



Universiteit
Leiden

The Netherlands

Datadozen: Analyse en historie

Kroonenberg, P.M.

Citation

Kroonenberg, P. M. (2005). *Datadozen: Analyse en historie*.

Leiden: Universiteit Leiden. Retrieved from

<https://hdl.handle.net/1887/3494>

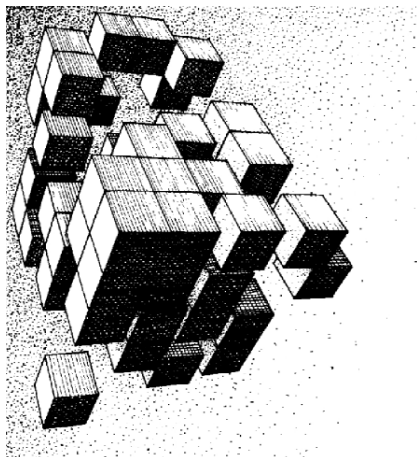
Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3494>

Note: To cite this publication please use the final published version (if applicable).

Datadozen: Analyse en historie



Rede uitgesproken door

Prof. dr. P.M. Kroonenberg

bij de aanvaarding van het ambt van bijzonder hoogleraar
in de multivariate analyse, in het bijzonder van driewegdata
aan de Universiteit Leiden
op vrijdag 2 september 2005

Illustratie: Siep Kroonenberg (1983)

*Meneer de Rector Magnificus,
Leden van het bestuur van de Stichting ter Bevordering van de Wetenschap der
Datatheorie,
Leden van het Curatorium van deze bijzondere leerstoel,
Zeer gewaardeerde toehoorders,
including those of you who are unfortunately not able to understand Dutch, and who
will have to content themselves with enjoying this special occasion by looking at the
pictures in the hand-out and by inspecting these august surroundings.*

Als jong medewerker was ik ooit op bezoek bij een collega in het VU ziekenhuis, alwaar ik mij vrolijk maakte over een bordje naast een deur waarop stond dat in die kamer een Lector in de biochemie van de mondholte huisde. Op eenzelfde manier kan ik mij voorstellen dat U zich vrolijk maakt over het esoterisch gehalte van mijn leeropdracht. "Multivariate analyse, in het bijzonder van driewegdata". Multivariate analyse, oftewel het analyseren van meer variabelen tegelijk dat kan nog net en wordt ook veelvuldig toegepast in vele takken van wetenschap, maar "in het bijzonder van driewegdata" klinkt in Uw oren wellicht zoals destijds "van de mondholte" in de mijne klonk. Is het nu echt nodig om voor zo iets een leerstoel in het leven te roepen?

In deze rede zal ik proberen U ervan te overtuigen dat mijn specialisme lang niet zo esoterisch is als het U misschien lijkt. Meer in het bijzonder zal ik betogen dat de onderzoeksweld niet zonder drieweganalyse kan, ook al heeft nog niet iedereen dat in de gaten. Bovendien zal blijken dat driewegdata zich overal manifesteren als men er maar oog voor heeft. Gelijk jonge aanstaande ouders plotseling overal zwangere vrouwen ontwaren.

Voordat ik er toe over ga U uit te leggen waar het nu allemaal over gaat, wil ik nog iets zeggen over deze zaal en de beperkingen die deze oplegt aan mijn rede. In het huidige tijdsgewricht kan een beetje wetenschapper al haast niet meer zonder een gelikte visuele presentatie met allerlei bewegende beelden en spannende geluidseffecten. Echter, de Leidse traditie heeft het, waarschijnlijk gedeeltelijk geïnspireerd door deze middeleeuwse kloosterzaal, dat zulks bij inaugurale redes uit den boze is. Voor een vak als het mijne is dit zowel een handicap als een uitdaging. Veel multivariate analyse steunt in belangrijke mate op het presenteren van gegevens en uitkomsten in de vorm van grafieken en ruimtelijke afbeeldingen, en die kunnen nu juist zo mooi via lichtbeelden getoond worden. Als tussenoplossing heb ik U een boekje ter hand laten stellen met illustraties behorende bij mijn betoog.¹

In deze rede zal ik U eerst uitleggen wat driewegdata zijn en hoe ze aangepakt worden. Met een gedetailleerd voorbeeld hoop ik U een gevoel te kunnen geven waarom drieweganalyse mij zo fascineert. Vervolgens zal ik U iets vertellen over de oorsprong van drieweganalyse, de stormachtige ontwikkeling in de laatste tien jaren en wat ik in de komende vijf jaar hoop te doen in het kader van deze leerstoel. Het derde en laatste deel van mijn rede bestaat uit een substantieel dankwoord.

Driewegdata en hun analyse

Wat zijn driewegdata?

Heel kort gezegd, zijn driewegdata gegevens die niet meer op één systeemkaartje passen, maar een hele doos nodig hebben. Wellicht vindt U dat niet verhelderend, maar mijn eerste illustratie kan U hierbij helpen. Als we even aansluiten bij de pedagogische wetenschappen, zien we dat vele databestanden eruit zien als de scores van kinderen op een aantal variabelen, zoals intelligentie, cijfers voor taal, rekenen en gymnastiek, etc. (zie Figuur 1).

	2000			
	IQ	Taal	Rekenen	Gym
Barry	110	8	9	5
Marie	90	5	6	9
Annie	130	9	8	7
Karel	95	7	6	8
Toos	100	7	8	9
Lea	101	6	5	5
Erik	97	6	6	9

Figuur 1. Tweewegdata

	2000				2001				2002				2003				2004				2005			
	IQ	Taal	Rekenen	Gym	IQ	Taal	Rekenen	Gym	IQ	Taal	Rekenen	Gym	IQ	Taal	Rekenen	Gym	IQ	Taal	Rekenen	Gym	IQ	Taal	Rekenen	Gym
Barry	110	8	9	5	111	7	8	6	110	9	8	5	109	8	7	6	110	9	8	5	111	7	8	6
Marie	90	5	6	9	90	5	6	9	90	5	6	9	90	5	6	9	90	5	6	9	90	5	6	9
Annie	130	9	8	7	130	9	8	7	130	9	8	7	130	9	8	7	130	9	8	7	130	9	8	7
Karel	95	7	6	8	95	7	6	8	95	7	6	8	95	7	6	8	95	7	6	8	95	7	6	8
Toos	100	7	8	9	100	7	8	9	100	7	8	9	100	7	8	9	100	7	8	9	100	7	8	9
Lea	101	6	5	5	101	6	5	5	101	6	5	5	101	6	5	5	101	6	5	5	101	6	5	5
Erik	97	6	6	9	97	6	6	9	97	6	6	9	97	6	6	9	97	6	6	9	97	6	6	9

Figuur 2. Driewegdata

Dit soort gegevens kan men op één systeemkaartje noteren in de vorm van een tabel met de kinderen in de rijen en de variabelen in de kolommen. Er zijn hier dus twee "wegen": kinderen en variabelen. Wanneer de gegevens een onderdeel zijn van een longitudinaal onderzoek, waarbij de kinderen ieder jaar opnieuw met dezelfde variabelen worden gemeten, dan heeft men driewegdata gekregen met als wegen: kinderen, variabelen en als derde weg jaren of meetmomenten. Nu is één kaartje niet meer voldoende, maar heeft men voor ieder jaar een apart kaartje nodig. En een stapeltje kaartjes bergt men op in een doos, of te wel driewegdata passen in dozen, zoals U in Figuur 2 ziet.

Enmaal bekend met het idee, komen we plotseling overal data in dozen tegen. De *plantenveredelaar* heeft diverse variëteiten pinda's (eerste weg) geplant op verschillende locaties (tweede weg) en meet eigenschappen van de geoogste planten, zoals de opbrengst, de kwaliteit, het percentage verzadigde en het percentage onverzadigde oliën (derde weg).² De *chemicus* heeft tien bakken met verschillende suikeroplossingen, stuurt licht van verschillende golflengten door elk van de bakken en meet de verstrooiing van het licht via de uittredende golflengten aan de andere kant van de bak.³ Een *kerkhistoricus* die de polemieken over het modernisme in het midden van de 19^e eeuw wil beschrijven kan een doos maken met polemisten (eerste weg), hun tegenstanders (tweede weg) en gebruikte argumenten (derde weg).⁴ Een *medische farmaceut* heeft de ruimtelijke structuur van een aantal varianten van cocaïne afgeleid uit kris-

tallografische metingen en wil weten in hoeverre de ruimtelijke structuren van deze cocaïnemoleculen overeenkomen, en in hoeverre ze verschillend zijn. De moleculen vormen de eerste weg, hun constituerende atomen de tweede en de ruimtelijke coördinaten de derde.⁵

Gewaardeerde toehoorders, misschien heeft u zich wel eens afgevraagd hoe een zendmast die niet alleen van uw mobieltje signalen ontvangt maar ook van talloze andere mobieltjes, precies weet dat het antwoord op uw ontboezemingen bij uzelf terecht moet komen en niet bij uw buurman of buurvrouw die dit antwoord maar beter niet kan weten? Ook dit proces kan via driewegmodellen beschreven worden.⁶

Wat is "Multivariate analyse, in het bijzonder van driewegdata"?

