

Vrije Universiteit Brussel

NIET-METRISCHE STATISTISCHE TECHNIKEN BIJ DE VERWERKING VAN VRAGENLIJSTGEGEVENS, MET EEN TOEPASSING BIJ OPINIE-ONDERZOEK

[NONMETRICAL STATISTICAL TECHNIQUES WHEN PROCESSING QUESTIONNAIRE-DATA, WITH
AN APPLICATION ON OPINION-RESEARCH]

ROLAND PEPERMANS¹

Questionnaires are frequently used in psychology. To deal with their data, classical synthetic statistics, such as factor analysis, cluster analysis, etc. are mostly used. But by using those techniques, one is working on a higher measurement level than those questionnaire-data usually allow. In fact, it would be better to use nonmetric statistics when processing those data. This article shows some basic elements of some non-metric techniques, especially of the multidimensional scaling method MINISSA-I (Lingoes, 1973) which is based on the SSA-I as a variant of the *Smallest Space Analysis* (SSA) (Lingoes, 1965; Guttman, 1968). It treats ranks of similarity measures between the empirical data (ex. gamma-coefficients) to create a spatial configuration of the data in y dimensions. Some notions are given concerning the alienation-coefficient which measures the goodness of fit of our theoretical model. Some aspects of a study on the opinions of labourers and employees about the psychologist are presented. MINISSA-I is used to deal with our data. Seventy-seven items of the questionnaire are represented in an 8-dimensional space (values of the goodness of fit and the distance-scores on the 8 dimensions are shown in Tab. 1 and Tab. 2). From a nonmetric factor analysis NONFACT (Lingoes, 1973), based on the SSA-III (Lingoes, 1966; Lingoes & Guttman, 1967), seven clusters of items are withheld. Labourers and employees of the sample have a clinical image (mental health and psycho-social) of the psychologist, although industrial psychologists are not unknown. There was almost no rejection of the psychologist in the sample.

INLEIDING

PROBLEEMSTELLING

Wanneer men in het psychologisch onderzoek met vragenlijsten werkt, o.a. bij opinie-onderzoek, attitudemetingen, belangstellingslijsten, e.d. dan doet zich veelal het probleem gelden van een massa gegevens overzichtelijk te structureren. Bepaalde klassieke statistische technieken, zoals faktoranalyse, klusteranalyse, e.a. kunnen hier gebruikt worden om een synthetisch beeld van de vragenlijstitems naar voor te brengen. Zij gaan a.h.w. een interne samenhang binnen de reeks

¹ Ik dank hier Prof. J. Borgers en Drs. E. Rosseel van het Centrum voor Organisatie- en Konsumentenpsychologie aan de V.U.B. voor hun steun, hulp en kritische opmerkingen i.v.m. het vermelde onderzoek.

items (variabelen) bepalen en komen aldus tot een groepering der items. Niet-metrische statistische technieken kunnen met hetzelfde doel worden aangewend, maar hebben een andere theoretische basis. Dit artikel zal proberen enkele aspecten van een paar technieken te belichten, zonder echter het voornemen te hebben de gedetailleerde mathematische grondslagen ervan te ontvouwen.

MEETNIVEAUS

Wanneer we synthetische technieken toepassen voor het bepalen van een interne structuur van vragenlijstitems, dan zien we dikwijls over het hoofd dat we werken op niet-metrisch niveau, d.w.z. met nominale (de klassifikatieschaal) of ordinale schalen. Het nominale meetniveau wordt hierbij gekenmerkt door een klassifikatie der variabelen in elkaar uitsluitende klassen. Het ordinale meetniveau situeert zich in de hiërarchie der meetschalen, net boven het nominale meetniveau en wordt gekenmerkt door een orderrelatie tussen de variabelen.

Wanneer we werken op een meetniveau waar we, in tegenstelling tot het nominale en ordinale meetniveau, wel iets weten over de afstand tussen de variabelen, dan begeven we ons op metrisch niveau, t.t.z. op het niveau der interval- en ratioschalen². Het verschil tussen deze beide ligt in het feit dat ratioschalen, naast het afstandsaxioma, ook nog gekenmerkt worden door een nulpunt. Beide schalen komen bij vragenlijsten echter minder frekwent voor. Veelal situeren deze zich op het ordinale niveau. De vijf-, zes-, zeven-, acht- en negen-puntenschalen (om het bij *the magical number seven, plus or minus two* van Miller (1956) te houden) als antwoordmogelijkheden bij vragenlijstuitspraken zullen zeker bekend zijn. Veelal vormen zij duidelijke ordinale meetschalen.

Laat ons het voorbeeld geven van een vijfpuntenschaal: – helemaal akkoord (1); – meer akkoord dan niet akkoord (2); – twijfel (3); – meer niet akkoord dan akkoord (4); – helemaal niet akkoord (5). Stel dat deze vijfpuntenschaal de mogelijke antwoordmogelijkheden op de uitspraken van een vragenlijst voor opinie-onderzoek weergeeft. Wanneer een respondent bv. item 1 beantwoordt met alternatief (2), item 2 beantwoordt hij met alternatief (3) en item 3 met alternatief (4), dan weten we echter niet of de mate waarin de respondent item 1 “verkiest” boven item 2, even “sterk” is als de mate waarin hij item 2 boven item 3 “verkiest”. De alternatieven kunnen wel gekodeerd worden door ze bijvoorbeeld een cijfer van 1 tot 5 toe te kennen, maar die cijfers zijn geen metrische afstandsindices: de afstand tussen bijvoorbeeld alternatief 1 en 3 is niet per se gelijk aan de afstand tussen alternatief 3 en 5.

Dit voorbeeld toont aan dat een vragenlijst waar deze antwoord-

² Gadourek (1972) ziet nog een onderscheid tussen intervalschalen met of zonder beginpunt. Dit onderscheid wordt elders echter weinig of niet gemaakt, zodat ook wij dit niet zullen behandelen.

modaliteiten gebruikt worden (en de variaties zijn schijnbaar eindeloos) van het ordinale niveau zijn. Specifiek aangepaste technieken dringen zich dan ook op.

