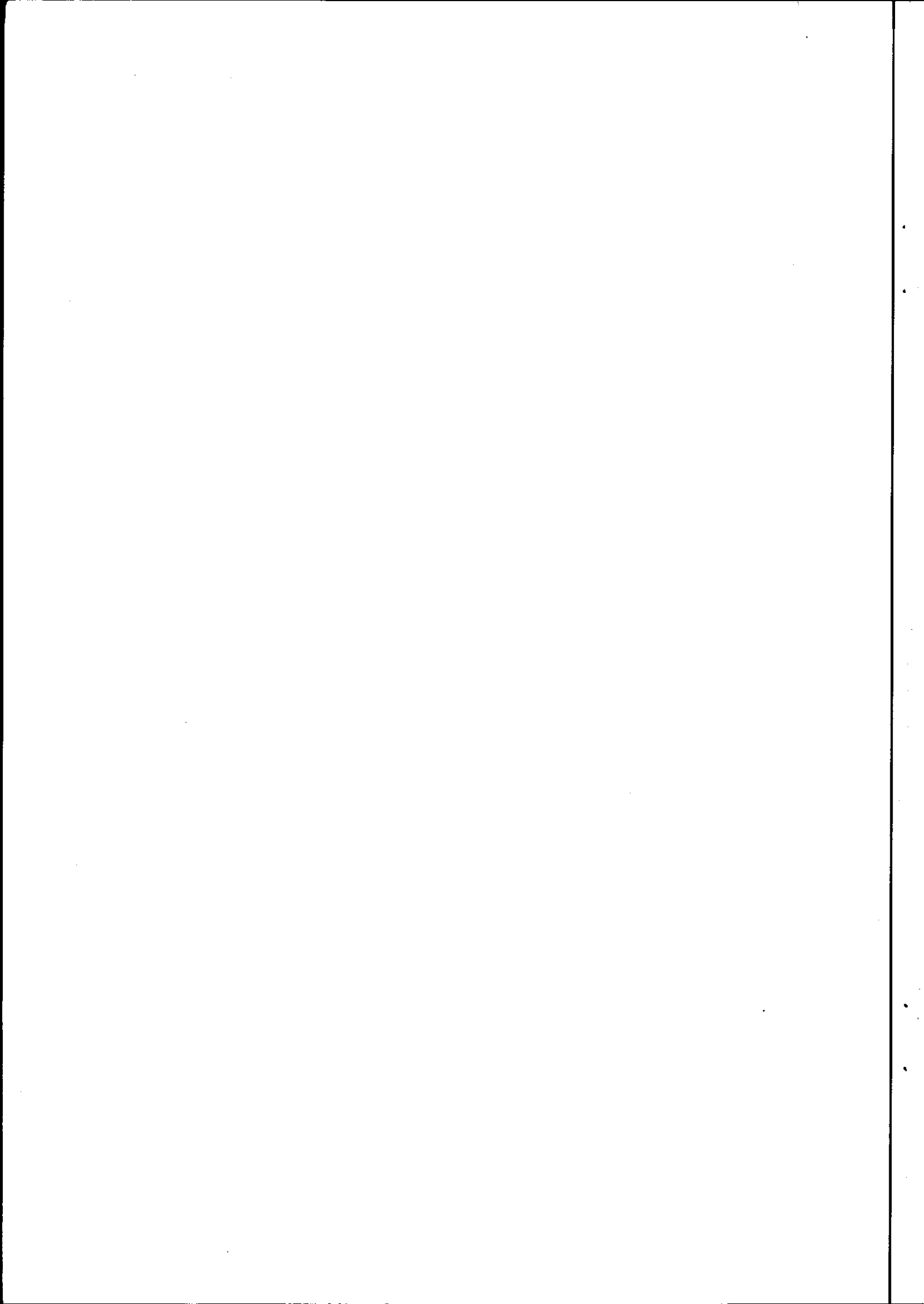
ENIGE TOEPASSINGSERVARINGEN MET DE METHODE VAN MEERDIMENSIO-
NAAL ONTVOUWEN IN HET OPENLUCHTRECREATIE-ONDERZOEK

P.M.A. Klinkers

Nota's van het Instituut zijn in principe interne communicatie-
middelen, dus geen officiële publikaties.
Hun inhoud varieert sterk en kan zowel betrekking hebben op een
eenvoudige weergave van cijferreeksen, als op een concluderende
discussie van onderzoeksresultaten. In de meeste gevallen zullen
de conclusies echter van voorlopige aard zijn omdat het onderzoek
nog niet is afgesloten.
Bepaalde nota's komen niet voor verspreiding buiten het Instituut
in aanmerking

16 MEI 1988

JSN 264980 *



INHOUD

1. Inleiding	blz. 3
2. Bruikbaarheid van de methode	3
3. Specifieke eigenschappen van de methode	4
4. Het principe van meerdimensionaal ontvouwen	5
5. Het computerprogramma UNFOLD	8
6. Een concreet voorbeeld	9
7. Conclusies t.a.v. verdere toepassingsmogelijkheden	17
Literatuur	19
Begrippenverklaring	21
Bijlage 1	

1 INLEIDING

In deze nota worden enkele ervaringen besproken met de methode van meerdimensionaal ontvouwen en de toepassing ervan via het computerprogramma Unfold.

In het volgende hoofdstuk (2) wordt kort aangegeven wat de bruikbaarheid van de methode is in relatie tot het type onderzoek en het type vraagstelling. Vervolgens worden in hoofdstuk 3 de specifieke eigenschappen van de meetmethode besproken, alsmede de mate waarin deze te onderscheiden is van andere analysetechnieken. In hoofdstuk 4 wordt het principe van meerdimensionaal ontvouwen nader toegelicht. Voor de liefhebbers wordt in hoofdstuk 5 beschreven hoe in het programma Unfold de ontvouwingmethode is geformaliseerd. Enige wiskundige voorkennis is bij het lezen van dit hoofdstuk (helaas) onontbeerlijk. In hoofdstuk 6 wordt een concreet voorbeeld uit het recreatieonderzoek besproken, waarna in hoofdstuk 7 tot slot enkele conclusies worden geformuleerd.

Voor de verklaring van een aantal gebruikte begrippen wordt verwezen naar de verklarende begrippenlijst achter in deze nota.

2 BRUIKBAARHEID VAN DE METHODE

De methode van meerdimensionaal ontvouwen is een analysetechniek die afkomstig is uit de sociale wetenschappen (psychometrie), en is bedoeld als een techniek voor de verwerking en analyse van onderzoeksgegevens. De methode is onder meer bruikbaar voor het ontdekken van structuur in onderzoeksgegevens zonder een vooraf vastgelegd categorieënsysteem. In technisch gericht onderzoek is de relevante meeteenheid (of het categorieënsysteem) bij de aanvang van het onderzoek al bekend en bepaald: hoogte in meters, temperatuur in graden celcius, etc. In sociaal wetenschappelijk onderzoek is dit meestal niet het geval; de onderzoeker is geïnteresseerd in een verschijnsel (b.v. het oordeel van recreanten over een recreatievoorziening) maar heeft niet de relevante meeteenheid voorhanden (interviewvragen). Het kiezen van een geschikt lijkend categorieënsysteem geschiedt doorgaans dan ook op basis van een kleine dosis kennis/theoretisch inzicht en een grote dosis intuïtie/praktisch inzicht. Met behulp van de meerdimensionale ontvouwingstechniek nu kan de keuze van een categorieënsysteem gefundeerd worden op een empirische basis. Uiteraard ontslaat dit de onderzoeker niet van zijn verantwoordelijkheid om na te gaan in hoeverre het gevondene past binnen de theoretische stand van zaken.