Drieweganalyse is niets anders is dan de analyse van data die in dozen moeten, maar daar komt U niet veel verder mee. Het is wellicht beter de vraag te herformuleren (een bekende truc van statistici als ze de oorspronkelijke vraag niet kunnen of willen beantwoorden). Eén mogelijkheid is: "Welk type onderzoeksvragen kan met drieweganalyse aangepakt worden?" Een dergelijke vraag heeft het voordeel dat hij geformuleerd is in termen van het inhoudelijk onderzoek in plaats van puur in termen van de techniek zelf.

Laten we het voorbeeld van de kinderen die een aantal jaren worden gevolgd weer bij de kop nemen. Welke vragen zou de onderzoeker in gedachten hebben gehad toen zij haar gegevens ging verzamelen? Er zijn natuurlijk vele mogelijkheden, maar in deze rede wil ik me beperken tot die vragen die aangepakt kunnen worden met drieweganalyse. De *kernvragen* in dit geval zouden kunnen zijn:

1. Hoe hangen de variabelen onderling samen?
2. Welke trends zijn er te ontdekken over de tijd?
3. Zijn er verschillende typen kinderen?

Drie vragen, één voor elke weg. Zulke vragen zijn weliswaar interessant maar ze hebben slechts betrekking op één weg tegelijk. Drieweganalyse is met name in het leven geroepen om ingewikkelder vragen te beantwoorden, zoals:

4. Verandert de samenhang tussen de variabelen over de tijd? Het is bijvoorbeeld bekend dat bij heel jonge kinderen intelligentie nog erg ongedifferentieerd is, maar dat met het ouder worden sommige kinderen zich beter ontwikkelen op bepaalde aspecten van intelligentie dan andere. Met andere woorden, over de tijd heen verandert de structuur oftewel de onderlinge samenhang van de onderdelen van een intelligentietest.
5. Een nog complexere vraag is: Verandert de structuur van de variabelen over de tijd op een andere manier voor verschillende groepen kinderen, bijvoorbeeld voor jongens en meisjes, of voor kinderen die in meer of minder ernstige mate verstandelijk gehandicapt zijn?

Bij dergelijke complexe vragen, waarbij alle drie aspecten of wegen van de data betrokken zijn, komt drieweganalyse pas goed tot zijn recht. *Plantenveredelaars* zijn bijvoorbeeld geïnteresseerd in specifieke aanpassing van gewassen, dat wil zeggen zij

willen graag weten welke variëteiten van een gewas het met name goed doen op bepaalde eigenschappen in locaties met specifieke kenmerken. In concreto, waar kun je het beste welke soort pindaplanten laten groeien zodat er pinda's verkregen worden die met name geschikt zijn om pindakaas van te maken? Richten bepaalde scholen *modernistische theologen* zich tegen specifieke andere scholen, gebruikmakend van bepaalde typen argumenten?

De eerste kennismaking met technieken voor drieweganalyse is vaak even schrikken vanwege de complicaties die het begrijpen en het juist interpreteren van die technieken met zich mee brengt. Maar dat is onvermijdelijk aangezien het hier gaat om complexe technieken die bedoeld zijn om complexe vragen te beantwoorden. Voor het toetsen van het verschil in lengte tussen jongens en meisjes draait een beetje pedagoog zijn hand niet om, maar driewegvragen zijn gewoon complexer. Echter die complexiteit geeft een onderzoek meer diepgang.

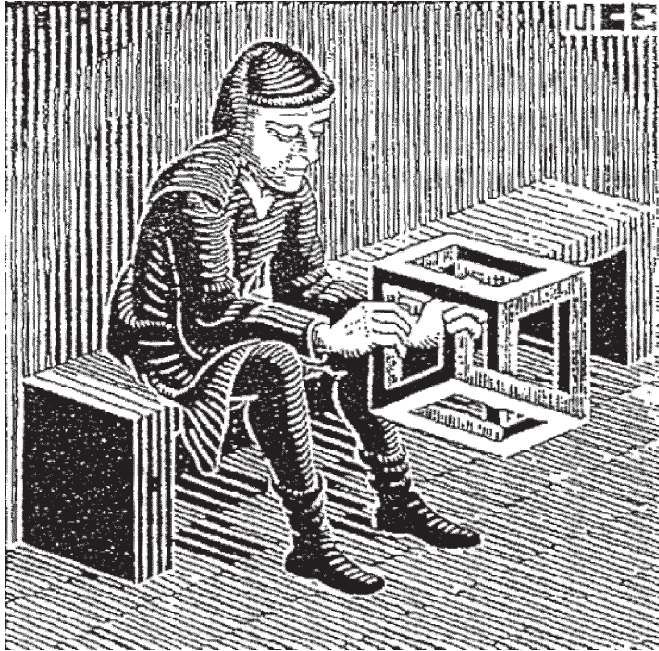
Ik sta iedere keer weer verrast hoe vaak onderzoekers zich tevreden stellen met univariate antwoorden in situaties die inherent multivariaat zijn. Geen wonder dat Jan Blokker⁷ ooit zei dat de meeste antwoorden uit sociaal-wetenschappelijk onderzoek al bekend waren bij het kruidenvrouwtje uit Stroe, en dat velen bij het lezen van de uitkomsten van sociaal-wetenschappelijk onderzoek denken: "Maar dat wisten we toch allang?" De adder die zich hier onder het gras bevindt, is dat, als de uitkomsten anders waren uitgevallen, andere mensen of eventueel dezelfde óók hadden gezegd dat ze dit allang wisten.

Als bewijs hiervan mag een truc van Prof. Köbben⁸ dienen, die in een rede een aantal resultaten van onderzoek naar het gedrag van allochtonen presenteerde, waarop in de zaal goedkeurend en bevestigend geknikt werd. Even later keerde hij terug tot dit onderwerp. Hij vroeg de zaal of men hem werkelijk geloofd had, want in werkelijkheid waren de resultaten geheel tegengesteld aan wat hij eerder beweerd had.

De kern van de zaak is dat het multivariate karakter van verschijnselen het mogelijk maakt dat in de ene omstandigheid het ene univariate resultaat behaald wordt, terwijl in een andere omstandigheid het tegenovergestelde gevonden wordt. Kortom, univariate resultaten zijn vaak misleidend in een multivariate wereld, en het multivariate karakter moet doorgrond worden door technieken die deze complexiteit recht doen, zoals technieken uit de gereedschapskist van de driewegdata-analist.

In Figuur 3 ziet U een typisch voorbeeld van zo'n driewegdata-analist die in dit geval met enige vertwijfeling zit te kijken naar de complexiteit van de hem aangeleverde datakubus. Deze houtsnede van M.C. Escher zou overigens prima passen in Jan Luykens bundel uit 1694 met emblemata, *Het Menselyk Bedryf*, samen met de Visser, de Bakker en de Smit.

De Driewegdata-analist



Figuur 3. M.C. Eschers "Man met kuboïde".

© 2005 The M.C. Escher Company B.V. Baarn - Holland.

Alle rechten voorbehouden. www.mcescher.com.

Hoe analyseerde men driewegdata voordat er drieweganalyse was?

Kort gezegd: Door de doos plat te slaan of uit te rekken. In beide gevallen is het de bedoeling om van driewegdata tweewegdata te maken door één van de wegen te verdoezelen. In plaats van te kijken naar de interacties tussen drie typen eenheden (of wegen) hoeft men dan immers maar naar twee ervan te kijken.

Platslaan wordt typisch gedaan door gemiddelden te nemen over alle kaartjes in het doosje en één kaartje met gemiddelden over te houden. Dit kunnen gemiddelden zijn over alle tijdstippen zodat men niet meer weet wat de ontwikkelingen over de tijd zijn, maar men kan ook gemiddelden over alle personen nemen, zodat de individuele verschillen uit het zicht verdwijnen.

Uitrekken doet men door òf alle kaartjes naast elkaar te leggen zodat de verbinding tussen dezelfde variabelen op de verschillende meetmomenten verwaarloosd wordt, òf door alle kaartjes onder elkaar te leggen zodat de verbindingen tussen de scores van de personen op verschillende meetmomenten doorbroken wordt. In alle gevallen worden de gegevens en hun analyse verschrompeld van drieweg naar tweeweg.