MOGELIJKE VERWERKINGSTECHNIKEN

De klassieke synthetische statistische technieken (faktoranalyse, klusteranalyse, e.a.), die veelal in de psychologie toegepast werden en worden, werken veelal op het metrische niveau, waar het afstandsaxioma wel geldt (cfr. de interval- en ratioschalen). Wanneer men echter een faktoranalyse toepast op ordinale, of zelfs nominale, gegevens, dan nemen we echter impliciet aan dat de onderlinge afstanden tussen de x alternatieven van een x -puntenschaal bij een vragenlijst gelijk zijn. Daardoor passen we eigenlijk een niet-lineaire monotone transformatie toe (als we o.a. produkt-moment-korrelaties berekenen), omdat de werkelijke afstanden tussen de x schaalwaarden onbekend zijn. Sommige auteurs (o.a. Brand-Koolen, 1972) hebben proberen aan te tonen dat de toepassing van een dergelijke metrische faktor-analytische procedure op ordinale gegevens toch tot bevredigende resultaten leidt. Doch de mathematische zekerheid wordt volgens ons bij dergelijke werkwijzen toch enigszins geweld aangedaan. Technieken aan de niet-metrische gegevens aangepast drongen en dringen zich dan ook op. Het sukses van Siegels werk over niet-parametrische toetsen voor de gedragswetenschappen (Siegel, 1956), wijst trouwens in de richting van een duidelijke behoefte aan nonmetrische verwerkingstechnieken. Ook i.v.m. de vragenlijstverwerking liet en laat die behoefte zich voelen. Alhoewel dergelijke technieken reeds enige jaren bestaan, is de bekendheid ervan nog niet tot de meerderheid der psychologen en statistici uit onze kontreien doorgedrongen. Dit artikel zal dan ook een voorbeeld van een dergelijke techniek naar voor brengen, evenals gegevens over een onderzoek waarbij twee niet-metrische technieken werden toegepast.

MULTIDIMENSIONELE SCHAALVERWERKINGSTECHNIKEN

SITUERING

Het analyseren van multidimensionele data, zoals men bij vragenlijsten kan bekomen, via niet-metrische modellen, kan twee objectieven nastreven: enerzijds kan men ze gebruiken voor de beschrijving der data waarbij men een voorstelling wenst te bekomen van de relaties tussen de gegevens en anderzijds kan men proberen de bekomen relaties te verklaren door het bepalen van de onderliggende interne structuur der waargenomen relaties tussen de gegevens.

Algemeen houden de multidimensionele schaaltechnieken in dat men uit de matrix der empirische gegevens een structuur haalt die we via een spatiale voorstelling in een aantal dimensies proberen weer te geven. De niet-metrische technieken vertrekken hiervoor van inter-

stimulus-similariteit-skores, minstens van het ordinale niveau, die in een ruimtelijke metrische structuur worden omgezet, waarbij de afstanden tussen de ruimtelijke punten een omgekeerde rangorde krijgen van de rangorde der similariteit-skores (Stefflre, 1972). Hierbij interpretern we de gelijkheid of similariteit tussen de data als een nabijheids- of afstandsrelatie. De omzetting van de gelijkheid-skores in metrische afstanden gebeurt via een zogenaamde afstandsfunctie (Shepard, 1962a). Dit is een monotone transformatie die resulteert in een ruimtelijk model, waarbij een *goodness of fit*-maat kan bepaald worden, die aangeeft in hoever de configuratie in het metrisch model overeenstemt met de te verwachten reële configuratie achter de data (Cliff, 1972).

Essentieel werken deze non-metrische technieken met n objecten $A_1, A_2, \dots, A_n$ zoals personen, ideeën, vragenlijstitems, e.a. De empirische gegevens die men dan analyseert zijn de reeds genoemde gelijkheid-skores (of koëfficiënten van overeenkomst) zoals korrelatiekoëfficiënten, gamma-koëfficiënten, voorkeurskores, e.a. Dergelijke impliciete afstanden moeten dan omgezet worden in expliciete (Euklidische) afstanden. In dit verband spreekt men ook van *proximity measures* en van *analysis of proximities* (Shepard, 1962a). De geometrische configuratie die men aldus bekomt, moet echter een zo laag mogelijke dimensionaliteit hebben, maar moet anderzijds ook zoveel mogelijk de waargenomen relatie tussen de empirische afstanden behouden. Dit betekent dat de monotoniteit tussen de similariteit-skores van de oorspronkelijke n objecten en de afstanden tussen de ruimtelijk voorgestelde punten, zoveel mogelijk moet blijven gelden. Bovenvermelde gegevens zijn een algemene situering van de methode.

In de volgende paragraaf zullen we verder ingaan op de techniek die wij hanteerden. Immers, verschillende multidimensionele schaalverwerkingstechnieken hebben zich ontwikkeld. Een eerste evolutie situeerde zich op het metrische niveau, waarbij naar steeds fijnere vormen en algoritmen van de bestaande berekeningswijzen werd gezocht. Anderzijds, bij de evolutie op niet-metrisch niveau, was het Shepard (1962a, 1962b) die met een eerste konkrete aanpak naar voor kwam, n.l. de reeds vermelde *proximity analysis*, alhoewel men zeker niet de rol van Torgerson (1958) hierbij mag vergeten. Kruskal (1964a, 1964b) gaat verder die analyse uitwerken door conceptuele en programmeerverbeteringen aan te brengen en door ze uit te breiden naar niet-Euklidische ruimten en afstanden. Hij gaat ook veel verder in de ontwikkeling van een *goodness of fit*-maat, n.l. de *stress*. Hoe kleiner die stress, hoe beter de monotone relatie tussen de ruimtelijke configuratie en de reële psychologische configuratie der n objecten. Doch het bezwaar tegen de methode is dat het resultaat ervan in grote mate afhangt van de uitgangskonfiguratie; ook de behandeling van *ties* is zeker niet optimaal.

Na Kruskal (1964a, 1964b) ontwikkelden zich nog andere in computerprogramma's omgezette algoritmen die proberen om zo efficiënt en robuust mogelijk de verwerking der psychologische data aan te

CLINIQUE
Rue St-Jacques
LIÈGE

pakken. In dit verband kunnen we, louter exemplatief, naast de MD-SCAL van Kruskal (1964a, 1964b) TORSCA van Young en Torgerson (1967) vermelden, evenals NMSPOM van De Leeuw (1969) en de GUTTMAN-LINGOES NONMETRIC PROGRAM SERIES van Guttman en Lingoës (o.a. Guttman, 1968; Lingoës, 1973) waarin verschillende algoritmen worden uitgewerkt.