Uit het voorgaande is duidelijk geworden dat de besproken analysetechniek met name bruikbaar is voor sociaal-wetenschappelijk onderzoek, waar de mens als onderzoekseenheid geldt (sociale geografie, sociologie, psychologie, economie). Toepassing ervan is mogelijk zowel in verkennend onderzoek, waarbij een totaal nieuw en onbekend probleemveld wordt onderzocht alsook in toetsend onderzoek, waarbij de onderzoeker over genoeg voorkennis beschikt om relevante hypothesen te kunnen formuleren. In het laatste geval kan b.v. als onderzoeksvraag gelden in hoeverre een gehanteerd categorieënsysteem aansluit bij de ervarings- en belevingswereld van betrokken personen.

Roskam (1983) merkt op dat het toetsend gebruik bij dergelijke

meetmethoden eigenlijk op de eerste plaats zou moeten komen.

3 SPECIFIEKE EIGENSCHAPPEN VAN DE METHODE

In het vorige hoofdstuk is uiteengezet wat de bruikbaarheid van de meetmethode is, of zou kunnen zijn. Thans zal worden ingegaan op de specifieke eigenschappen van de methode en de mate waarin deze te onderscheiden is ten opzichte van andere analysemethoden.

De ontvouwingmethode behoort tot de familie van MeerDimensionale Schalingsmethoden, meestal aangeduid als M.D.S. Op een tweetal punten onderscheiden deze methoden zich van de meer traditionele analysetechnieken zoals regressie-analyse, factoranalyse e.d.

Allereerst bieden deze methoden de mogelijkheid tot analyse van gegevens wanneer meerdere eigenschappen van zaken (of personen) in onderlinge combinatie de reactie bepalen van de onderzochte persoon. B.v. de voorkeur die een recreant heeft voor een strandbad, wordt bepaald door een combinatie van vormgevingsfactoren (ziet de omgeving er aantrekkelijk uit) en functionele factoren (welke voorzieningen zijn er). Bij M.D.S.-technieken worden factoren onderzocht naar de mate waarin ze van invloed zijn, al of niet in combinatie met elkaar. Bij de meeste gangbare analysetechnieken is juist sprake van het zoeken naar onafhankelijke beïnvloedende factoren (onderlinge ongecorrleerdheid).

Op zichzelf is de ene techniek niet beter of slechter dan de ander, het is hooguit zo dat de ene techniek beter past bij een bepaald type vraagstelling. B.v. in het geval van de voorkeur van recreanten ten opzichte van strandbaden, is het weinig zinvol om te onderzoeken welke functionele factoren de keuze van een strandband bepalen, onafhankelijk van vormgevingsfactoren. Immers, de recreant kan zich in zijn keuze niet beperken tot louter functionele factoren, aangezien deze voor hem/haar niet los staan van vormgevingsfactoren. Voor een dergelijke vraagstelling is een techniek zoals regressie-analyse dan ook minder geschikt.

Een tweede kenmerk van M.D.S.-technieken is, dat het niet-metrische analysetechnieken zijn. De gangbare analysetechnieken zijn metrische technieken, d.w.z. het gebruik van deze technieken veronderstelt dat de gegevens van interval meetniveau zijn.

Ter toelichting op het begrip meetniveau: methodologisch onderscheidt men 4 soorten meetniveau's, afhankelijk van de wijze waarop de gemeten gegevens ingedeeld kunnen worden:

1. Ratio-metniveau

Dit is het hoogste meetniveau, d.w.z. het kent de meeste restricties. Als gegevens op dit niveau gemeten zijn, kunnen de gegevens ingedeeld worden in klassen met onderling vaste afstanden tot elkaar. Tevens is er een vast beginpunt of nulpunt bekend; dit is de ondergrens waar beneden geen waarnemingen kunnen optreden. Voorbeeld: lengte in meters, waarbij 0 meter een vast beginpunt is (er bestaat geen negatieve lengte).

2. Interval-metniveau

Dit komt overeen met het ratio-metniveau, echter nu kan de keuze van een zinvol nulpunt op meerdere wijzen geschieden. B.v. het

meten van temperatuur kan geschieden in graden Celcius of graden Fahrenheit, elk met een ander nulpunt.

3. Ordinaal-meetniveau

Hierbij wordt weer een restrictie verlaten: er is nog wel een indeling naar klassen of groepen in een oplopende reeks, echter de klassen bezitten geen vaste afstand tot elkaar. Voorbeeld: antwoordcategorieën op de vraag of men tevreden is over het aantal surfplassen in de woonomgeving kunnen er als volgt uitzien: 1) zeer ontevreden, 2) ontevreden, 3) neutraal, 4) tevreden, 5) zeer tevreden. In dit geval is het onjuist om te zeggen dat bijvoorbeeld het interval 1 tot en met 5 vijf maal zo lang is als interval 1; wel is er sprake van een rangordening van laag naar hoog.

4. Nominaal-meetniveau

Dit is het laagste meetniveau, waarbij sprake is van een indeling in klassen of groepen op grond van een bepaald kenmerk of definitie. Er is geen rangordening aan te brengen in de klassen c.q. antwoordcategorieën. B.v. een indeling in geslacht (man-vrouw) of de mening van personen over een stelling (mee eens - oneens).

Bij deze indeling in meetniveau's is er sprake van een hiërarchie; de hogere meetniveau's stellen de meeste restricties. Zijn de gegevens van niveau 1 (ratio) of 2 (interval), dan noemt men dit ook wel metrisch meetniveau. De klassieke analysetechnieken (w.o. alle regressie-achtige technieken) leveren resultaten die pas zinvol geïnterpreteerd kunnen worden als de gegevens minstens op interval-niveau gemeten zijn, of als anderszins mag worden aangenomen dat het interval-niveau van toepassing is op de gegevens.

Bij de M.D.S.-technieken zijn de gestelde restricties ten aanzien van meetniveau minder zwaar: zowel gegevens van ordinaal- als nominaal meetniveau kunnen zinvol geanalyseerd worden.

4 HET PRINCIPE VAN MEER-DIMENSIONAAL ONTVOUWEN

In het navolgende wordt ingegaan op meerdimensionaal ontvouwen als analysetechniek: hoe werkt deze techniek en welke ideeën/veronderstellingen liggen er aan ten grondslag.

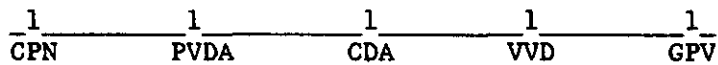
Ontvouwen (in het Engels: Unfolding) heeft betrekking op het ont-rafelen van preferentiegegevens, zodanig dat een ruimtelijke afbeelding wordt verkregen. Het model en de terminologie zijn afkomstig van C.H. Coombs, die beschouwd kan worden als een van de grondleggers omtrent de theorie van de moderne schaaltechnieken.

In sociaal-wetenschappelijk onderzoek is men geïnteresseerd in de relatie van een subject tot een object. Het subject van onderzoek is meestal een persoon of groep personen, het object van onderzoek omvat een veel breder terrein; het kan in principe alles zijn wat zich in de fysische werkelijkheid afspeelt. Een voorbeeld van een subject is een fietsende recreant, voorbeeld van een object is het fietspad waarop gefietst wordt. In aansluiting op de gangbare terminologie zal in het vervolg van deze nota het onderzoekssubject worden aangeduid met "individu", het onderzoeksobject met de term "stimulus" (meervoud "stimuli").