Soms kan dit geen kwaad, want het kan voorkomen dat een drieweganalyse leidt tot de conclusie dat men geen drieweganalyse nodig heeft. Bijvoorbeeld als er niets verandert in de tijd, of als alle personen opgevat mogen worden als aselechte trekkingen uit eenzelfde populatie. Of als alle theologen alle tegenstanders in dezelfde mate bestoken met alle relevante argumenten. Zodra dit niet het geval is, leidt het platslaan of uittrekken van driewegdata tot een onnodige en soms onaanvaardbare versimpeling.

Waarmee analyseert men dan driewegdata?

Sinds het begin van de jaren zestig is een serie technieken ontworpen, specifiek gericht op het recht doen aan driewegdata. Zij luisteren naar zulke fraaie namen als driewegprincipale componentenanalyse, driewegfactoranalyse, driewegclusteranalyse, parallelle factoranalyse, meerwegcovariantie-analyse, meerdimensionale schaaltechnieken voor individuele verschillen, Procrustesanalyse, multivariate longitudinale analyse en zo kan ik nog een tijdje doorgaan.

Het gros van deze methoden heeft een sterk exploratief karakter, dat wil zeggen, men zoekt naar de patronen tussen de elementen van de drie wegen zonder a priori patronen te postuleren en zonder toetsen op deze patronen los te laten. Gedeeltelijk komt dat omdat het moeilijk is van te voren zulke patronen te specificeren, en gedeeltelijk omdat toetsen veronderstelt dat men iets weet over de verdelingen van de scores, en dat is bij driewegdata uiterst zelden het geval. Het is wel goed mogelijk om de stabiliteit van de geschatte waarden van de parameters te bepalen via herhaald steekproeftrekken uit de steekproef zelf, maar die ontwikkelingen staan bij drieweganalyse nog in de kinderschoenen.⁹

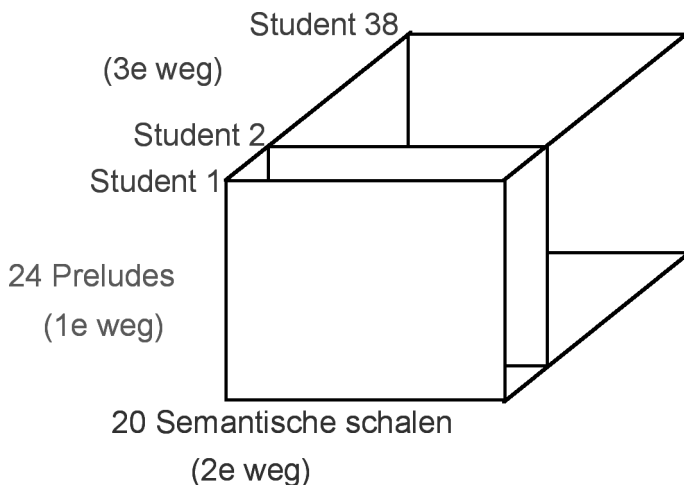
Wat vinden Japanse studenten van de Preludes van Chopin?

Om U enig idee te geven van de aard van de resultaten met driewegtechnieken verkregen, wil ik U graag een analyse voorleggen van gegevens die geproduceerd zijn door Japanse studenten in reactie op het horen van (gedeeltes van) de 24 preludes van de Pool Chopin gespeeld door een Fransman. Het aardige van deze gegevens is dat we hier een huwelijk hebben van psychologie, semantiek, crosscultureel onderzoek, muziek, en statistiek. Dat dit een voorbeeld naar mijn hart is, kunnen sommigen van U beamen die dit voorbeeld in iets andere vorm al eens eerder hebben zien passeren. Verder maakt het voorbeeld gebruik van een type gegevens, semantische differentiaal, dat de aanleiding is geweest tot de ontwikkeling van de drieweganalyse, zoals ik later nog zal toelichten. De analyses zijn tot stand gekomen in nauwe samenwerking met professor Takashi Murakami van de Nagoya Universiteit in Japan.¹⁰

De *onderzoeksvragen* waren als volgt:

1. Muziek kan worden gekarakteriseerd in *muziektechnische* termen zoals toonsoort, tempo en modus, d.w.z. majeur/mineur. In hoeverre zijn *leken* gevoelig voor zulke aspecten van de muziek, wanneer we ze vragen om muziek te beschrijven in

- gewone bijvoeglijke naamwoorden zoals luid, somber, gevoelig, etc.?
- Is het mogelijk om tegelijkertijd via diezelfde adjectieven hun voorkeur na te gaan voor bepaalde typen muziek, niet uitgedrukt in de adjectieven maar in de technische aspecten van de muziek zoals toonsoort, tempo en modus?



Figuur 4. Datadoos voor de Chopinstudie

Aan het onderzoek deden 38 Japanse studenten mee die bekend waren met klassieke muziek (de derde weg). Hen werden de 24 preludes voorgespeeld (de eerste weg), en na iedere prelude (of de eerste 80 seconden ervan) moesten ze twintig zogenaamde semantische beoordelingsschalen invullen (de tweede weg). Een lijstje hiervan is te vinden in Tabel 1. Merk op dat deze schalen bestaan uit twee tegenovergestelde begrippen (onrustig – kalm; snel – langzaam, sterk – zwak, etc.). Bij wijze van oefening mag U zelf oordelen over prelude nummer 22 die door Casper de Jonge zal worden voorgespeeld [*Orgelspel*¹]. Uw antwoorden kunt U de volgende keer inleveren in daartoe bestemde datadoos.

Waar we naar op zoek zijn is de samenhang tussen de preludes en de beoordelingsschalen in relatie met individuele verschillen tussen de studenten. De centrale uitkomst van het onderzoek was dat de studenten het grotendeels eens waren over de technische, cognitieve aspecten van de preludes, maar dat ze verschilden in de affectieve aspecten, namelijk het soort muziek waar hun voorkeur naar uitging. Dit zal ik nu toelichten.

Tabel 1. Semantische differentialen
 (Nederlandse vertaling van de Engelse
 vertaling van de Japanse termen; scores van 1 t/m 7)

Kalm (= 1)	-	Onrustig (= 7)
Lieflijk	-	Streng
Stil	-	Luidruchtig
Lyrisch	-	Dramatisch
Sereen	-	Heftig
Zwak	-	Sterk
Langzaam	-	Snel
Ingehouden	-	Schreeuwerig
Licht	-	Zwaar
Vrolijk	-	Somber
Helder	-	Donker
Teer	-	Robuust
Gelukkig	-	Droevig
Helder	-	Ondoorzichtig
Warm	-	Koud
Klein	-	Groot
Fijn	-	Grof
Dun	-	Dik
Onaantrekkelijk	-	Aantrekkelijk
Oninteressant (= 1)	-	Interessant (= 7)

Consensus over de muziektechnische aspecten van de preludes

Schalen

Laten we eens kijken naar de *Semantische ruimte* (Plaat I). We zien hier de relaties tussen de 20 bipolaire schalen zoals die uit de oordelen van de studenten naar voren kwamen. Iedere schaal is afgebeeld als een pijl met bij de punt het ene adjectief (*onrustig*) en aan de andere kant het antoniem (*kalm*). Voor de overzichtelijkheid heb ik slechts van een viertal pijlen *beide* polen laten zien. Wat betekent het dat sommige pijlen vlak bij elkaar staan en andere grote hoeken met elkaar maken? Bekijken we eerst de pijlen *onrustig* en *dramatisch*. Deze liggen vlak naast elkaar omdat de studenten over het algemeen preludes ruwweg dezelfde scores geven op deze twee schalen. Dus *onrustige* preludes zijn ook *dramatisch* en *kalme* preludes zijn ook *lyrisch*. Als de pijlen loodrecht op elkaar staan, bijvoorbeeld *snel* en *koud*, dan hebben volgens de studenten de bijbehorende schalen niets met elkaar te maken, oftewel *snelle* preludes kunnen zowel *warm* als *koud* zijn en hetzelfde geldt voor *langzame* preludes.

Preludes

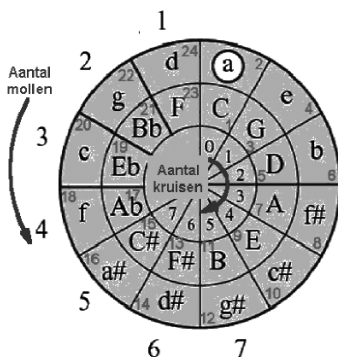
In Plaat II zien we de configuratie van de preludes zoals die voortvloeide uit de scores die de studenten aan de preludes hebben toegekend. Onderaan ziet U de schalen nog

eens afgebeeld, en ten behoeve van de interpretatie moet U de schalen bovenop de preludes in de figuur denken.