ENKELE THEORETISCHE BASISPRINCIPES VAN SSA, SSA-I EN MINISSA-I

De techniek die wij verder zullen toepassen, MINISSA-I, steunt enerzijds op de SSA-I als uitloper van de *Smallest Space Analysis* (SSA) van Guttman (1968) met o.a. aanpassingen door Lingoës (1971) en Lingoës & Roskam (1973) en anderzijds (maar in mindere mate) op de reeds genoemde MD-SCAL van Kruskal. MINISSA-I verenigt de voordelen van beide programma's tot een geheel dat snel en goed de *ties* behandelt. Het streeft daarbij, evenals de SSA, naar het voldoen aan twee voorwaarden: enerzijds de monotoniteitsvoorwaarde en anderzijds moet men tot een voorstelling komen in een zo klein mogelijk aantal dimensies (m). De afstanden die daarbij in het model gehanteerd worden zijn Minkowski-afstanden en hebben *city block* en Euclidische afstanden als bijzondere metrische gevallen. De *goodness of fit*-maat die specifiek is voor de SSA en voor de MINISSA-I, is de aliënatiecoëfficiënt, alhoewel bij MINISSA-I volgens het Lingoës-programma (1973), ook Kruskals stress (1964a) wordt berekend. Hoe kleiner die aliënatiecoëfficiënt is, hoe beter de gevonden ruimtelijke configuratie overeenkomt met de impliciete structuur van de oorspronkelijke objecten. Toch is het zo dat de aliënatiecoëfficiënt (evenals de stress) slechts afneemt tot een voor de gegevens specifiek minimum is bereikt om dan weer te stijgen.

Voor een zeker y aantal dimensies gaan we de ruimtelijke punten zodanig "itereren" tot we de aliënatiecoëfficiënt voor die y aantal dimensies zo klein mogelijk hebben. Daarna wordt bij $y+1$ dimensies hetzelfde gedaan. Aldus wordt gestreefd naar een zo klein mogelijke aliënatiecoëfficiënt, maar anderzijds ook naar een zo klein mogelijk aantal dimensies. Tussen deze twee alternatieven moet dan uiteindelijk door de onderzoeker naar een aanvaardbaar evenwicht worden gezocht.

Om dit te bereiken, kunnen we vertrekken van similariteit-skores tussen de empirische objecten: s_{ij} als gelijkheid-skore voor de objecten i en j . De similariteit-skores die gebruikt kunnen worden zijn velerlei, o.a. Pearson-korrelatiecoëfficiënten, voorwaardelijke kansen, gamma-coëfficiënten, enz.³ Elk object wordt aldus "vergeleken" met alle andere $n-1$ objecten waardoor we $n-1$ *proximity*-metingen krijgen die de impliciete afstanden tussen de waargenomen objecten weergeven.

³ In de toepassing van een multidimensionele techniek die verder in dit artikel zal behandeld worden, zal een gemotiveerde keuze van één der similariteit-skores uiteenzet worden.

De ruimtelijke afstanden tussen de objecten i en j noteren we als d_{ij} . Deze afstand is slechts funktie van de koördinaten behorend tot de overeenkomstige objecten A_i en A_j , waarbij elke A_i m koördinaten x_{ia} heeft met $i = 1, 2, \dots, n$ en $a = 1, 2, \dots, m$. De afstand tussen A_i en A_j wordt dan gegeven door:

$$d_{ij} = d[(x_{i1} - x_{j1})^2, (x_{i2} - x_{j2})^2, \dots, (x_{im} - x_{jm})^2] \quad <1>$$

Hierbij houdt men automatisch rekening met de symmetrie $d_{ij} \equiv d_{ji}$ en met het niet-negatief zijn der afstanden:

$$\begin{aligned} d_{ij} &\geq 0 \\ (i, j &= 1, 2, \dots, n) \end{aligned} \quad <2>$$

We willen hier wel opmerken dat $<1>$ en $<2>$ geen metriciteit bepalen, maar slechts een semi-metriciteit, aangezien de driehoeksongelijkheid door d kan verstoord worden (Guttman, 1968)⁴. Uiteindelijk krijgen we een volstrekt monotone relatie als de rangorde der s_{ij} (similariteiten) precies de inverse is van de rangorde der d_{ij} (ruimtelijke afstanden) of:

$$\begin{aligned} s_{ij} > s_{kl} &\Rightarrow d_{ij} < d_{kl} \\ (i, j, k, l &= 1, 2, \dots, n) \end{aligned} \quad <3>$$

$$\text{en } s_{ij} = s_{kl} \Rightarrow d_{ij} = d_{kl}$$

Onder de genoemde voorwaarde spreekt Guttman (1968) van *strong monotonicity*. Bij psychologische data echter, en vooral bij vragenlijstgegevens, is het veelal zo dat deze volstreekte monotoniciteit mogelijk verstoord kan worden door *ties* in de similariteit-skores. Daartoe verzwakt Guttman (1968) zijn voorwaarden en spreekt van *semi-strong monotonicity*, waarbij de *ties* kunnen ontbonden worden:

$$\begin{aligned} s_{ij} > s_{kl} &\Rightarrow d_{ij} < d_{kl} \\ (i, j, k, l &= 1, 2, \dots, n) \end{aligned} \quad <4>$$

$$\text{en } s_{ij} = s_{kl} \Rightarrow d_{ij} \neq d_{kl}$$

Deze *semi-strong monotonicity* laat toe het aantal dimensies van de schaalverwerking te beperken. Werken we echter op het niveau van de *weak monotonicity*, dan bestaat de mogelijkheid dat $d_{ij} = d_{kl}$ als $s_{ij} \neq s_{kl}$, wat in feite tegen de minimum dimensionaliteit ingaat. De genoemde principes vormen in feite de basis waarop de *Smallest Space Analysis* steunt. Wanneer we verder de SSA toepassen zoals Guttman (1968) ze naar voor brengt (en waarnaar we voor verdere informatie ook verwijzen), dan omvat de eerste benaderende ruimtelijke voorstelling van de empirische gegevens de m latente vectoren (we wensen immers een m -dimensionele structuur te bekomen) van een