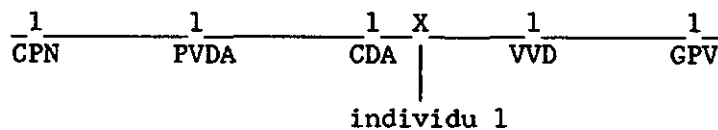
In het ontvouwingsmodel worden de relaties tussen individuen en

stimuli beschreven door zowel de stimuli als de individuen als punten in een ruimtelijk model af te beelden. Daarbij geldt dat de afstand tussen individu en stimulus de mate van voorkeur van het individu voor de stimulus weergeeft. Dit betekent dat een hoge voorkeur van een individu voor een stimulus in het model tot uitdrukking komt door een relatief kleine afstand tussen individu-punt en stimuluspunt en een lage voorkeur voor een stimulus tot uitdrukking komt in een relatief grote afstand van individu-punt tot stimuluspunt. Een individu-punt is daarbij op te vatten als 'ideaalpunt': als iemand stimulus A boven B prefereert, betekent dit dat A dichterbij het ideaal van die persoon ligt dan B. In de afbeelding correspondeert dit met een kleinere afstand van punt A tot het ideaalpunt.

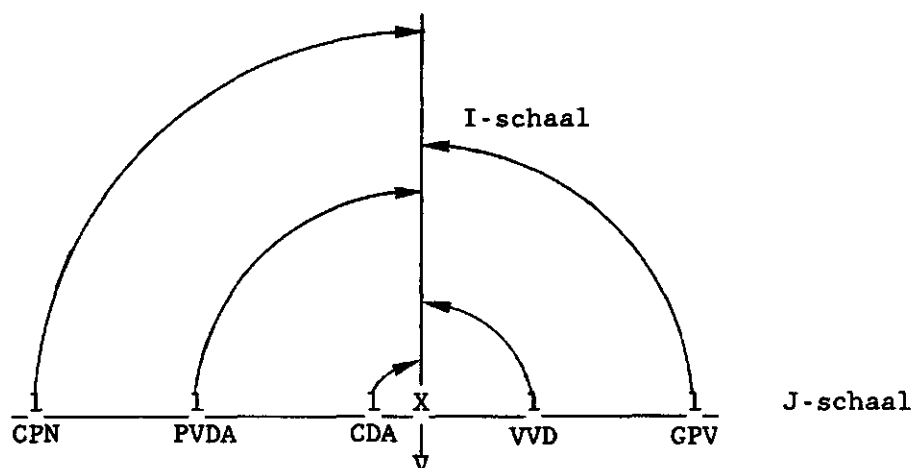
Het voorgaande kan aan de hand van een eenvoudig voorbeeld worden geïllustreerd. Gesteld dat aan twee individuen wordt gevraagd naar hun voorkeursrangordening van 5 politieke partijen, b.v. PVDA, GPV, VVD, CDA, CPN. Voorzover wordt gekeken naar het kenmerk progressiviteit van deze partijen, kan dit als volgt in een schaal worden weergegeven:



Bovenstaande schaal is te beschouwen als een dimensie, waarop alle partijen een bepaalde positie innemen. Indien de twee individuen zich in hun voorkeur uitsluitend laten leiden door de mate van progressiviteit van een partij, zullen zij bovenstaande schaal (min of meer bewust) als criterium hanteren bij de bepaling van hun voorkeur. Het is derhalve mogelijk ook de individuen een plaats op de schaal te geven, rekening houdend met het principe dat afstand gelijkenis (voorkeur) uitdrukt. B.v. individu 1 heeft als voorkeursrangordening CDA, GPV, PVDA, VVD, CPN. Zijn plaats op de schaal is als volgt:



Individu 1 is het dichtste bij het stimuluspunt van het CDA geplaatst, daarna het dichtste bij de VVD enz. Uit het voorgaande is ook te zien hoe de rangordening van individu 1 is te relateren aan de algemene schaal: iemands voorkeursrangorde is te genereren door de schaal dubbel te klappen met als scharnierpunt het ideaalpunt van het individu. Met andere woorden het ideaalpunt van een individu op de schaal kan worden afgeleid uit de rangordening van de stimuluspunten door het individu. Dit is in de volgende figuur visueel weergegeven.



De algemene dimensie, die op alle individuen van toepassing is, wordt aangeduid als de J-schaal (Joint-scale). Iemands individuele schaal, die ontstaat door de J-schaal om een bepaald punt "dubbel te vouwen", wordt I-schaal genoemd (Individual-scale).

In het voorgaande is aangegeven hoe vanuit een algemene dimensie een individuele voorkeursrangorde kan worden afgeleid. In de praktijk doet zich echter het omgekeerde probleem voor, n.l. een aantal individuele voorkeursrangordeningen die zodanig moet worden ontvouwd, dat een algemene schaal zichtbaar wordt. Welke dimensies meespelen bij het opstellen van een rangordening door individuen, is dan niet bekend; juist dát moet door middel van het onderzoek worden vastgesteld.

Allereerst kan worden getracht de voorkeursrangordeningen in een een-dimensionale schaal onder te brengen. Lukt het niet goed om voor een aantal voorkeursrangordeningen één onderliggende algemene schaal te vinden, dan kan men trachten een meerdimensionale oplossing te construeren. In zo'n oplossing worden de stimuli en proefpersonen niet als punten op een rechte lijn maar als punten in een ruimte met 2 of meer dimensies afgebeeld. Toegepast op het bovenstaande voorbeeld van politieke partijen is het voorstelbaar dat, wanneer de rangordeningen van de individuen in twee dimensies worden ontvouwen, er behalve de dimensie progressiviteit nog een tweede dimensie meespeelt, n.l. de confessionaliteit van de partij.

Om te beoordelen in hoeverre een oplossing (met een bepaald aantal dimensies) goed bij de gegevens past, wordt een zgn. verlies-functie gedefinieerd, meestal aangeduid onder de naam stress. Daarvan bestaan verschillende varianten, doorgaans gebaseerd op het "kleinste-kwadraten-kriterium" (d.w.z. een minimale som van gekwadrateerde afwijkingen tussen de gevonden oplossing en de oorspronkelijke gegevens).

Meerdimensioneel ontvouwen is zonder gebruik van een computeralgoritme vrijwel ondoenlijk. Daarom zal in de volgende paragraaf nader worden ingegaan op de mogelijkheden van het programma Unfold ten behoeve van (meerdimensioneel) ontvouwen.

5 HET COMPUTERPROGRAMMA UNFOLD

Het programma UNFOLD maakt deel uit van de KUNST-library van de Kath. Universiteit Nijmegen. Een handleiding bij het programma is te vinden in Thissen (1982). In het navolgende zullen de wiskundige achtergronden van dit programma worden belicht.

UNFOLD voert een niet-metrische analyse uit op data, d.w.z. het gaat uit van data op nominaal of ordinaal meetniveau. Voor bijvoorbeeld een aantal proefpersonen (rijen in het bestand) is een rangordening gegeven van een aantal stimuli, die de kolommen van het bestand vormen (b.v. merken auto's). UNFOLD zoekt een weergave van rijen en kolommen als punten in een ruimte, zodanig dat per rij-punt de volgorde van de afstanden ten opzichte van de kolompunten zo goed mogelijk overeenkomt met de oorspronkelijke rangordening in die rij.

Het zoeken van een ruimtelijke weergave geschiedt door middel van een iteratieve rekenprocedure. Elke iteratie omvat de volgende berekeningswijzen.

Het programma gaat uit van een bepaalde startconfiguratie; indien de gebruiker deze niet opgeeft, wordt deze door het programma gemaakt. Hierna wordt de configuratie van stimuli en individuen genormaliseerd door de oorsprong te leggen in de centroide van de stimuli, en bovendien de som der kwadraten van de stimuluscoördinaten gelijk te maken aan 1. Dit laatste gebeurt door de coördinaten met een constante te vermenigvuldigen.