We kunnen nu afleiden dat de zojuist door U gescoorde prelude, nummer 22 in G mineur, aangeduid met het tempo *Molto agitato*, met name beoordeeld werd als luid en heftig. Prelude nr. 20 in C mineur, met als tempo *Largo*, vindt U rechtsonder in de Preluderuimte. Deze prelude werd door de studenten als robuust, zwaar, droevig en somber aangemerkt [*Orgelspel*]. Prelude Nummer 7 in A majeur, *Andantino*, linksonder gelegen in de Preluderuimte, werd als kalm, lyrisch en sereen beoordeeld [*Orgelspel*]. Helaas kan ik U geen echt snelle preludes laten horen omdat die niet geschikt zijn voor orgel. In dit verband is het aardig op te merken dat Preludes 4, 6 en 20, die vlak bij elkaar in de Preluderuimte liggen, alle drie op de begrafenis van Chopin op een orgel werden gespeeld.¹²

Technische merites

Bekijken we nu in de Preluderuimte de configuratie van de preludes op hun technische merites, dan zien we dat de studenten een duidelijk indeling hebben gemaakt in tempo (snel en langzaam) en in modus (majeur en mineur). Met andere woorden, op basis van de semantische schalen hebben de studenten een organisatie aangebracht in de preludes die overeen blijkt te stemmen met een indeling op muziektechnische aspecten. Maar dit is nog niet alles, want bij verdere inspectie kan men zelfs, mits muzikaal onderlegd, in grote lijnen de toonsoorten zoals die in de kwintencirkel zitten (zie Figuur 5) terugvinden in de verdeling van de preludes over het vlak. Zelfs twee anomalieën komen zo aan het licht. Dit zijn de nummers 9 en 10, want 9 is in majeur en ligt in het mineurgedeelte, en 10 is in mineur en ligt in het majeuredeelte van de Preluderuimte. Ik ga echter nu niet in op de redenen waarom deze preludes zo afwijkend beoordeeld werden.¹³



Figuur 5. Kwintencirkel.¹⁴

Noot: De nummers aan de buitenkant refereren aan het aantal mollen, die aan de binnenkant aan het aantal kruisen. De nummers in de vakjes komen overeen met de volgorde van de preludes in de partituur.

Waar studenten in verschillen

Wat we nog niet besproken hebben, is waarin de studenten nu *verschillen*. Dit laat ik U niet in een plaatje zien, aangezien dit gemakkelijk in woorden verteld kan worden. De studenten hielden met name óf van snelle stukken in majeur óf van langzame stukken in mineur. Dit had U natuurlijk allang gedacht, maar wees gerust, dit is geen aanzet tot een truc à la Köbben.

Historie van de drieweganalyse

De historie van technieken voor driewegdata-analyse begint bij de recent overleden Ledyard R Tucker, die onder andere aan de University of Illinois en de Educational Testing Service (het Amerikaanse CITO) in Princeton heeft gewerkt. Een andere belangrijke peetvader is Douglas Carroll, voor een groot deel van zijn academische leven werkzaam bij Bell Laboratories en nu verbonden aan de Rutgers University. Een derde grondlegger van driewegtechnieken is Richard Harshman die reeds lang verbonden is aan de University of Waterloo in London, Canada. Ik zal mij in deze rede voornamelijk concentreren op het werk van Tucker, omdat zijn werk de basis is geweest voor vrijwel al het mijne.

Hoewel natuurlijk de wiskunde de moeder aller (exacte) wetenschappen is en de (mathematische) statistiek haar dochter, is het voor een wiskundige wellicht verrassend te zien hoe zo'n ondergeschoven kindje als de psychometrie, de leer van het meten in de psychologie, aanleiding heeft gegeven tot belangrijke statistische technieken en vernieuwingen. Hierbij kan men denken aan factoranalyse, meerdimensionale schaaltechnieken, item-response technieken, covariantiestructuurmodellen, maar ook de drieweganalyse.

Hoe graag was ik niet in het begin van de jaren zestig, toen ik, onwetend en onschuldig, net de middelbare school binnenstapte, een student geweest aan de University of Illinois. Het Departement Psychologie moet een broedplaats van briljante ideeën zijn geweest. Bijna tegelijkertijd werd daar de semantische differentiaal ontwikkeld door Charles Osgood,¹⁵ de gedragsdifferentiaal door Harry Triandis,¹⁶ en de drieweganalyse door Ledyard Tucker.¹⁷ Vele van hun collega's produceerden andere belangrijke bijdragen aan de psychometrie en de psychologie zelf. Zou men over de kennis beschikken van wat er zich daar allemaal afspeelde, dan lijkt mij hier ruim materiaal voor een prachtige academische roman getiteld *Onder Psychometrici*.

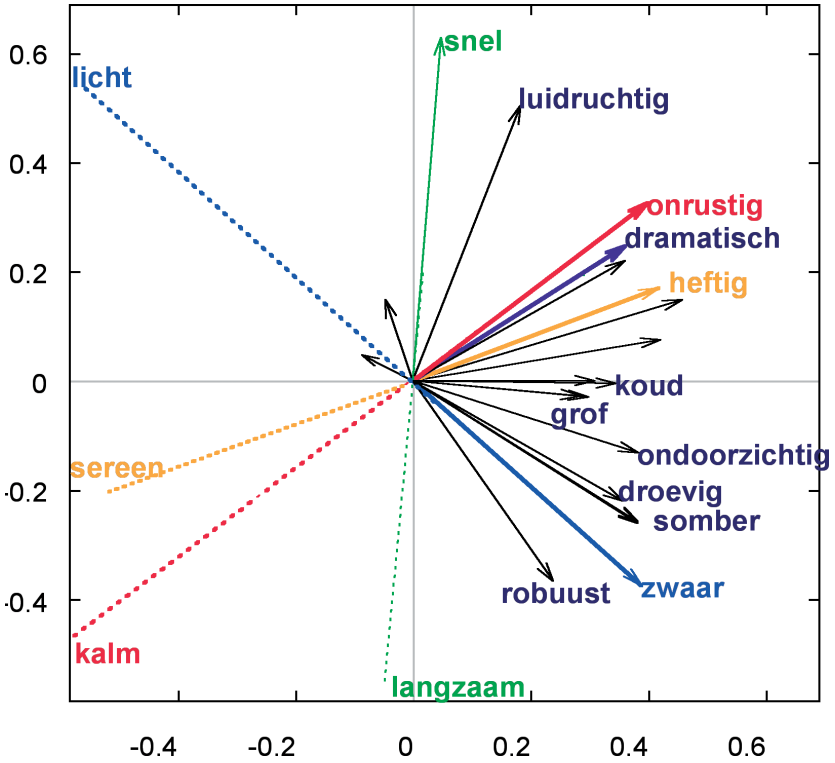
Hoe ontstond drieweganalyse?

In een interview met Neil Dorans in 2004¹⁸ zei Tucker hierover het volgende:

"Driewegfactoranalyse ontstond uit mijn werk aan meerdimensionale schaaltechnieken. Terwijl ik werkte bij de Educational Testing Service, zag ik dat Charles Osgood van de University of Illinois in zijn semantische differentiaal onderzoek driewegdata verzameld had – concepten,

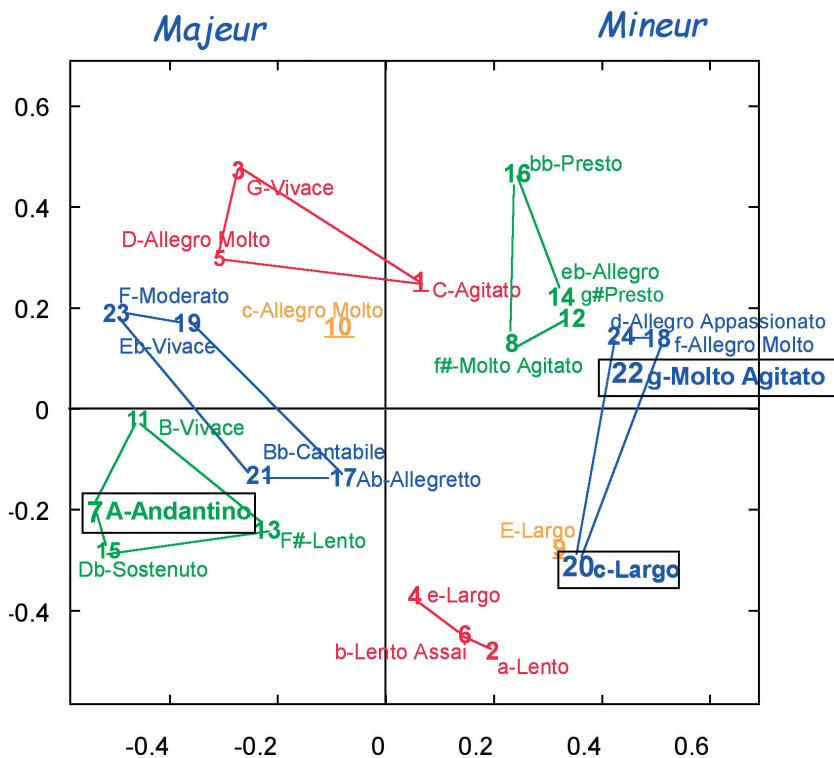
PLAAT I

Semantische ruimte



PLAAT II

Preluderuimte (consensus studenten)