⁴ Dit is echter voor onze dimensionaliteitsbepaling geen bezwaar, aangezien die onafhankelijk is van de metriciteit.

gemodificeerde initiële matrix der rangnummers (symbool: D) der afstanden d_{ij} , n.l. $\{D_{ij}\}$, waarbij de rangnummers als niet-negatief worden beschouwd:

$$\begin{aligned} D_{ij} &\geq 0 \\ D_{ii} &= 0 \end{aligned} \quad (i, j = 1, 2, \dots, n) \quad <5>$$

Het blijkt tevens dat de eerste primaire configuratie, die we bij een SSA krijgen, een goed beeld geeft van de uiteindelijke, optimale configuratie. Bij die eerste voorstelling krijgt men ook reeds een idee van de bovenlimiet voor het minimum aantal dimensies, gebaseerd op de latente wortels van de aangepaste matrix $\{D_{ij}\}$. De uiteindelijke oplossing wordt tenslotte bekomen na verschillende modifikaties van de initiële $\{D_{ij}\}$ (Guttman, 1968).

Voor het bepalen van de *goodness of fit*-maat die specifiek is voor de Guttman-Lingoes programma's, n.l. de aliënatiecoëfficiënt (κ) steunt men op de monotonociteitscoëfficiënt μ (Guttman, 1968, p. 478):

$$\mu = \frac{u}{\sqrt{v w}}, \text{ waarbij} \quad <6>$$

$$u = \sum_{i=1}^n \sum_{j=1}^n e_{ij} \hat{D}_{ij} d_{ij} \quad <7>$$

$$v = \sum_{i=1}^n \sum_{j=1}^n e_{ij} \hat{D}_{ij}^2 \quad (i, j = 1, 2, \dots, n) \quad <8>$$

$$w = \sum_{i=1}^n \sum_{j=1}^n e_{ij} d_{ij}^2 \quad <9>$$

waarbij geldt:

$$\begin{aligned} e_{ij} &= 1 \text{ als } D_{ij} \text{ gegeven is voor het koppel} \\ &\quad (A_i, A_j) \quad (i \neq j) \\ &= 0 \text{ als } i = j \text{ of als } D_{ij} \text{ niet gegeven is} \end{aligned} \quad <10>$$

$\hat{D}_{ij}$ zijn de *fitted values* van D_{ij} , via het model. Stellen we $\hat{D}_{ij} = d_{ij}^*$, dan kunnen niet-perfekte *fits* bestudeerd worden (in de *semi-strong monotoniciteitsvoorwaarde*), waarbij d_{ij}^* de *rank-image* van d_{ij} wordt genoemd waarmee wordt bedoeld dat een waarde van $\{d_{ij}\}$ via een permutatie aanleiding geeft tot een element van $\{d_{ij}^*\}$, waarbij de rangorde van $\{D_{ij}\}$ wordt behouden, maar waarbij de *ties* worden opgeheven.

De eigenlijke aliënatiecoëfficiënt wordt dan bepaald via $<6>$:

$$\kappa = \sqrt{1 - \mu^2} \quad <11>$$

In $\langle 11 \rangle$ is $0 \leq \kappa \leq 1$ omdat ook $0 \leq \mu \leq 1$. $\mu = 1$ als de monotonieit behouden wordt, d.w.z. als de rangorde der impliciete afstanden tussen de empirische objekten behouden blijft in de rangorde van de in de analyse berekende ruimtelijke afstanden. In het tegengestelde geval is $\mu = 0$. Dus hoe groter de monotoonheid, hoe groter de waarde van μ wordt en hoe kleiner de waarde van κ . Hoe kleiner κ hoe beter echter ook de ruimtelijke voorstelling van de objekten (o.a. items uit een vragenlijst) in een bepaald aantal dimensies.

EEN TOEPASSING BIJ OPINIE-ONDERZOEK

ONDERZOEKSOPZET

In het kader van een licentieverhandeling in de Psychologische Wetenschappen aan de Vrije Universiteit van Brussel, werden niet-metrische statistische technieken (een multidimensionele schaalverwerking van het type MINISSA-1 en een niet-metrische faktoranalyse NONFACT) toegepast op de verwerking van gegevens bekomen bij een vragenlijst die de bedoeling had te peilen naar de visie van arbeiders, technische bedienden en bedienden op de psycholoog (Pepermans, 1978).

Voor de opstelling van de gehanteerde vragenlijst werden eerst een aantal geleide individuele en groepsinterviews afgenomen, ter exploratie van het onderzoeksterrein. Hieruit resulteerden, na een kwalitatieve, vrije inhoudsanalyse (Gadourek, 1972), 287 uitspraken die nog aan twee vooronderzoeken werden onderworpen. Het doel van die vooronderzoeken was na te gaan in hoever personen uit onze doelpopulatie konden en wilden antwoorden op de items en in hoever alles eenduidig te interpreteren was. Tot slot kregen we dan een gestandaardiseerde vragenlijst met 93 uitspraken. Op deze items moest men, in het uiteindelijk onderzoek, antwoorden door te kiezen tussen vijf mogelijke antwoordalternatieven gaande van "helemaal akkoord" tot "helemaal niet akkoord", zoals die eerder al vermeld werden (zie p. 190). Die vragenlijst werd (samen met nog twee andere technieken, n.l. een status-rangorde en de *Adjective Check List* van Gough, waarop we in verdere artikels hopen te kunnen terugkomen) beantwoord door een steekproef van 113 respondenten uit de doelpopulatie van arbeiders, technische bedienden en bedienden.

VERWERKING DER GEGEVENS

Vooraleer een multidimensionele schaalverwerking toe te passen, werden eerst 16 van de 93 gehanteerde items van verdere verwerking uitgesloten. Dit diende te gebeuren omdat deze uitspraken bijna volledig op de "akkoord"- of op de "niet-akkoord"-alternatieven scoorden. Zij vertoonden weinig variatie in hun antwoorden, waardoor, bij verdere (niet-parametrische) statistische analyses, zij veelal in een ruimtelijke voorstelling een centrale groep zouden vormen die ofwel geen eenduidige interpretatie zou toelaten ofwel enigszins "vervalste"

resultaten zou opleveren. Die items konden enkel geïnterpreteerd worden op basis van hun voor-de-hand-liggende inhoud. Dit betekende dat op die 16 items de psycholoog vooral in een psycho-sociale functie werd gezien, waarbij hij vooral met mensen zou praten, alhoewel hij ook tests zou gebruiken; tevens zou hij in zijn relaties (op die 16 uitspraken) als een stabiel persoon worden gezien.