Vervolgens worden de afstanden (D_{ij}) berekend volgens de zgn. Minkowski-metrick. Het algemene model hiervan is als volgt:

$$D_{ij} = \left[\sum_k |X_{ik} - Y_{jk}|^r \right]^{1/r} \quad k = 1, \dots, n$$

waarin n = totaal aantal op te geven dimensies
 X_{ik} = coördinaat van individu i op dimensie k
 Y_{jk} = coördinaat van stimulus j op dimensie k
 r is de Minkowski-metrick.

Het meest gebruikelijk als keuze voor r is:

- $r = 2$: euclidisch afstandsmodel
- $r = 1$: city-blok afstandsmodel
- $r = \infty$: dominantie afstandsmodel

Hierna worden de zgn. fitting-values (\hat{D}_{ij}) bepaald, zodanig dat zij zwak monotoon stijgend of -dalend zijn met de data en tegelijkertijd de verschillen tussen D_{ij} en \hat{D}_{ij} zo klein mogelijk zijn. Op basis hiervan wordt de representatie verbeterd, waarna hernieuwd fitting values berekend worden. Hierna wordt opnieuw de representatie aangepast etc. Dit proces wordt in de literatuur aangeduid als "the ping-pong game". Tijdens dit proces wordt gewerkt met twee soorten fitting-values; aanvankelijk Guttman's fitting algoritme (dat werkt op basis van zgn. rank images). De hieruit resulterende fitting-values heten DSTAR. Daarna wordt Kruskal's fitting algoritme gebruikt (op basis van monotone regressie). Hieruit resulteren fitting-values die DHAT genoemd worden. De "ping-pong-game" geschiedt achtereenvolgens tussen de representatie en de DSTAR-fitting values, en daarna tussen de representatie en de DHAT-fitting values.

Na elke iteratie, d.w.z. na elke "ping" in de "ping-pong-game", wordt de zogenaamde stress bepaald; dit is een maat die aangeeft in hoeverre de berekende ruimtelijke representatie als geheel past bij de data. Er zijn verschillende stress-formules in omloop, zoals eerder werd gesteld zijn deze meestal gebaseerd op het "kleinste kwadraten criterium". UNFOLD maakt gebruik van de volgende stress-formule:

$$s = \sqrt{\frac{1}{n} \sum_i \frac{\sum_j (D_{ij} - \bar{D}_i)^2}{\sum_j (D_{ij} - \bar{D}_i)^2}}$$

waarbij \bar{D}_i de gemiddelde afstand is van individu i tot de stimuluspunten

Het formuledeel achter het eerste sommatieteken heeft betrekking op individu i ; de vierkantswortel hieruit wordt genoemd de stress voor individu i . Gesommeerd over alle individuen geeft dit dus de totale stress voor de totale configuratie.

Hoewel de stress-waarde geen toetsingsgrootheid is, geldt ze wel als een criterium voor de mate waarin de ruimtelijke representatie afwijkt van de data; de representatie is optimaal wanneer de "stress" ervan zo klein mogelijk is; een stresswaarde van max. 0,05 geldt doorgaans als acceptabel.

Voorzover de berekende stresswaarde het vooraf bepaalde minimum bereikt (0,01) of te weinig vermindert ten opzichte van de vorige iteratie, wordt de berekening in deze fase beëindigd. Zoniet, dan begint UNFOLD aan de volgende iteratie.

Behalve wanneer de euclidische metriek gehanteerd wordt als afstandsmodel, resulteert een figuur waarbij het assenstelsel "vast" ligt, en zijn daarmee de dimensies gegeven waarlangs de gegevens geïnterpreteerd dienen te worden. Alleen een verschuiving van de oorsprong en een spiegeling langs de dimensies zijn in deze gevallen toegestaan. Ingeval van een euclidische metriek is rotatie van het assenstelsel in de eindconfiguratie wel mogelijk; alleen de afstanden tussen de punten onderling ligt vast. De onderzoeker heeft dan zelf de mogelijkheid te bepalen welke dimensies hij kiest als ingang tot de analyse.

6 EEN CONCREET VOORBEELD

In deze paragraaf zal een voorbeeld worden behandeld uit het recreatie-onderzoek.

In het kader van een onderzoek naar de druktebeleving van fietsende recreanten (Klinkers en Van Hoorn, 1987) is aan 1089 fietsers gevraagd naar hun mening over het aantal aanwezige fietsen op het wegvak op dat moment (te veel, goed, te weinig). Met behulp van deze vraag is dus informatie verkregen over het oordeel van de fietsers over de intensiteit van het aanwezige fietsverkeer.

Daarnaast is aan dezelfde groep fietsers (individuen) gevraagd om 4

uitspraken (stimuli) te rangordenen naar de mate waarin dit aansloot bij hun eigen mening. De 4 uitspraken hebben alle betrekking op het onderwerp fietsen, en zijn

1. drukke fietspaden kun je beter vermijden
2. hoe meer mensen op een fietspad, hoe meer vreugd
3. het is gezellig om andere fietsers tegen te komen
4. fietsen is alleen maar leuk op rustige wegen en paden

Deze 4 uitspraken zijn theoretisch op $4!=24$ verschillende manieren te rangordenen; daarvan bleken er 22 voor te komen. Deze 22 rangordeningen representeren verschillende standpunten ten opzichte van de 4 uitspraken. In tabel 1 is weergegeven met welke frequentie de 24 mogelijke rangordeningen voorkomen.

Tabel 1. Waargenomen frequentie van de mogelijke rangordeningen.

nr.	Rangordening	abs. freq.	rel. freq.
1	1234	3	0.3
2	1324	13	1.2
3	1342	43	3.9
4	1423	2	0.2
5	1432	78	7.2
6	2134	1	0.1
7	2314	7	0.6
8	2341	20	1.8
9	2413	1	0.1
10	2431	9	0.8
11	3124	14	1.3
12	3142	43	3.9
13	3214	52	4.8
14	3241	62	5.7
15	3412	110	10.1
16	3421	136	12.5
17	4123	7	0.6
18	4132	180	16.5
19	4213	9	0.8
20	4231	18	1.7
21	4312	139	12.8
22	4321	142	13.0
23	1243	0	0.0
24	2143	0	0.0
	totaal	1089	100 %