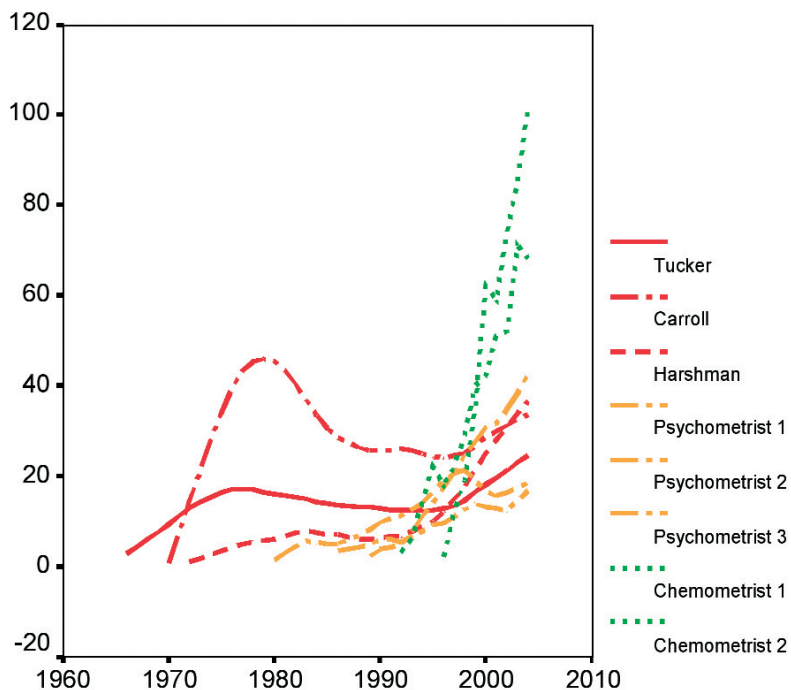


PLAAT III

Ontwikkeling van de drieweganalyse 1963 – 2004

Populariteit van de drieweganalyse

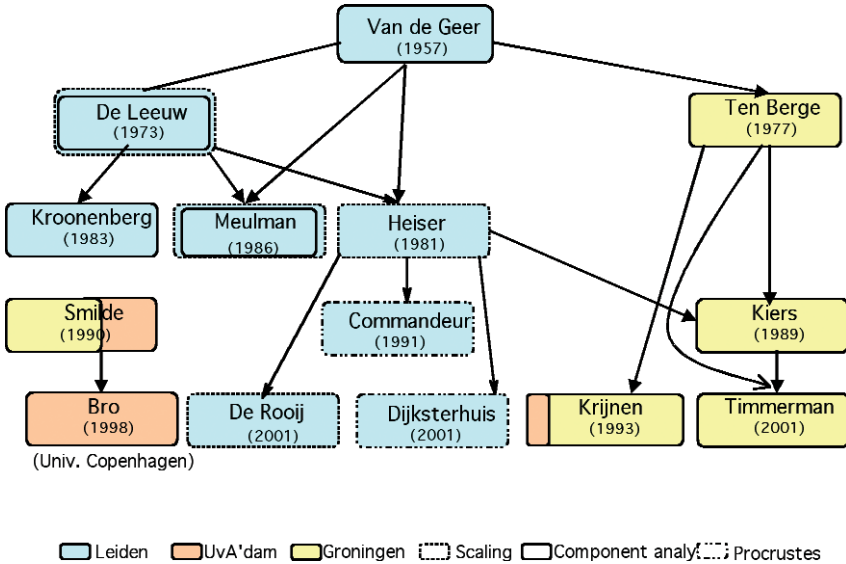
(citaties per jaar)



Noot: De gegevens voor de figuur zijn verzameld met behulp van het *Web of Science*.

PLAAT IV

Nederlandse driewegpromotiegenealogie



Noot:

Jaartallen refereren aan het jaar van promotie; alleen namen van onderzoekers die primair actief zijn op drieweggebied en wier proefschrift ook drieweg-georiënteerde onderdelen bevatten, zijn opgenomen.

Andere Nederlanders met specifieke bijdragen op drieweggebied:

Wansbeek (econometrie; meerweg structurele modellen); Verhees (promotor Wansbeek); Neudecker, Magnus (wiskunde); Niesing (promotores. Kiers, Ten Berge), Bijlsma (promotor. Smilde), Faber (analytische chemie; statistiek); Westerhuis (proces- & analytische chemie); Van Houwelingen, Weesie (mathematische statistiek); Groenen, Nierop (promotor: Heiser); Oort (meerweg structurele modellen).

schalen en personen. Ik had het idee dat hij zijn gegevens op een andere manier zou moeten analyseren dan hij deed. Hij telde op over personen en gooide de individuele verschillen weg. [hij plette dus zijn datadoos] Daarom ontwikkelde ik driewegfactoranalyse, een erg algemeen model om individuele verschillen te kunnen onderzoeken." (p. 8)

Tucker illustreerde zijn bezwaar tegen het platmaken met het voorbeeld van de firma Chrysler, die een auto ontwierp voor de gemiddelde koper zonder acht te slaan op het belang van individuele verschillen. Helaas werd de auto niet verkocht omdat die *gemiddelde* koper niet bestond, en Chrysler ging hier bijna aan failliet. De noodzakelijke aandacht voor individuele verschillen is nog steeds de leidraad van vrijwel al het werk op het gebied van drieweganalyse.

Hoe ver zijn we nu met drieweganalyse?

Een volledige schets van de ontwikkeling van drieweganalyse zou niet passen in het kader van deze rede, te meer daar ik dan verder in de techniek zou moeten duiken dan vandaag mijn bedoeling is. Om toch enige indruk te geven van het wel en wee over de jaren heen, heb ik voor U een figuur gemaakt (Plaat III) met de enigszins grof berekende citatiecurven van de voornaamste protagonisten, ingedeeld in grondleggers (Tucker, Carroll en Harshman), psychometrici, en chemometrici, waarbij de laatste twee groepen bijna geheel uit Nederlanders bestaan. Deze voorstelling van zaken is niet ingegeven door chauvinisme, maar het is een feit dat in de jaren 80 en 90 diverse Nederlanders voortrekkers op dit terrein zijn geweest.

Het meest opvallende aan deze grafiek is de bijzonder sterke stijging in citaties in de jaren negentig, en het feit dat deze stijging bijna geheel voor rekening komt van de chemometrici. Het is in feite zelfs zo dat de toename aan citaties van de psychometricisten ook te danken is aan de stormachtige ontwikkelingen in de chemometrie. Toen men zich daar eenmaal de kracht van drieweganalyse, met name de modellen ontwikkeld door Richard Harshman, gerealiseerd had, was er geen houden meer aan. Eén van de voornaamste oorzaken was dat men in de (analytische) chemie modellen heeft die van zichzelf al de vorm van een driewegmodel hadden. Men hoefde daar alleen maar de parameters te schatten, in plaats van uit te zoeken of het model wel van toepassing was en in welke vorm dat zou zijn. Bovendien waren er ook direct al belangrijke commerciële toepassingen.¹⁹

Een interessant aspect van deze ontwikkeling is dat de psychometrici ook zijn gaan publiceren in chemometrische tijdschriften omdat in die hoek hun grootste klanten zaten.²⁰ Betekent dit echter dat drieweganalyse vanaf nu uitsluitend een chemische zaak is? Zeker niet, toepassingen in andere wetenschappen, zoals landbouw, signaalverwerking, medicijnen en de wiskunde verschijnen ook regelmatig, meestal met mondjesmaat, maar ook daar is er een groeiende belangstelling.²¹

Tellen de sociale en gedragswetenschappen dan niet meer mee? Is er nog wel emploti voor driewegmethoden in deze disciplines? Kunnen we niet al het werk over

laten aan anderen en hun technieken gewoon gebruiken? Is het eigenlijk wel nodig om een leerstoel in de sociale wetenschappen te hebben voor multivariate analyse in het bijzonder van driewegdata? Zoals U kunt verwachten is mijn antwoord een nadrukkelijk "Ja!".