Onze multidimensionele schaalverwerking vertrok van 77 items (de *n* objecten). De interstimulus-similariteit-skores die we gebruikten, waren gamma-koëfficiënten tussen alle variabelen (Nie et al., 1975). Door het gebruik van deze gelijkheidsindices tussen onze 77 objecten, bekwamen we een vierkante 77×77 datamatrix van onderlinge generale data die het gebruik van de MINISSA-1 verantwoordt (Roskam, 1970a, 1970b). We kozen de gamma-maat boven de Pearson-produkt-moment-korrelatiekoëfficiënt, omdat voor de berekening van deze laatste de relaties tussen de items die we vergelijken lineair moeten zijn, tevens wordt hij door ties verstoord. Voor de Pearson-korrelatie zijn in feite gegevens van het interval-niveau vereist, terwijl wij waarschijnlijk op ordinaal niveau werken, waarbij de gamma-koëfficiënt wel kan toegepast worden zonder restrikties. Deze koëfficiënt is een associatiemaat die tevens de richting van de associatie aangeeft. De maximum- en minimumwaarde is respectievelijk $+1$ en -1 ⁵. De aldus bekomen impliciete afstanden tussen alle 77 items, moesten dan omgezet worden in expliciete ruimtelijke afstanden. We hanteerden hiervoor de reeds genoemde MINISSA-1 zoals het programma in de Guttman-Lingoes serie (Lingoes, 1973) is gegeven, doch enigszins aangepast aan de CDC-komputer. We dienden hiermee een zo klein mogelijk aantal dimensies na te streven, waarbij de monotoonheid zo hoog mogelijk is, dus de aliënatiekoëfficiënt zo laag mogelijk (we kozen liefst lager dan .10)⁶. Voor verschillende aantallen dimensies werd die koëfficiënt berekend (zie Tab. 1). Daaruit bleek dat, na 50 iteraties, we met 8 dimensies een "voldoende goede" monotoonheid bekwamen (verdere dimensionering zou in ons geval waarschijnlijk tē moeilijk interpreteerbare resultaten opleveren).

dimensions	alienationcoefficient	stress
4	.1692	.1614
6	.1230	.1177
7	.1089	.1042
8	.0958	.0920

TAB. 1. VALUES OF THE STRESS (Kruskal, 1964a) AND THE ALIENATIONCOEFFICIENT (Guttman, 1968) FOR DIFFERENT DIMENSIONS.

⁵ Verdere gegevens over deze associatiemaat zou de uiteenzetting onnodig verzwaren en uitgebreid maken. We zullen volstaan met het vermelden van enige referentieliteratuur: Mueller, Schuessler & Costner (1970) en Nie et al. (1975).

⁶ Dit naar analogie van de genormaliseerde stress van Kruskal (1964a, p. 3).

Op basis van de afstandskores op de 8 dimensies van de multi-dimensionele ruimte werden 7 puntenklusters bepaald (voor de afstandskores verwijzen we naar Tab. 2). De gevormde puntenklusters (dit zijn geen factoren!) die we verder zullen vermelden (zie p. 201), geven een interne samenhang tussen de variabelen weer. Zij bepalen a.h.w. een aantal groepen die de structuur van het imago van de psycholoog in de steekproef omschrijven. Voor de afbakening van die groepen is een duidelijk overzicht vereist van de puntenvoorstelling in de verschillende dimensies. Dit kan enerzijds gebeuren door de verschillende voorstellingen van alle mogelijke 2-dimensionele ruimten (vlakken) binnen de 8-dimensionele ruimte te beschouwen. Anderzijds kan men ook de coördinatenstelsels van alle variabelen, voor de 8 dimensies,

TAB. 2. DISTANCESCORES OF 77 QUESTIONNAIRE-ITEMS ON 8 DIMENSIONS. As the reader will see, 16 items of the 93 used in the questionnaire had no distancescore. It were these items that were eliminated before any multidimensional analysis took place, because of statistical reliability (see text).

item	dimensions							
	1	2	3	4	5	6	7	8
1	39.400	6.946	-72.746	19.161	-53.139	-97.922	-40.292	-61.179
2	-32.874	41.410	-60.943	-18.270	-52.021	-74.421	4.509	-22.732
3	49.231	42.726	-11.682	-19.989	-33.171	5.571	-27.642	-100.000
4	50.319	-48.466	-81.657	-18.078	9.009	-56.909	-6.042	-53.085
5	-42.762	-34.537	-8.348	-90.835	-2.625	-12.028	-5.980	5.475
6	80.534	.805	-43.229	-19.852	9.069	-59.136	11.419	-76.531
7	23.861	36.817	7.801	23.808	-10.637	6.851	-24.332	-44.269
9	1.737	-26.884	-57.907	-93.601	13.999	-84.859	-24.201	-16.490
10	45.513	-53.321	-60.403	-51.518	-27.509	-66.847	8.235	-49.201
11	83.167	-3.252	6.077	-34.533	-10.105	-80.652	-15.379	-62.171
12	22.364	-47.394	-59.711	-54.015	-14.106	-63.950	-57.027	-86.895
14	-2.220	28.802	-67.450	-65.280	-15.286	10.388	-36.326	-62.835
15	-56.182	-18.278	-.641	-56.820	-55.637	-7.152	-60.915	-48.076
17	-11.286	-73.469	-55.563	-9.710	-47.410	10.992	-60.452	-46.778
18	-40.510	-61.090	15.659	-3.325	-23.244	-37.106	-47.773	-58.101
19	-38.470	-52.435	-54.836	-.085	-2.891	-30.852	-18.801	-88.345
21	55.281	-1.458	29.031	2.186	29.781	-38.076	14.313	-36.673
22	40.006	-45.244	-27.614	-32.420	8.287	26.404	51.602	-30.715
23	86.665	-74.335	-11.386	-7.085	-56.220	-26.047	16.367	-39.766
24	-64.577	-12.100	3.465	-2.147	-15.840	5.666	2.107	-72.528
25	87.639	-50.953	-36.711	-27.117	-14.290	-29.523	-27.354	-47.036
27	100.000	-11.451	-42.282	-2.336	-61.435	-46.958	-4.581	-47.142
28	-82.679	-8.057	-15.438	-8.571	-19.557	3.593	-22.283	-38.638
29	-14.473	8.338	-36.673	-49.489	-34.882	-56.358	-100.000	-18.152
30	6.714	-79.809	-73.707	-41.449	-14.500	-16.689	-18.610	-62.402
32	57.386	-57.329	-50.987	-6.022	-24.431	-5.546	-28.723	-20.783
33	-37.761	-73.533	9.810	-25.209	-23.835	-7.576	8.616	-45.432
34	26.480	-3.526	-46.705	-31.320	-80.566	-21.253	6.264	-74.432
35	-51.715	-7.127	-18.452	-46.221	-80.632	-38.243	-14.698	-90.818