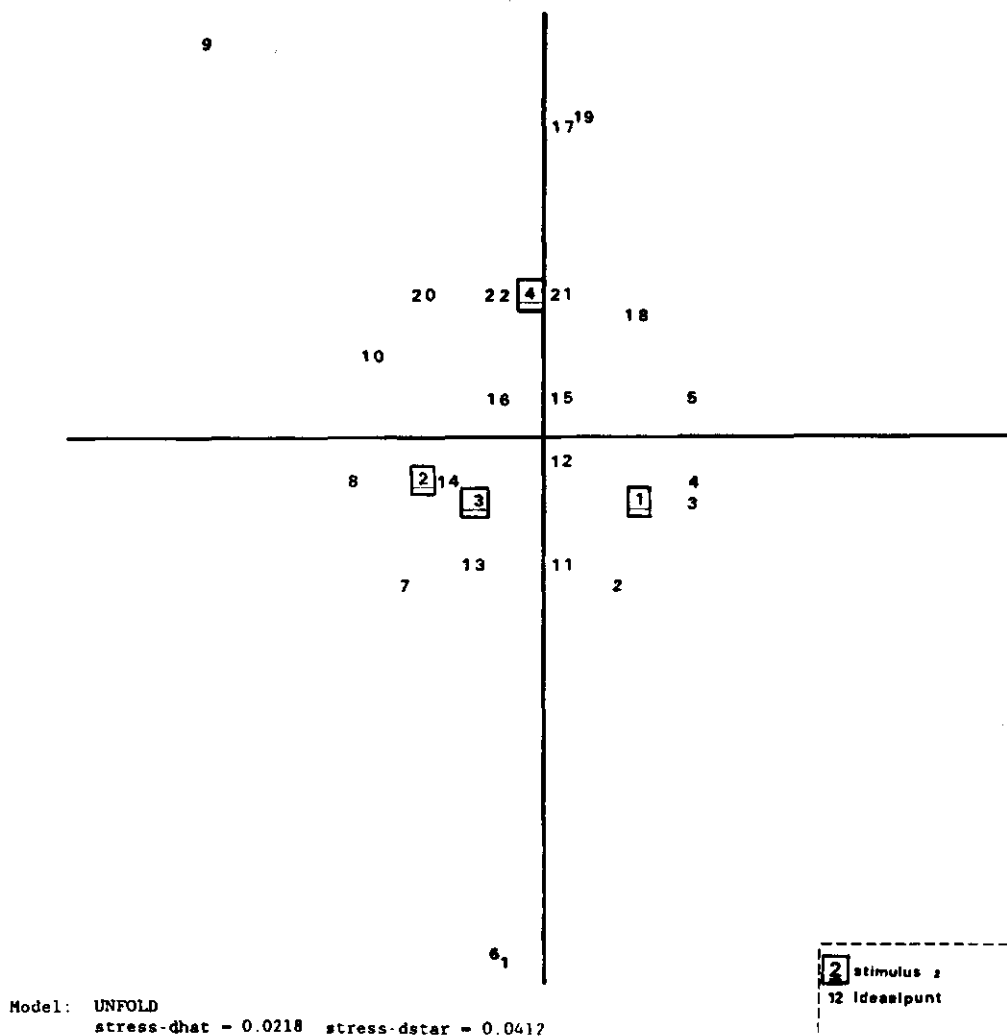
Wanneer 4 stimuli in één dimensie worden ontvouwen, zijn wiskundig gezien maximaal 7 verschillende rangordeningen toegestaan (zie bijlage 1 voor een uitgewerkt voorbeeld). Indien 4 stimuli in twee dimensies worden ontvouwen zijn 18 verschillende rangordeningen toegestaan. In deze beide gevallen geldt dat een aantal rangordeningen elkaar uitsluiten als mogelijkheid om tegelijkertijd voor te komen, d.w.z. het betreft niet willekeurig 7 respectievelijk 18 verschillende rangordeningen. In 3 dimensies ontvouwen van 4 stimuli zijn alle 24 rangordeningen toegestaan (Bennett en Hays, 1960).

Om de optimale oplossing voor bovenstaand probleem te bepalen, zijn m.b.v. Unfold 4 verschillende oplossingen uitgerekend. Deze zullen besproken worden in de volgorde waarin ze zijn bepaald, zodat de lezer zich een beeld kan vormen van het analyseproces, i.c. de overwegingen die ertoe hebben geleid dat er (al of niet) op een bepaald moment naar een verbeterde oplossing is gezocht.

Vooraf is het nuttig op te merken dat de tweede oplossing (weergegeven in figuur 2) als optimale configuratie naar voren kwam. Dit wordt duidelijk wanneer de verschillende oplossingen inhoudelijk met elkaar worden vergeleken. Daarom zal in het navolgende elke configuratie kort besproken worden.

Eerste configuratie: een gedegeneerde oplossing

In eerste instantie is getracht het model te ontvouwen in 2 dimensies volgens het euclidische afstandsmodel. De hieruit resulterende configuratie is weergegeven in figuur 1.



Figuur 1. Twee-dimensionale ontvouwing van 4 stimuli en 1089 individuen (verdeeld over 22 ideaalpunten), volgens de euclidische metriek

In figuur 1 vertegenwoordigt elk ideaalpunt een aantal individuen, overeenkomstig de weergegeven frequenties in tabel 1.

Bij het interpreteren van de figuur gaat het er om, te letten op de relatieve onderlinge afstanden van stimulus- en ideaalpunten tot elkaar; er kan geen absolute betekenis worden gehecht aan de afstanden binnen de figuur. Dit is een rechtstreeks gevolg van het meetniveau waarop de gegevens geïnterpreteerd worden, namelijk interval-meetniveau (zie hoofdstuk 3).

Hoewel er in de oplossing, behorend bij figuur 1, sprake is van een lage stress-waarde van DHAT en DSTAR van resp. 0,0218 en 0,0412, blijkt uit de figuur dat zowel de stimuli als de ideaalpunten allen nagenoeg op een kluitje liggen in de oorsprong van de figuur. In een dergelijke situatie spreekt men van een gedegenererde oplossing; er is sprake van zeer weinig differentiatie zowel tussen de stimuli als de ideaalpunten.

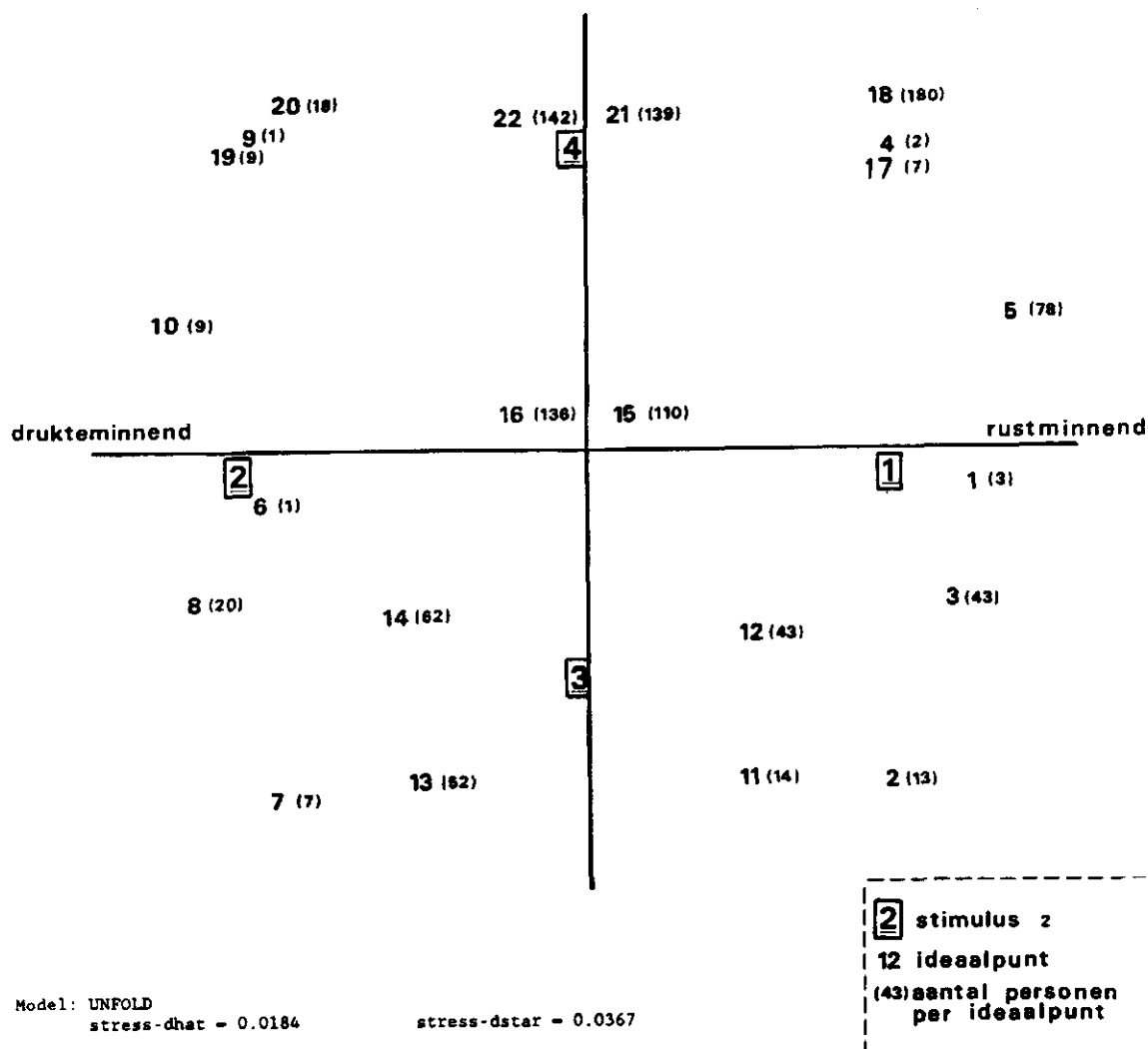
Tweede configuratie: een bevredigende oplossing