Met de instelling van deze leerstoel zal het juist mogelijk zijn om de eigenheid van sociaal-wetenschappelijke gegevens aan te pakken. Zoals met name mijn naaste data-theoretische collega's niet na laten te benadrukken, worden veel sociaal-wetenschappelijke data gekenmerkt doordat ze categorisch zijn en niet continu, zoals in de meeste natuurwetenschappelijke disciplines. De Leidse Datatheoriegroep heeft belangrijke bijdragen geleverd in het analyseerbaar maken van grote verzamelingen categorische data en het ligt dan ook op mijn weg om deze lijn voort te zetten in de drieweganalyse.

Een andere reeds lang gevestigde trend is de opzet van grootschalige longitudinale onderzoeken naar de ontwikkeling van kinderen. Dit gebeurt niet alleen overzee, maar ook hier in Leiden, en met name in mijn eigen afdeling Algemene en Gezinspedagogiek en Datatheorie. Ook hier krijgen de onderzoekers te maken met complexe bestanden van multivariate gegevens, vaak met verschillende meetniveaus en met grote verschillen tussen (groepen van) individuen. Drieweganalyse en longitudinale data-analyse met een serieuze rol voor individuele verschillen zijn voor elkaar geschapen. Ook hier kan ik leunen op werk gedaan in Nederland waarbij ik speciaal mijn Groningse collega's wil noemen.²²

De conclusie uit het voorgaande mag zijn dat een indringende aandacht voor drieweganalyse, speciaal gericht op gegevens die kenmerkend zijn voor de sociale en gedragswetenschappen, van groot belang zal zijn voor de inhoudelijke ontwikkeling van die wetenschappen en voor de data-analyse zelf.

Waarom heeft Nederland zo'n vooraanstaande plaats in de drieweganalyse?

Professor Van de Geer, de oprichter van de Stichting die deze leerstoel heeft ingesteld, was dé pionier op het gebied van de data-analyse in Nederland en met name in Leiden. Zijn dochter, Prof. Sara van de Geer verwoordde dit vijf jaar geleden in haar oratie als: "De datatheorie in Leiden is jouw geesteskind"²³.

Daarnaast is hij *ook* de grondlegger geweest van de *drieweganalyse* in Nederland. Dit klinkt misschien een beetje als 'wereldberoemd op het Rapenburg', maar zo is het niet bedoeld. Van de Geer heeft didactische stukken over drieweganalyse geschreven die helaas nooit zijn gepubliceerd, maar die mij als student wel hielpen te begrijpen waar het nu allemaal over ging.²⁴

Ter illustratie van de invloed die hij heeft gehad op de drieweganalyse in Nederland heb ik een promotiegenealogie geconstrueerd (Plaat IV), waaruit zijn nestoraat duidelijk naar voren komt. Ik zal dit schema niet in detail bespreken, maar slechts memoreren dat hij aan de basis heeft gestaan van zowel de Leidse als de Groningse school. U ziet overigens aan de dwarsverbanden dat er eigenlijk zoiets bestaat als een Nederlandse school van drieweganalyse. Ik had ook een dergelijke figuur kunnen maken van personen die met elkaar gepubliceerd hebben, maar dat zou

een onbeschrijflijke wirwar van lijnen te zien geven. Mocht dat op U enigszins incetueus overkomen, dan wil ik dat ten stelligste bestrijden. Het is juist deze intensieve samenwerking die de huidige ontwikkelingen heeft gestimuleerd. Met name het kader van de Interuniversitaire Onderzoekschool voor de Psychometrie en Sociometrie (IOPS) is hier debet aan, en het toont de kracht van zulke nationale onderzoekscholen in plaats van de nu steeds meer in zwang komende lokale varianten.

Toen ik als student toegepaste wiskunde besloot dat ik toch ook wel eens wiskunde wilde maken in plaats van alleen reproduceren, wendde ik mij tot één van mijn docenten wiskunde met de vraag of die een onderwerp voor me wist om een scriptie over te schrijven. Als ik hem destijds goed begrepen heb, was hij enigszins verbaasd over die vraag en antwoordde, wij schrijven 1975, dat ik toch niet kon verwachten dat ze voor iedere student zo maar een onderwerp hadden. Dus wendde ik me tot Jan de Leeuw, mijn baas als student-assistent, die mij een aantal blaadjes oranje grafiekpapier gaf met op elk een paar formules die betrekking hadden op een "ping-pong"-algoritme voor drieweganalyse.²⁵ En zoals U ziet, hier sta ik nu en ik heb het onderwerp nooit vaarwel kunnen zeggen.

Dankwoord

Aan het derde en laatste deel van mijn rede gekomen, wil ik een dankwoord richten tot degenen die hebben bijgedragen aan het feit dat ik hier nu sta.

Zij die hebben bijgedragen aan de instelling van de leerstoel

Allereerst natuurlijk dank aan allen die hebben bijgedragen tot de instelling van deze leerstoel, in het bijzonder het bestuur van de Stichting ter Bevordering van de Wetenschap der Datatheorie onder de bezielende leiding van Professor Willem Heiser, en uiteraard het College van Bestuur van deze universiteit dat er in toegestemd heeft deze leerstoel hier te vestigen. Het Curatorium van de leerstoel, in het bijzonder de hoogleraren Heiser, Meulman, Smilde en Ten Berge, zal toezien op mijn vervulling van deze leeropdracht, en ik zie hen niet zo zeer als een verzameling schoolmeesters die me bij de les zullen houden, maar als een verzameling collega's die mij zullen stimuleren in mijn pogingen de drieweganalyse verder uit te breiden en tot nut te laten zijn in inhoudelijke disciplines, met name de sociale en gedragswetenschappen.

Zij die hebben bijgedragen aan mijn datatheoretische vorming

Eén van de belangrijkste personen in mijn opleiding tot statisticus is de veel te vroeg overleden Albert Verbeek geweest: vriend, leermeester en collega. Van hem heb ik geleerd waar het echt om gaat in de statistiek en ik probeer zijn inzichten nog iedere dag over te dragen aan mijn studenten.

Zoals hierboven al gememoreerd, heb ik mijn vorming op het gebied van de multivariate analyse, in het bijzonder van driewegdata, vooral te danken aan Jan de Leeuw in persoon en Professor Van de Geer in geschrift.

Hooggeleerde Meulman, Beste Jacqueline, ik heb tot nog toe op bescheiden schaal met je samen gewerkt op onderzoekstechnisch terrein, maar des te meer in onderwijs aan promovendi.²⁶ Ik hoop dat we in de toekomst met name op het gebied van optimale schaling en drieweganalyse een goede symbiose zullen vinden die zal leiden tot een aantal mooie technieken, toepassingen en dito publicaties. De enige poging in die richting was van de hand van Richard Sands, student aan de University of North Carolina²⁷. Hij is nu Chairman en Chief Executive Officer van Constellation Brands, 's werelds grootste wijnproducent, ook niet onaardig, maar dat haalt het natuurlijk niet bij de verlokkingen van de drieweganalyse.

Zij die hebben bijgedragen aan mijn sociaal-wetenschappelijke vorming

Bij mijn aanstelling bij Pedagogiek in 1975 werd ik verkozen boven een andere kandidaat, naar ik later vernam omdat men dacht dat het makkelijker voor een statisticus was om sociale wetenschappen te leren dan andersom. Of ik deze verwachting nu heb vervuld is aan anderen om te oordelen, maar, wellicht tot mijn schande, moet ik zeggen dat, na zo'n 30 jaar, ik me nog steeds meer statisticus voel dan pedagoog. Ik ben een vreemde eend in de bijt, maar wel één die zich nu best lekker voelt in pedagogisch water.

Hooggeleerde Van IJzendoorn, Beste Rien, één van de dingen die ik het meest heb gewaardeerd in je houding tegenover mij is de stimulans en de vrijheid die je me hebt gegeven om mijn specialisme uit te oefenen. Je stond, en staat nog steeds, coulant tegenover mijn analyses van katoenoogsten,²⁸ pindaplantjes,²⁹ cocaïnemoleculen,³⁰ en zieke blauwe krabben.³¹ Je moet soms het sombere idee hebben gehad dat ik alles analyseerde wat los of vast zat, zolang het maar in een doosje paste en géén sociale wetenschappen was. Gelukkig wist je me regelmatig met mooie pedagogische databestanden op het rechte pad te houden.³² Het is grotendeels aan jou te danken, dat ik geen grote neiging heb gehad om mijn driewegkunstjes permanent elders te vertonen. Het lijkt me hier een typisch voorbeeld van veilige gehechtheid en "secure-base behaviour". Vele korte en langer durende exploraties in het buitenland brachten me vroeger of later weer terug in de moederschoot, bij mijn alma mater.