item	dimensions							
	1	2	3	4	5	6	7	8
36	-33.599	-9.807	-25.359	-38.782	-100.000	-60.967	-15.378	-58.682
37	69.474	-25.423	30.561	-25.610	-39.097	-37.998	-61.535	-50.451
38	-51.104	13.835	-3.560	-20.244	-64.705	-3.379	42.714	-46.997
39	-60.488	-14.045	22.588	27.160	-12.702	-60.580	5.116	-50.076
40	-90.545	-10.149	-41.370	-45.071	12.540	-23.575	4.250	-28.219
42	-60.335	-59.165	-28.310	2.418	-12.426	-6.510	-6.947	-22.054
43	29.389	-51.658	-40.680	-24.655	-23.099	-46.745	.361	21.114
44	-10.284	-91.075	-14.012	-27.933	-5.363	-24.477	-29.535	-85.316
45	-59.241	-28.760	-75.357	-36.610	-43.546	-15.607	-10.297	-35.312
47	-11.108	1.504	-100.000	-.553	-29.906	-16.488	-19.729	-56.067
48	46.408	-30.554	-85.114	-23.124	-63.650	10.284	-4.620	-39.971
49	-.889	-60.175	-63.813	9.742	-38.772	-23.353	-58.481	-74.901
50	68.585	1.245	-61.238	-49.887	-6.798	-17.730	5.816	-73.515
51	-60.293	-22.235	-55.596	-68.379	-23.623	-40.634	33.028	-66.828
52	-50.365	-58.623	-67.828	3.502	-38.423	-58.082	-5.417	-30.213
53	1.967	2.060	-18.313	61.490	-53.692	-38.550	-33.294	-51.735
54	-89.094	-2.239	-62.823	-10.053	-32.912	-27.046	-5.892	-21.884
55	-42.291	-51.769	12.879	35.203	-2.375	-44.290	-35.779	-31.296
56	.593	-85.155	5.226	-36.666	15.256	-25.798	-14.963	-30.327
57	64.436	-54.151	49.491	-28.056	-38.472	-54.442	2.839	-34.947
59	-100.000	-17.602	-41.723	4.362	-10.987	-61.729	-32.420	-47.464
60	-77.364	5.619	-22.219	-10.598	-26.243	-73.328	-7.945	-16.512
61	9.914	16.358	-34.947	52.443	-46.770	-37.434	20.296	-34.331
62	16.356	-100.000	-40.104	-12.663	-41.761	-38.012	-13.476	-40.628
63	-47.128	-62.942	-27.515	2.202	-30.637	-44.945	25.053	5.188
64	-44.650	-57.262	-11.167	-9.178	-96.471	-36.066	1.888	-66.181
65	-68.700	-61.707	-43.933	-13.352	-36.033	-51.929	-45.510	-40.000
66	-74.665	-21.458	-17.726	-38.975	-83.376	-43.889	-2.482	-23.550
67	96.576	18.973	-53.552	5.289	-12.050	-14.829	-7.058	-23.452
68	49.069	-9.990	-33.876	-12.120	-91.827	-41.708	8.945	8.464
69	-67.629	21.523	-46.171	-59.678	-31.812	-5.932	-30.885	-18.871
70	80.026	-26.013	-62.134	-19.219	-34.978	-38.697	-9.730	10.042
71	37.809	-87.900	-48.628	-13.934	-45.709	-82.104	-.500	-26.293
72	46.086	49.291	34.087	-91.138	-17.458	-56.836	-5.566	-45.582
74	-87.814	-35.514	1.238	-56.647	-26.318	-50.039	-12.599	-39.345
76	-34.660	40.725	-21.468	-14.824	28.748	-56.473	-42.710	-2.666
77	74.074	-64.696	-17.563	-50.073	-60.476	-61.650	-37.227	-9.344
78	-87.852	.707	-40.711	-36.956	-59.972	-31.589	-.586	-16.703
79	-68.000	-62.597	-13.559	-59.047	-48.892	-64.577	-1.282	-53.498
80	54.367	-29.434	-33.653	-100.000	-78.592	-18.822	-13.488	-42.569
81	10.797	-80.432	30.260	-35.376	-46.821	10.332	-28.171	-58.276
85	-27.504	-4.609	-68.735	-17.536	-24.416	-68.329	41.476	-91.357
87	-18.248	-44.298	-5.082	-16.982	-3.854	-100.000	-14.403	-79.120
88	84.748	16.895	29.705	-28.471	-39.283	-14.569	-32.162	-15.343
90	6.066	-62.182	39.739	-32.469	-57.017	-58.207	34.485	-67.467
91	-90.725	11.202	10.857	-44.188	-28.458	-62.902	-7.364	-78.782
92	-61.658	-32.635	1.634	1.091	-57.281	-40.062	-48.518	-9.219
93	84.026	11.502	8.535	-21.256	-66.873	-44.325	-37.555	-35.810

in ogenschouw nemen (zie Tab. 2). Er geldt hier dat hoe meer de coördinatenstelsels van twee variabelen i en j met elkaar "overeenkomen", hoe kleiner de metrische afstand in het theoretisch model tussen de punten i en j binnen de 8 dimensies is. De klusters die men aldus kan vormen, kunnen dan achteraf, afhankelijk van de inhoud der items, benoemd worden. Het zal duidelijk zijn dat we geen voorstelling kunnen geven van de 7 klusters die we vormden, omdat we dan een 28-tal 2-dimensionele ruimten zouden moeten konstrueren, wat wellicht een bijkomend artikel zou vereisen.