In een tweede poging is nogmaals getracht de 4 stimuli (uitspraken) en 22 verschillende rangordeningen in twee dimensies te ontvouwen, maar nu volgens het city-blok afstandsmodel. Uit theoretisch oogpunt zijn er geen redenen te noemen waarom de ene metriek te verkiezen zou zijn boven de andere. Uit ervaringsfeiten is echter wel gebleken dat de city-blok metriek beter past bij stimuli met duidelijk onderscheidbare dimensies. Bij stimuli met niet duidelijk onderscheidbare dimensies (b.v. bij muziek-klanken zijn toonhoogte en volume zeer moeilijk afzonderlijk waarneembaar) lijkt in het algemeen een euclidische metriek beter te passen (Roskam, 1983). In het onderhavige geval bleek de city-blok metriek een betere oplossing te leveren dan de euclidische metriek. In figuur 2 is weergegeven welke figuur resulteert na een ontvouwing in twee dimensies volgens de city-blok metriek.

De tweede oplossing, zoals weergegeven in figuur 2, mag bevredigend genoemd worden. De berekende stress-waarden zijn aanvaardbaar laag: (0,0184 voor DHAT en 0,0367 voor DSTAR). Ook de configuratie is veel gedifferentieerder dan figuur 1, d.w.z. zowel de stimuluspunten als de ideaalpunten vertonen een grotere spreiding, zowel ten opzichte van elkaar als ten opzichte van de dimensies.

Afgaande op de inhoud van de stimuli, lijkt de horizontale dimensie weer te geven de voorkeur voor veel respectievelijk weinig mensen op een weg/pad. De verticale dimensie is, afgaande op de inhoud van de stimuli, minder duidelijk te interpreteren; dit wordt nog eens extra bemoeilijkt door het geringe aantal stimuli waarop deze dimensie is gebaseerd.

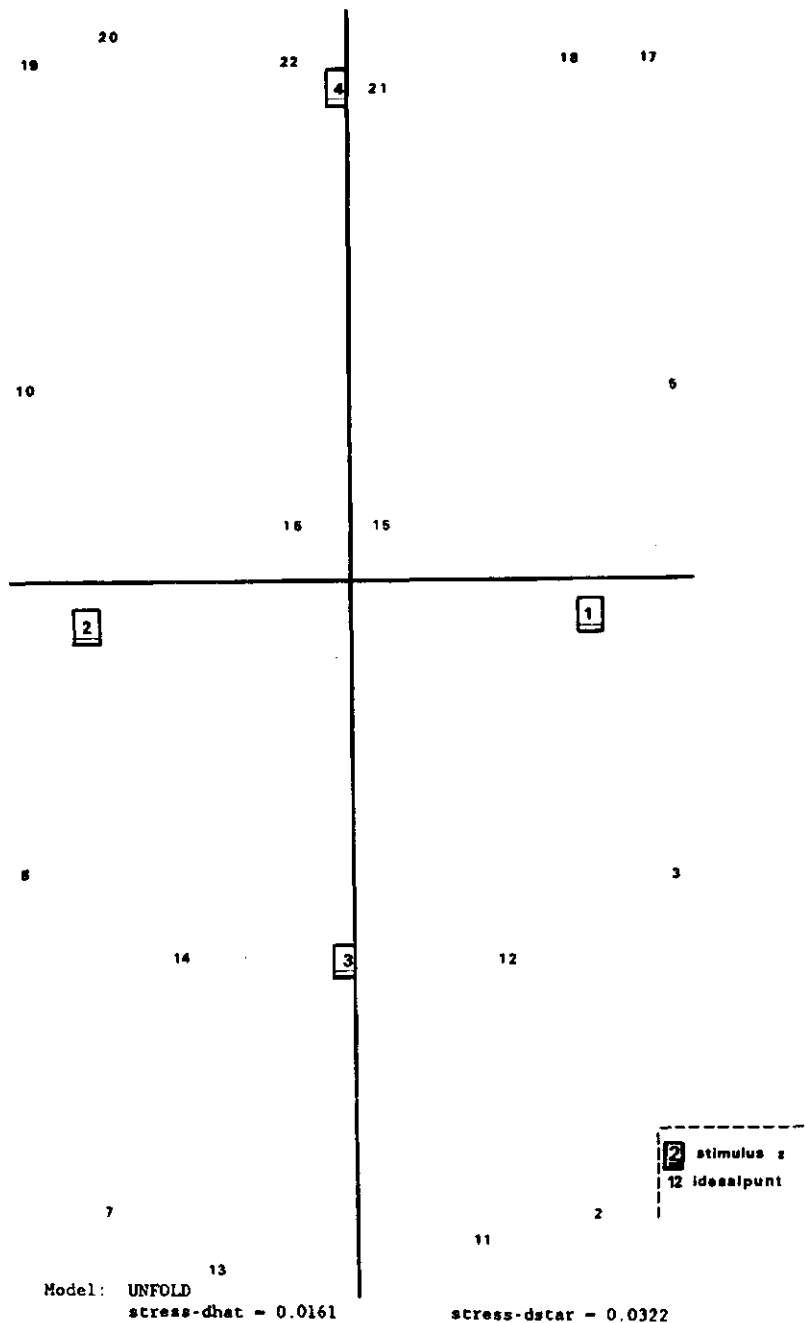
In dit verband is tevens een ééndimensionale oplossing geprobeerd (city-blok afstandsmodel, 1089 individuen). De resulterende stress-waarden waren uitzonderlijk hoog (0,2371 voor DHAT en 0,4317 voor DSTAR) en maakten daardoor deze oplossing onaanvaardbaar.



Figuur 2. Twee-dimensionale ontvouwing van 4 stimuli en 1089 individuen (verdeeld over 22 ideaalpunten), volgens de city-blok metriek

Derde configuratie: mislukte poging tot verbetering

Om na te gaan in hoeverre een verbeterde oplossing mogelijk is, is nogmaals een ontvouwing in twee respectievelijk één dimensie geprobeerd met de city-blok metriek, maar nu nadat de 4 ideaalpunten met de laagste frequentie (en relatief hoge stress-waarde per ideaalpunt) waren verwijderd, n.l. de punten 1, 4, 6 en 9 (totaal 7 individuen is 0,64 % van de populatie). De oplossing in één dimensie vertoonde nog steeds hoge stress-waarden (DHAT is 0,2173 en DSTAR is 0,4069). Figuur 3 toont een afbeelding van de oplossing in twee dimensies.



Figuur 3. Twee-dimensionale ontvouwing van 4 stimuli en 1082 individuen (verdeeld over 18 ideaalpunten), volgens de city-blok metriek

De stress-waarden van de oplossing, behorende bij figuur 3 is niet noemenswaardig lager geworden in vergelijking tot figuur 2 (DHAT is 0,0161, dit was eerst 0,0184 en DSTAR is 0,0322, dit was eerst 0,0367). Inhoudelijk gezien levert deze configuratie ook niet veel nieuws op; de configuratie is meer uitgerekt op de tweede dimensie, hetgeen juist de dimensie betreft welke moeilijk te interpreteren

valt. De posities van de stimulus- en ideaalpunten tot elkaar is in figuur 3 echter niet gewijzigd, vergeleken met figuur 2.