Studenten: Wat U van mij kunt verwachten

Dames en heren studenten pedagogische wetenschappen. Van mijn specialisme hebt U, in uw dagelijkse praktijk van student-zijn, niet erg veel gemerkt en ik ben bang dat U dat ook in de toekomst niet zult doen. Gedeeltelijk komt dit, omdat voor de gegevens waar U mee werkt zelden dozen nodig zijn. Waar ik U meer mee van dienst kan zijn, is mijn ervaring met het hanteren van allerlei soorten data, drieweg of niet. De verscheidenheid aan gegevens die in de loop der jaren door mijn handen zijn gegaan, zal, naar ik verwacht, nuttig voor U zijn wanneer U in de fase van uw masterproject bent aangeland en U zich, al dan niet vertwijfeld, afvraagt wat U er mee doen moet.³³

Professor Van de Geer meldde in zijn oratie in 1964 dat "bij toespraken als deze gemiddeld 2.01% der woorden rechtstreeks tot u [studenten] wordt gericht, een per-

centage dat door zijn schrielheid in fel contrast staat met de dagelijkse gang van zaken.¹⁹³⁴ Helaas moet ik bekennen dat ik in deze rede, geheel onbewust, ook niet veel verder gekomen ben dan dit magere percentage.

Familie: De link met het verleden die de toekomst waarborgt

Ik vind het ontzettend jammer dat mijn beide ouders deze plechtigheid niet meer mee hebben kunnen maken. Ook betreur ik het zeer dat mijn moeders broer, Lipke Holthuis, wegens gezondheidsredenen deze plechtigheid niet kan bijwonen. Hij heeft in hoge mate bijgedragen aan mijn wetenschappelijke carrière, niet alleen als wetenschapper maar ook als suikeroom in magere studentenjaren. Daarnaast heeft hij door zijn voorbeeld mij een onbedwingbare lust gegeven de wereld te verkennen. Van de generatie van mijn ouders zijn gelukkig wel mijn vaders broer Kees en zijn vrouw Lydia aanwezig. Ook de jongste generaties zijn ruim vertegenwoordigd en wellicht zullen zij in de toekomst nog wel eens data in dozen stoppen. En wie weet, worden voor hen vier- en meerdimensionale datadozen wel heel gewoon. Deze hyperdozen bestaan allang en worden ook al uitgebreid onderzocht, maar daarover mag ik U wellicht een andere keer onderhouden.

Ik heb gezegd

Noten

- 1 De illustraties uit het uitgereikte boekje zijn alle opgenomen in deze gedrukte uitgave van de rede.
- 2 Dit voorbeeld is geënt op diverse driewegtoepassingen in de landbouw, zie bijvoorbeeld:
Basford, K.E., Kroonenberg, P.M., & DeLacy, I.H. (1991). Three-way methods for multiattribute genotype by environment data: An illustrated partial survey. *Field Crops Research*, 27, 131-157.
- 3 Een inleidend overzichtsartikel met een gedegen uitleg van het Parafac model met toepassingen in de chemie is te vinden in:
Bro, R. (1997). PARAFAC. Tutorial & applications. *Chemometrics and Intelligent Laboratory Systems*, 38, 149-171.
Smilde, A.K., Bro, R., & Geladi, P. (2004). *Multi-way analysis. Applications in the chemical sciences*. Chichester, UK: Wiley.
- 4 Dit voorbeeld is slechts denkbeeldig, maar is geënt op werk gedaan in het kader van het onderzoeksproject van de Leidse Faculteit der Godgeleerdheid onder leiding van Prof. E.G.E. van der Wall, "Christendom en moderniteit".
- 5 Kroonenberg, P.M., Dunn III, W.J., & Commandeur, J.J.F. (2003). Consensus molecular alignment based on generalized Procrustes analysis. *Journal of Chemical Information and Computer Science*, 45, 69-97.
- 6 Sidiropoulos, N.D., Bro, R., & Giannakis, G.B. (2000). Parallel factor analysis in sensor array processing. *IEEE Transactions on Signal Processing*, 48, 2377-2388.
- 7 Hoewel Jan Blokker en het "kruidenvrouwtje van Stroe" in mijn geheugen onlosmakelijk met elkaar verbonden zijn, heb ik geen bewijs kunnen vinden dat deze twee inderdaad aan elkaar gekoppeld zijn. In een persoonlijke brief van Jan Blokker meldt hij "Het lijkt me niet uitgesloten dat ik zoiets ooit heb geschreven, maar ik heb geen flauw idee wanneer, en in welke ruimere context".
- 8 Köbben, A.J.F. (1980). *Het heilig vuur*. Oratie ter aanvaarding van de Cleveringa-leerstoel aan de Rijksuniversiteit Leiden (Een bewerkte versie onder dezelfde titel is verschenen in *Intermediair*, 16 (49), 1-7). Waarschijnlijk ligt de oorsprong van deze confrontatie bij P.F. Lazarsfeld (1949). "The American Soldier". An expository review. *Public Opinion Quarterly*, 13, 377-404 (m.n. p. 380). Dit artikel wordt aangehaald in J. Goudsblom (1974). *Balans van de sociologie*. Utrecht: Het Spectrum (Aula pocket Nr. 518). Met dank aan Prof. Leo van der Kamp voor deze referenties.
- 9 Kiers, H.A.L. (2004). Bootstrap confidence intervals for three-way methods. *Journal of Chemometrics*, 18, 22-36.
- 10 Murakami, T., & Kroonenberg, P.M. (2003). Individual differences in semantic differential studies and their analysis by three-mode models. *Multivariate Behavioral Research*, 38, 87-96.

- 11 In eigen bewerkingen voor orgel werden tijdens de oratie de preludes 22, 20 en 7 van Chopin ten gehore gebracht door Drs. C.C. de Jonge. Van elke prelude werd maximaal 1 minuut gespeeld.
- 12 Bron (gezien: 28/8/2005): <http://www.geocities.com/darnbi/english/prel.htm>. Over Prelude Nr.4 : "It was fittingly played, with Nos. 6 and 20, by the famous organist Lefébure-Wely, at Chopin's funeral service at the Madeleine Church in Paris".
- 13 De waarschijnlijke oorzaak is dat in sommige langere preludes modulaties voorkomen en studenten slechts de eerste 80 seconden te horen kregen, zodat ze hierdoor slechts een beperkt beeld konden krijgen van de aard van deze preludes.
- 14 Bewerkte versie van een oorspronkelijke afbeelding te vinden op <http://www.uncletim.com/answers/circle.htm> (laatst gezien: 19/7/2005); volgens de website ontleend aan: Tim Gillespie, *Uncle Tim's Building Blocks*, Fifth printing 2004, ISBN 0-9647059-7-4.
- 15 Osgood, C.E., Suci, G.J., & Tannebaum, P. (1967). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- 16 Triandis, H.C. (Ed.) (1972). *The analysis of subjective culture*. New York: Wiley.
- 17 Tucker, L.R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31, 279-311.
- 18 Dorans, N. (2004). *A conversation with Ledyard R Tucker*. Princeton: Educational Testing Service (www.ets.org/research/dload/tucker.pdf; gezien 20/8/2005; zie ook het In Memoriam door Dorans in *Journal of Educational and Behavioral Statistics*, 29, 377-378).
- 19 Zie het boek van Smilde, Bro en Geladi, noot 3.
- 20 Vooral Kiers, Ten Berge en Harshaman, en recent ook Kroonenberg, hebben diverse publicaties het licht doen zien in het *Journal of Chemometrics en Chemometrics and Intelligent Laboratory Systems*.
- 21 Er zijn nu al twee specialistische kleinschalige conferenties die zich op het terrein van de drie- en meerweganalyse richten: TRilinear methods in Chemistry And Psychology (TRICAP) sinds 1993 en de Tensor workshops sinds 2004.
- 22 De Groningse onderzoekers Ten Berge en Kiers kregen tezamen met Leidse onderzoekers Heiser en Kroonenberg, een vijfjarige subsidie voor het project: "Meerweganalyse van multivariate longitudinale gegevens" (1996-2000). Dit heeft onder andere proefschriften opgeleverd van Timmerman en De Rooij, alsmede een serie artikelen over het onderwerp.
- 23 Van de Geer, S.A. (2000). *Een zeker toeval*. Rede uitgesproken ter gelegenheid van de aanvaarding van het ambt van hoogleraar in de Kansrekening en Statistiek aan de Universiteit Leiden (1 december 2000).
- 24 Van de Geer, J.P. (1974). Toepassing van drieweg-analyse voor de analyse van multiple tijdreeksen. In C.J. Lammers (Ed.), *Groei en ontwikkeling van de ziekenhuisorganisaties in Nederland*. Interim rapport, Sociologisch Instituut, Universiteit Leiden.