Bij het vormen en benoemen van de puntengroepen lieten we ons echter ook nog leiden door een niet-parametrische faktoranalyse (Lingoes, 1966; Lingoes & Guttman, 1967). Op deze supplementaire techniek en het doel ervan, zal hierna verder worden ingegaan.

NIET-METRISCHE FAKTORANALYSE : NONFACT

Deze niet-metrische analyse heeft hetzelfde doel als de klassieke faktoranalyse, n.l. het bepalen van factoren die de variabelen waarvan we vertrokken, zo goed mogelijk beschrijven. Die klassieke faktoranalyse zoekt echter naar een ruimte met een minimum aantal dimensies om de korrelaties tussen de items uit te drukken, terwijl de niet-metrische faktoranalyse steunt op de rangorde der korrelaties en deze bij de voorstelling wil bewaren.

De analyse kan goed gebruikt worden als hulp bij, ter controle, of als uitgangspunt van non-metrische multidimensionele schaalverwerkingstechnieken (zie o.a. Daman, 1978; Pepermans, 1978). Verder vermeldt Lingoes (1973) dat deze faktoranalyses best voorbehouden blijven tot *proximity measures* en in 't bijzonder kovariaties en korrelaties. De analyse die wij toepasten ging uit van Pearson-korrelatiecoëfficiënten tussen de 77 variabelen (vragenlijstuitspraken). Hierop werd de faktoranalyse NONFACT (Lingoes, 1973), steunend op de SSA-III (Lingoes, 1966; Lingoes & Guttman, 1967) en aangepast aan de CDC-computer, uitgevoerd. Dit leidde na 2 iteraties tot 43 factoren met een wortel groter dan nul⁷. Hiervan hadden 24 factoren een wortel groter dan de kommunaliteit (= .7985). Op die 24 factoren + 1 faktor werd dan een genormaliseerde Varimax-rotatie uitgevoerd. Voor het benoemen van de uiteindelijke 12 factoren steunden we ons op een criterium dat wel eens bij een parametrische faktoranalyse wordt gebruikt⁸. Volgens Harman (1967) zouden er namelijk minstens 3 variabelen nodig zijn met een hoge lading om een faktor te benoemen. Wij stelden .30 en -.30 als criterium voor het behoren van een variabele tot een faktor (voor de preciese faktorladingen verwijzen we naar het oorspronkelijk werk (Pepermans, 1978)).

⁷ Na één iteratie was de aliënatiecoëfficiënt .0644 en na twee iteraties was het .0628. Verdere iteraties bleken hier niet nodig.

⁸ Voor niet-metrische faktoranalyses, waarover weinig literatuur voorhanden is, bestaan geen dergelijke gegevens. Deze voor de metrische faktoranalyse blijken echter wel te voldoen.

Aldus benoemden we 12 factoren van de 25 die bij de genormaliseerde Varimax-rotatie betrokken waren. We zullen ze echter niet opnoemen in dit artikel, aangezien deze techniek van ondergeschikt en slechts richtinggevend belang was. Algemeen was er wel een overeenkomst met de MINISSA-I-resultaten, doch een volledige konformiteit mag men niet eisen of verwachten, aangezien we o.a. met andere similariteitscoëfficiënten werkten en aangezien de MINISSA-I een afstandsanalyse is en NONFACT een vektorenanalyse.

ONDERZOEKSRESULTATEN ⁹

Uit de genoemde nonmetrische statistische analysetechnieken toegepast op onze gestandaardiseerde vragenlijst konden, zoals reeds gezegd, 7 puntenklusters worden afgeleid:

- *Verwerping*: Deze puntengroep geeft een mogelijke afwijzing en negering van de psycholoog weer.
- *Kaderrol* (in vraag gesteld): Qua inhoud merken we hier een elitaire positie aan de psycholoog toegeschreven met een lichte afwijzing en in vraag stelling van de kaderrol van de psycholoog (dus vooral bedrijfsgericht).
- *Bedrijfsfunctie*: Hier gaat het duidelijk om de adviserende rol van de psycholoog in het bedrijf, met een neutrale ondertoon.
- *Psycho-sociale functie*: De inhoud der items geeft hier een duidelijke neutrale functie aan op het gebied van het sociale en de intermenselijke relaties.
- *Geestelijke gezondheidsfunctie*: Het betreft hier een meer individuele klinisch-psychologische functie i.v.m. zenuwachtigheid, gewetensproblemen, e.d.
- *Psychopathologie*: Men vermeldt hier een diepte-psychologisch (psychiatrisch), zielkundig aspect van de psycholoog.
- *Maatschappelijke afwijzing*: Hier merken we een meer algemene afwijzing, door de maatschappij, van de psycholoog, terwijl de eerste puntengroep een individuele verwerping inhield.

Op basis van de antwoordenverdeling op de items van elke puntengroep kon verder worden nagegaan in hoever onze respondenten al dan niet akkoord gaan met de inhoud der bekomen imago-komponenten. Het bleek dat er vooral een akkoordtendens was op de component geestelijke gezondheidsfunctie, met op het sekundair vlak een instelling gericht op de intermenselijke problematiek (psycho-sociale functie). De relatie met de psychiater ("psychopathologie")

⁹ We zijn er ons van bewust dat we slechts heel oppervlakkig op de eigenlijke inhoudelijke aspecten van het vermelde onderzoek ingaan. We hebben hier zeker niet de bedoeling algemeen het methodologische van een onderzoek te laten primeren, doch in het kader van dit artikel (en enkel in het kader daarvan) leek ons de statistische methodologie toch het belangrijkste. Voor verdere gegevens i.v.m. de inhoud van het onderzoek verwijzen we naar Pepermans (1978).

wordt wel gezien, alhoewel het onderscheid psychiater-psycholoog toch niet zó erg duidelijk is in onze steekproef. De bedrijfsfunctie van de psycholoog lijkt ook weinig duidelijk, alhoewel men die functie wel lijkt te aanvaarden. Algemeen bleek de afwijzing van de psycholoog in onze steekproef weinig op te treden (d.w.z. veelal niet akkoord-antwoorden op de items uit de kluster), alhoewel wel een maatschappelijke afwijzing door onze respondenten werd aangegeven. Of dit laatste al dan niet een projectiefenomeen is, zou verder onderzoek moeten uitwijzen. Verder onderzoek is ook zeker noodzakelijk op het gebied van de extrapolatie van onze resultaten, bekomen op een willekeurige steekproef van arbeiders, technische bedienden en bedienden, naar de totale doelpopulatie.