Vierde configuratie: laatste poging tot optimalisatie

Tot slot is de ontvouwing nogmaals herhaald in twee respectievelijk één dimensie met het city-blok model, maar nu na verwijdering van nog eens zes extra punten met lage frequenties en relatief hogere stress-waarden: de punten 2, 7, 10, 11, 17 en 19. In totaal betreft het hierbij 59 individuen is 5,4 % van de (oorspronkelijke) populatie. In één dimensie ontvouwen vertoonde de oplossing nog steeds aanzienlijke stress-waarden; 0,1572 voor DHAT en 0,3145 voor DSTAR. Weliswaar is deze stress-waarde grotendeels te wijten aan de bijdrage van één ideaalpunt, namelijk het punt 13 had in de eendimensionele oplossing een stress-waarde van 0,6879, terwijl alle overige punten in deze oplossing ten hoogste 0,047 vertoonde. Gezien het aantal individuen met dit ideaalpunt (52 personen is 4,8 % van de oorspronkelijke populatie) leek het niet verantwoord dit punt uit de populatie te verwijderen. Als conclusie moet derhalve gelden dat ontvouwen in één dimensie van de 4 uitspraken niet zinvol is.

Ontvouwen in twee dimensies levert een configuratie op, welke is weergegeven in figuur 4 (zie volgende pagina). De ontvouwing in twee dimensies levert weliswaar extreem lage stress-waarden op (0,004 voor DHAT en 0,009 voor DSTAR), echter in de configuratie blijken de stimuli 3 en 4 tot elkaar te naderen terwijl de positie van de eerste twee stimuli nagenoeg gelijk blijft. Kennelijk is de horizontale dimensie stabielere dan de verticale dimensie.

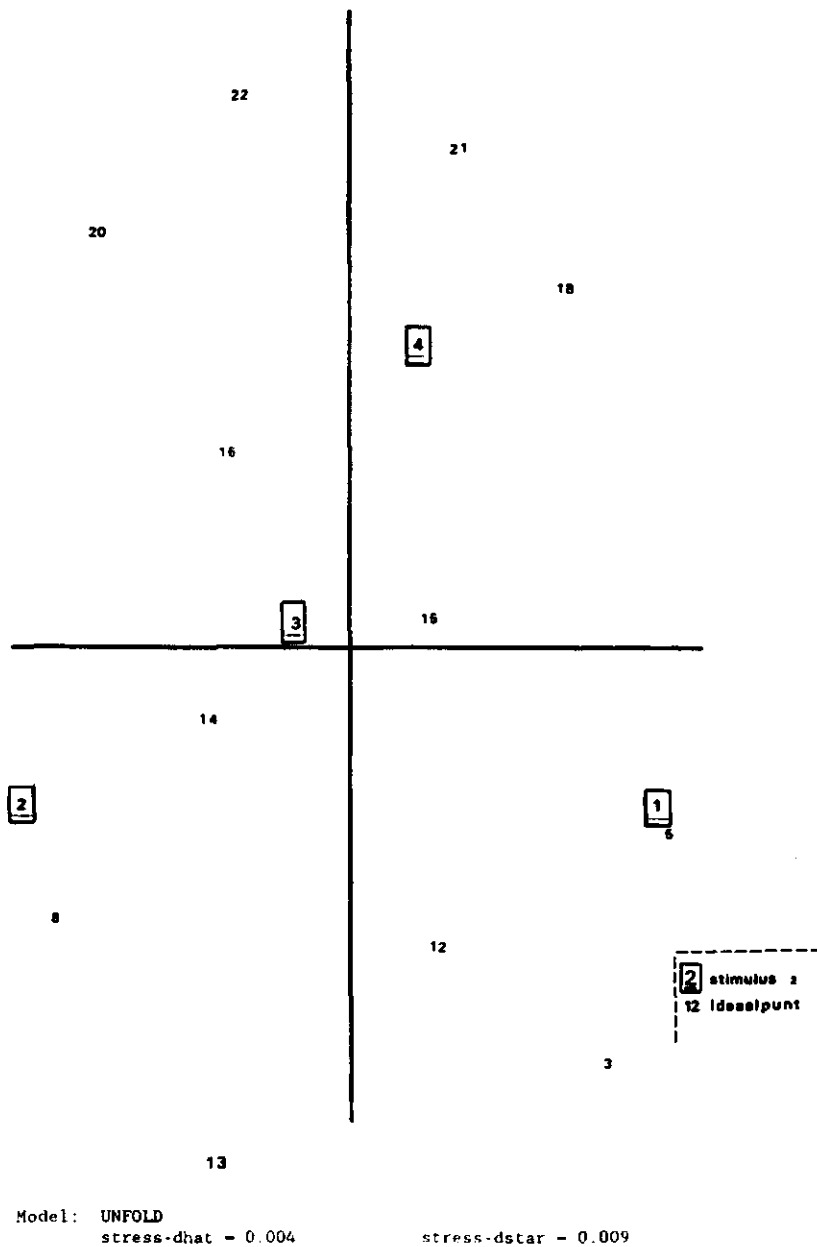
Conclusies ten aanzien van de gebruikte techniek

Concluderend kan gesteld worden dat bij de beoordeling van de 4 uitspraken van de fietsende recreanten twee dimensies een rol spelen. Hiervan is één dimensie duidelijk benoembaar: het gaat om de mate waarin de fietsende recreanten aan veel respectievelijk weinig mensen op een weg/pad de voorkeur geven.

Overigens blijkt, wanneer de populatie in 5 groepen wordt ingedeeld naar de ligging ten opzichte van de eerste dimensie, dat er een samenhang is tussen deze groeps-indeling en het oordeel over het aantal fietsen op het traject; namelijk de recreanten met een voorkeur voor weinig mensen op een weg/pad zeggen relatief vaker dat zij het aantal fietsen op het traject te veel vinden dan recreanten met een voorkeur voor veel mensen op een weg/pad (Klinkers en Van Hoorn, 1987). Een en ander duidt er op dat de eerste dimensie een redelijke (interne) geldigheid biedt.

De tweede dimensie is moeilijk eenduidig te interpreteren. Een moeilijkheid bij het benoemen van met name de tweede dimensie, is het (te) geringe aantal stimuli. Bij meer stimuli (dan 4) zijn er meer gegevens beschikbaar, hetgeen de interpretatie vergemakkelijkt.

Tevens moet hierbij worden opgemerkt dat het moeilijk interpreteren van een dimensie ook zijn oorzaak kan vinden in het model zelf. De veronderstelling die aan het ontvouwingsmodel ten grondslag ligt is, dat alle respondenten cognitief (d.w.z. in kennend opzicht) gelijk tegen de stimuli aankijken, doch dat alleen de evaluatie van de stimuli verschilt. Bijvoorbeeld ingeval het eerdergenoemde voorbeeld met de politieke partijen is een noodzakelijke voorwaarde dat alle respondenten evenzeer op de hoogte zijn van de mate van progressiviteit en/of confessionaliteit van een politieke partij. In de praktijk blijkt dit niet altijd het geval te zijn. Om aan dit soort



Figuur 4. Twee-dimensionale ontvouwing van 4 stimule en 1023 individuen (verdeeld over 12 idealepunten), volgens de city-blok metriek

bezwaren tegemoet te komen is het gewogen ontvouwingsmodel ontwikkeld, dat er kortweg gesproken vanuit gaat dat de ideaalsituatie van een respondent niet als één ideaalpunt in een n-dimensionale ruimte kan worden voorgesteld, maar veeleer als een 2-dimensionaal vlak in deze ruimte. Afhankelijk van het moment ligt het ideaalpunt van de respondent ergens in dit vlak. Het grote voordeel van dit gewogen ontvouwingsmodel is dat het minder stringente eisen aan de data stelt; het biedt als het ware aan de respondent een grotere onzekerheidsmarge, hetgeen duidelijk meer in overeenstemming is met reële keuze-

situaties. De relevante dimensies komen hierdoor duidelijker naar voren dan bij het gewone, ongewogen ontvouwingsmodel. Een nadeel van het gewogen ontvouwingsmodel is, dat er grotere aantallen respondenten en stimuli vereist zijn, omdat als het ware het informatiegehalte per respondent gedaald is (vanwege de zojuist genoemde onzekerheidsmarge). In het onderhavige voorbeeld is het stimulusaantal (4 uitspraken) te gering om dit model te kunnen toepassen.