- 25 Kroonenberg, P.M., & De Leeuw, J. (1977). *TUCKALS2: A principal component analysis of three-mode data*. Research Bulletin No. RB 001-'77, Department of Data Theory, Leiden University, Leiden The Netherlands. (Licht bewerkte versie van mijn doctoraalscriptie).
- 26 Ik denk hierbij met name aan de cursus "Toegepaste multivariate data-analyse" die Willem Heiser, Jacqueline Meulman en ik diverse malen hebben gegeven aan promovendi, tot groot genoegen van deelnemers en docenten.
- 27 Sands, R., & Young, F.W. (1980). Component models for three-way data: ALS-COMP3, an alternating least squares algorithm with optimal scaling features. *Psychometrika*, 45, 39-67.
- 28 Basford, K.E., Kroonenberg, P.M., DeLacy, I.H., & Lawrence, P.K. (1990). Multiattribute evaluation of regional cotton variety trials. *Theoretical and Applied Genetics*, 79, 225-234.
- 29 Kroonenberg, P.M., Harch, B.D., Basford, K.E., & Cruickshank, A. (1997). Combined analysis of categorical and numerical descriptors of Australian groundnut accessions using nonlinear principal component analysis. *Journal of Agricultural, Biological, and Environmental Sciences*, 2, 294-312.
- 30 Zie noot 5.
- 31 Kroonenberg, P.M., Basford, K.E., & Gemperline, P.J. (2004). Grouping three-mode data with mixture methods: The case of the diseased blue crabs. *Journal of Chemometrics*, 18, 508-518.
- 32 Zonder twijfel is onze meest geslaagde onderneming het 1988 artikel in *Child Development* (59, 147-156) getiteld "Cross-cultural patterns of attachment: A meta-analysis of the Strange Situation". De meest recente loot aan onze langdurige en productieve samenwerking is een heranalyse van gegevens uit de NICHD Study of Early Child Care and Youth Development over de relatie tussen kinderopvang en agressie.
- 33 Sinds anderhalf jaar heeft het Departement Pedagogische Wetenschappen een Scriptie-atelier ingericht, waarbij studenten in hun doctoraal- of masterfase aan hun scriptie of masterproject kunnen werken en advies kunnen krijgen van ervaren scriptie-ateliermedewerkers over het schrijven van een scriptie en alles wat daarmee verband houdt, en over het verwerken en analyseren van hun gegevens.
- 34 Van de Geer, J.P. (1964). *Kans en Onzekerheid*. Rede uitgesproken bij de aanvaarding van het ambt van gewoon hoogleraar in de Experimentele Psychologie en de Statistische Methode in de Psychologie aan de Rijksuniversiteit te Leiden (1 mei 1964).

Genealogie van Nederlandse “drieweg”-proefschriften¹

Gedragwetenschappen

- Commandeur, J.J.F. (1991). *Matching configurations*. (Proefschrift Rijksuniversiteit Leiden, Promotor: J.M.F. ten Berge, Ref.: Van de Geer, Cie.: Heiser, Kroonenberg).
- De Leeuw, J. (1973). *Canonical analysis of categorical data*. Leiden: DSWO Press (1984). (Proefschrift Rijksuniversiteit Leiden, Promotor: J.P. van de Geer).*
- De Rooij, M.J. (2001). Distance models for transition frequency data (Proefschrift Universiteit Leiden, Promotor: W.J. Heiser, Cie.: Kiers, Meulman).
- Dijksterhuis, G.B. (1995). *Multivariate data analysis in sensory and consumer science*. (Proefschrift Rijksuniversiteit Leiden, Promotor: W.J. Heiser).
- Heiser, W.J. (1981). *Unfolding analysis of proximity data*. (Proefschrift Rijksuniversiteit Leiden, Promotors: J.P. van de Geer, J. de Leeuw).*
- Kiers, H.A.L. (1989). *Three-way methods for the analysis of qualitative and quantitative two-way data*. Leiden: DSWO Press. (Proefschrift Rijksuniversiteit Groningen, Promotors: J.M.F. ten Berge, W.J. Heiser).
- Krijnen, W.P. (1993). *The analysis of three-way arrays by constrained Parafac methods*. Leiden: DSWO Press. (Proefschrift Rijksuniversiteit Groningen, Promotor: J.M.F. ten Berge, Ref.: Kiers, Cie.: Heiser).
- Kroonenberg, P.M. (1983). *Three-mode principal component analysis: Theory and applications*. Leiden: DWO Press. (Proefschrift Rijksuniversiteit Leiden, Promotor: J. de Leeuw, Cie.: Van de Geer).
- Nierop, A.F.M. (1993). *Multidimensional analysis of grouped variables: An integrated approach*. Leiden: DSWO Press. (Proefschrift Rijksuniversiteit Leiden, Promotor: W.J. Heiser, Ref.: Van de Geer, Cie.: Ten Berge).
- Niesing, J. (1997). *Simultaneous component and factor analysis methods for two or more groups: A comparative study*. Leiden: DSWO Press. (Proefschrift Rijksuniversiteit Groningen, Promotors: H.A.L. Kiers; J.M.F. ten Berge, Cie.: Heiser).
- Meulman, J.J. (1986). *A distance approach to nonlinear multivariate analysis*. (Proefschrift Rijksuniversiteit Leiden, Promotors: J.P. van de Geer, J. de Leeuw).*
- Ten Berge, J.M.F. (1977). *Optimizing factorial invariance*. (Proefschrift Rijksuniversiteit Leiden, Promotor: J.P. van de Geer).
- Timmerman, M.E. (2001). *Component analysis of multisubject multivariate longitudinal data*. (Proefschrift Rijksuniversiteit Groningen, Promotors: H.A.L. Kiers; J.M.F. ten Berge. Cie.: Heiser, Smilde).

1 Proefschriften gemarkeerd met een * gaan niet, slechts zijdelings, of in geringe mate over driewegonderwerpen, maar vormen een essentieel onderdeel van de genealogie. De verbanden tussen de onderzoekers zijn nog sterker dan door de genealogiefiguur wordt gesuggereerd, omdat andere driewegonderzoekers onderdeel uitmaakten van de promotiecommissie. Ook zij worden in deze lijst vermeld.

Scheikunde

- Bijlsma, S. (2000). *Estimating rate constants of chemical reactions using spectroscopy*. (Proefschrift Universiteit van Amsterdam, Promotor: A.K. Smilde, Cie.: Kiers).
- Bro, R. (1998). *Multi-way analysis in the food industry*. (Proefschrift Universiteit van Amsterdam, Promotor: A.K. Smilde, Cie.: Kiers, Kroonenberg)².
- Smilde, A.K. (1990). *Multivariate calibration of reversed-phase chromatographic systems*. (Proefschrift Rijksuniversiteit Groningen, Promotor: D.A. Doornbos. Cie.: Wansbeek)³
- Nilsson, J.F (1998). *Multiway calibration in 3D QSAR. Applications to dopamine receptor ligands*. (Proefschrift Rijksuniversiteit Groningen, Promotor: A.K. Smilde. Cie.: Ten Berge).

Economie

- Verhees, J. (1989). *Econometric analysis of multidimensional models*. (Proefschrift Rijksuniversiteit Groningen, Faculteit Economische Wetenschappen, Promotor: T.J. Wansbeek; Cie.: Ten Berge).

-
- 2 Rasmus Bro is de enige niet-Nederlander in de genealogie in Plaat VI. Hij was en is verbonden aan KVL - Den Kgl. Veterinær- og Landbohøjskole, Universiteit van Kopenhagen; sinds 2002 als hoogleraar in de chemometrie.
 - 3 Smildes proefschrift past niet in de genealogie en zijn proefschrift bevat nog geen verwijzingen naar het werk van zijn Groningse collega's, maar spoedig na het verschijnen van het proefschrift in 1990, is een intensieve samenwerking ontstaan tussen Smilde, Kiers en Ten Berge. De precieze begindatum hiervan kon ik niet achterhalen. Deze samenwerking is van fundamenteel belang geweest voor de kruisbestuiving van driewegonderzoek in de chemie en de psychologie en heeft tot een groot aantal gezamenlijke artikelen geleid. Deze samenwerking is ook de start geweest van een reeks conferenties op drieweggebied. De eerste was in 1993 in Epe, Gelderland – ThRee-way Methods in Chemistry (TRIC). Latere edities hebben plaats gevonden onder de naam TRICAP in Manson, WA (1996), Faaborg, Denmark (2000), Lexington, KT (2003) en, in 2006 op Kreta.