SLOTBESCHOUWINGEN

We hopen met deze gegevens een ietwat breder perspectief geopend te hebben van de mathematisch zekerder verwerking van niet-parametrische (ordinale) vragenlijstgegevens via meer aangepaste statistische technieken. We zijn er ons echter wel van bewust dat we zeker geen volledig overzicht van de niet-metrische technieken hebben gegeven, dat is trouwens ook nooit onze bedoeling geweest. We hebben duidelijk maar enkele technieken summier behandeld en deze technieken waren dan nog aangepast aan onze onderlinge generale data. Uit de door Roskam (1970a, 1970b) ontwikkelde datatheorie blijken echter nog andere psychologische data mogelijk met voor elk duidelijk aangepaste technieken, zowel binnen het afstandsmodel (waarvan wij MINISSA-I toepasten) als binnen het skalair-produktmodel (waarvan wij NONFACT toepasten) (Roskam, 1970b).

We hebben ook zeker niet de intentie gehad om de meer klassieke statistische analyses te verdoemen. Zij zullen in de toekomst nog een belangrijke rol blijven spelen als controle en hulp bij de ontwikkeling en toepassing van niet-metrische statistische technieken (Daman, 1978; Pepermans, 1978). We vonden het echter wel belangrijk enkele voorbeelden van voorhanden zijnde modellen wat meer in het voetlicht te plaatsen en naar een mogelijke toepassing ervan te verwijzen in het kader van de huidige ontwikkeling van de vragenlijsttechnieken, met toepassingen op alle gebieden van de psychologie (o.a. arbeidspsychologie, konsumentenpsychologie, sociale psychologie, enz.).

REFERENTIES

- BRAND-KOOLEN, M. *Faktoranalyse in het sociologisch onderzoek*. Leiden : H.E. Stenfert-Kroese N.V., 1972.
- CLIFF, N. Consistencies among judgements of adjective combinations. In A.K. ROMNEY, R.N. SHEPARD & S.B. NERLOVE (Eds.), *Multidimensional scaling. Theory and applications in the behavioral sciences*. Volume II. New York : Seminar Press, 1972, 163-182.
- DAMAN, J.-P. *Onderzoek naar het begrip "belangenverdediging". Studie uitgevoerd bij*

- de Sint-Niklase bevolking. Onuitgegeven licentieverhandeling, Vrije Universiteit Brussel, 1978.
- DE LEEUW, J. *The positive othant method for nonmetric multidimensional scaling* (stencil). Leiden University: Department of data theory for the social sciences, 1969.
- GADOUREK, I. *Sociologische onderzoekstechnieken*. Deventer: Van Loghum Slaterus, 1972.
- GUTTMAN, L. A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika*, 1968, 33, 4, 469-506.
- HARMAN, H.H. *Modern factor analysis*. Chicago: University of Chicago Press, 1967.
- KRUSKAL, J.C. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964a, 29, 1, 1-27.
- KRUSKAL, J.C. Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 1964b, 29, 2, 115-129.
- LINGOES, J.C. An IBM-7090 program for Guttman-Lingoes Smallest Space Analysis, I, II. *Behavioral Science*, 1965, 10, 193-194.
- LINGOES, J.C. An IBM-7090 program for Guttman-Lingoes Smallest Space Analysis, III. *Behavioral Science*, 1966, 11, 75-76.
- LINGOES, J.C. Some boundary conditions for a monotone analysis of symmetric matrices. *Psychometrika*, 1971, 36, 2, 195-203.
- LINGOES, J.C. *The Guttman-Lingoes nonmetric program series*. Ann Arbor: Mathesis Press, 1973.
- LINGOES, J.C., & GUTTMAN, L. Nonmetric factor-analysis. A rank reducing alternative to linear factor-analysis. *Multivariate Behavior Research*, 1967, 2, 485-505.
- LINGOES, J.C., & ROSKAM, E.E. A mathematical and empirical study of two multidimensional scaling algorithms. *Psychometric Monographs*, 1973, 38.
- MILLER, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 1956, 63, 81-97.
- MUELLER, J., SCHUESSLER, K., & COSTNER, H. *Statistical reasoning in sociology*. Boston: Houghton Mifflin, 1970².
- NIE, N.H. et al. *SPSS: Statistical package for the social sciences*. New York: McGraw Hill, 1975².
- PEPERMANS, R. *Het imago van de psycholoog bij arbeiders en bedienden*. Onuitgegeven licentieverhandeling, Vrije Universiteit Brussel, 1978.
- ROSKAM, E.E. Datatheorie en metrische analyse, I. Theorie der data en hun afbeelding. *Nederlands Tijdschrift voor Psychologie en haar Grensgebieden*, 1970a, 25, 1, 15-54.
- ROSKAM, E.E. Datatheorie en metrische analyse, II. Imperfekte data en de constructie van algorithmen. *Nederlands Tijdschrift voor Psychologie en haar Grensgebieden*, 1970b, 25, 2, 66-82.
- SHEPARD, R.N. The analysis of proximities: multidimensional scaling with an unknown distance function, I. *Psychometrika*, 1962a, 27, 2, 125-139.
- SHEPARD, R.N. The analysis of proximities: multidimensional scaling with an unknown distance function, II. *Psychometrika*, 1962b, 27, 3, 219-246.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. Kogakusha: McGraw Hill, 1956.
- STEFFLER, V.J. Some applications of multidimensional scaling to social science problems. In A.K. ROMNEY, R.N. SHEPARD & S.B. NERLOVE (Eds.), *Multidimensional scaling. Theory and applications in the behavioral sciences*. Volume II. New York: Seminar Press, 1972, 211-243.
- TORGERSON, W.S. *Theory and methods of scaling*. New York: Wiley, 1958.
- YOUNG, F.W., & TORGERSON, W.S. TORSKA: A Fortran IV program for Shepard-Kruskal multidimensional scaling analysis. *Behavioral Science*, 1967, 12, 498.