7 CONCLUSIES T.A.V. VERDERE TOEPASSINGSMOGELIJKHEDEN

Ontvouwen (één- of meerdimensionaal) is een analysetechniek die ten behoeve van recreatie-onderzoek goed toepasbaar is. Het biedt mogelijkheden om inzicht te krijgen in de factoren of dimensies die bepalend zijn voor de reacties van recreanten en/of hun opvattingen over recreatieve zaken.

Veel recreatie-onderzoek verkeert nog in een verkennend stadium; de onderzoeker is van tevoren niet op de hoogte van de factoren die meespelen bij een bepaalde vorm van recreatief gedrag, juist daarover wil hij door middel van onderzoek meer inzicht verkrijgen. Het ontvouwingsmodel is voor verkennend onderzoek zeer geschikt. Het biedt de mogelijkheid te zoeken naar een categorieënsysteem dat ook aansluit bij de belevingswereld van de recreant, en niet alleen bij de belevingswereld van de onderzoeker (zoals in veel onderzoek helaas het geval is).

Specifiek biedt het ontvouwingsmodel de mogelijkheid om preferenties te analyseren, een situatie die in recreatie-onderzoek regelmatig aan de orde is. Het model komt daarbij het beste tot zijn recht indien grotere aantallen stimuli gebruikt worden.

In bovenstaand voorbeeld is het model vooral gebruikt als schaalmodel voor individuen; daarnaast en met name is het ontvouwingsmodel geschikt als schaalmodel voor stimuli. Wat dit laatste betreft zijn er ook toepassingen denkbaar buiten het openluchtrecreatie-onderzoek, b.v. in sociaal-geografisch onderzoek, waarin ook vaak sprake is van het analyseren van de dimensies die aan voorkeuren ten grondslag liggen.

Het computerprogramma UNFOLD voldoet goed voor toepassing van het ontvouwingsmodel; het is gebruiksvriendelijk zowel wat betreft de aansturing van het programma als wat betreft de outputmogelijkheden.

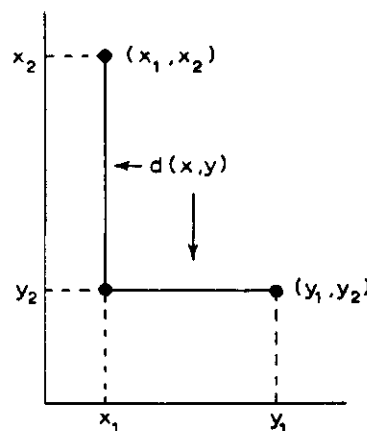
LITERATUUR

- Bennett, J.F. en W.L. Hays, 1960. Multidimensional unfolding: determining the dimensionality of ranked preference data. Psychometrika, 25 no. 1.
- Klinkers, P.M.A. en A. Van Hoorn, 1987. Drukbeleving van fietsende recreanten. Wageningen, I.C.W.-nota 1805.
- Roskam, E.E.Ch.I., 1983. Methodenleer, collegediktaat Methoden-II Studierichting Psychologie. Nijmegen.
- Thissen, M., 1982. Input-output beschrijving UNFOLD. Nijmegen, Groep Rekentechnische Dienstverlening van het Psychologisch Laboratorium.

BEGRIPPENVERKLARING

Centroïde het gemiddelde van de coördinaten van de punten in een figuur

City-blok metriek afstandsmodel, volgens welk de afstand tussen twee punten gelijk is aan de som van de afstanden op de afzonderlijke coördinaatassen. Men noemt dit het city-blok model, omdat deze afstand overeenkomt met de afstand die men moet lopen om van x naar y te komen in een stad met uitsluitend rechthoekig op elkaar staande straten. (zie onderstaande figuur).



Dimensie een ordening van stimuli naar een eigenschap die zij in meerdere of mindere mate bezitten. B.v. de dimensie donkerheidsgraad van de kleur ogen. Een willekeurige groep mensen kan worden ingedeeld naar de mate waarin zij een donkere kleur ogen bezitten.

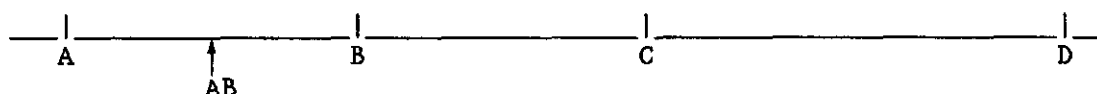
Euclidische metriek afstandsmodel volgens de stelling van Pythagoras, die inhoudt dat in een rechthoekige driehoek het kwadraat van de schuine zijde gelijk is aan de som van de kwadraten van de rechthoekszijden.

Dominantie-metriek afstandsmodel, waarbij de mate van ongelijkheid tussen stimuli gelijk is aan het grootste verschil op enige dimensie. De dominantie-metriek zegt in feite dat de afstand tussen x en y bepaald wordt door die dimensie waarop de coördinaatwaarden het meest verschillen. Met andere woorden, deze dimensie domineert alle anderen. De dominantie-metriek wordt weinig toegepast.

BIJLAGE 1: Het bepalen van het aantal toegestane rangordeningen (I-schalen) vanuit de algemene dimensie (J-schaal).

Omwille van het voorbeeld gaan we uit van een ontvouwing in één dimensie van 4 stimuli.

Gesteld dat we de J-schaal kennen; dit is de algemene dimensie, die op alle individuen van toepassing is. Deze zou er als volgt uit kunnen zien: kunnen zien:



Elke stimulus (weergegeven met de letter A, B, C of D) neemt op de dimensie een bepaalde plaats in. Stimuli die op deze dimensie veel met elkaar gemeen hebben, zijn dicht bij elkaar gesitueerd. Ook individuen kunnen op deze wijze op de schaal worden geplaatst, afhankelijk van de voorkeur die ze hebben voor de verschillende stimuli.

Het pijltje op de schaal vormt het exacte midden tussen A en B. Individuen die links van dit kruisje een plaats innemen op de schaal, hebben een voorkeur voor stimulus A boven B; ze zijn immers dicht bij A gesitueerd dan bij B. Omgekeerd hebben individuen die rechts van het kruisje een plaats innemen op de schaal, een voorkeur voor stimulus B boven A; ze zijn immers dicht bij B gesitueerd dan bij A. Het is makkelijk in te zien dat tussen de 4 stimuli in totaal 6 middenpunten liggen; behalve het reeds genoemde midden AB zijn dit respectievelijk de middens AC, AD, BC, CD EN BD.

Deze 6 middens verdelen derhalve de schaal in 7 gebieden, elk met een bepaalde volgorde van stimuluspunten (rangordening). Dit betekent dat alle individuen die zichzelf op een punt binnen een bepaald gebied plaatsen, dezelfde rangordening van stimuli prefereren. In een steekproef mogen derhalve uitsluitend deze 7 verschillende rangordeningen voorkomen.

In onderstaande figuur is de J-schaal weergegeven met de 6 middelpunten en de daarbij behorende 7 verschillende rangordeningen.

