

HET TALENTENPROJEKT NOG EENS  
OPNIEUW BESCHOUWD

Adriaan Meester  
Jan de Leeuw  
Vakgroep Datatheorie FSW/RUL

Maart 1983

## 1. Inleiding

Sinds jaar en dag is de rol die het onderwijs speelt in het bestendigen of het mogelijkkerwijs opheffen van maatschappelijke tegenstellingen voorwerp van discussie. Matthijssen (1975) probeert deze discussie enigszins te ordenen en rangschikt de verschillende definities (hij onderscheidt er vijf) die in de loop der jaren van de sociale ongelijkheid van onderwijskansen zijn gegeven, langs een continuüm dat volgens hem in toenemende mate de bewustwording over de aard van het ongelijkheidsverschijnsel weergeeft.

"De eerste definitie relateert nauwelijks iets van de vooronderstellingen die ten grondslag liggen aan ons onderwijs-maatschappij-systeem. Elke volgende definitie gaat in dit relativiseringsproces een stap verder, dat wil zeggen dat de achtereenvolgende definities toenemend 'fundamenteel' worden in het ter discussie stellen van de organisatie-principes van ons onderwijs in onze maatschappij" (p. 449). Sociale ongelijkheid van onderwijskansen beschouwt hij overigens "niet primair als een onderwijsprobleem, maar veeleer een algemeen maatschappelijk probleem dat izch óók in het onderwijs uitdrukt" (p. 429).

De eerste definitie is gericht op individuele mobiliteit, beschouwt de bestaande maatschappelijke orde als gegeven en het gelijkheidsprobleem dientengevolge als een toelatingsprobleem: alle kinderen van een bepaald intellectueel begaafdheidsniveau moeten dezelfde mogelijkheden hebben om toegelaten te worden tot een bepaald niveau van onderwijs. Dat impliceert het opsporen en activeren van hoger begaafden in de lagere sociale milieu's.

Deze preoccupatie met hoger begaafden en het reduceren van begaafdheid tot intellectualiteit is volgens Matthijssen kenmerkend voor de eerste definitie en de onderzoeken in het kader van het Talentenproject, evenals die van Peschar (1975) en van Sonnemans en hemzelf (Matthijssen en Sonnemans, 1958) passen allemaal in deze traditie. Van Heek formuleert het trouwens in zijn inleiding tot "Het Verborgene Talent" op soortgelijke wijze. Hij onderscheidt drie criteria voor het bezetten van individuele sociale posities in de samenleving: erfelijk privilege, gelijkheid en prestatie en hoewel elk van deze beginselen in zijn uiterste toepassing kwetsbaar is, neemt dit niet weg, "dat in onze samenleving aan het prestatiebeginsel terecht een overwegende betekenis wordt toegekend" (Van Heek, 1968, p. 4). In het

licht van Matthijssen's continuüm neemt van Heek in feite een ook voor dat moment (1968) nogal "ouderwets" standpunt in.

Al in 1959 benadrukte Idenburg in zijn kritiek op Michael Young's "The Rise of the Meritocracy 1870-2033" de eenzijdigheid van deze traditionele benadering van gelijkheid van kansen: "Deze visie veronderstelt een maatschappij, welke niet meer is dan het operatieterrein voor individuele opklimmingsdrift en begeerte naar invloed en macht" (Idenburg, 1959; p. 222).

Begaafdheid moet in Idenburg's opvatting meer multidimensioneel gedefinieerd worden en de ontwikkeling ervan is een maatschappelijk belang, méér dan een individueel belang (vgl. Idenburg, 1959, 1962). Nu is het uiteraard niet zo, dat van Heek zich van deze en andere opvattingen niet bewust is. Hij geeft er in zijn tweede hoofdstuk een uitgebreid overzicht<sup>1)</sup> van. Als het echter op de formulering van de probleemstelling en de operationalisatie van het begrip "schoolgeschiktheid" aankomt is hij weer helemaal bij de in de eerste zin geformuleerde drie beginselen terug. Erfelijk privilege, gelijkheid en prestatie en de optimale vermenging van die drie ingrediënten, dat staat centraal. We gaan daar in de volgende paragraaf nader op in.

## 2. Talentenreserve en schoolgeschiktheid

Allereerst van Heek's probleemstelling: "In hoeverre zijn de beschikbare 'reserves' aan schoolgeschikten voor v.h.m.o. in ons land inderdaad uitgeput wat betreft parate schoolgeschiktheid (schoolgeschiktheid bij het huidige school- en voorlichtingssysteem)?" (p. 55). Deze probleemstelling plaatst het Talentenproject ondubbelzinnig onderaan op Matthijssen's "continuüm". Het gaat om de hoger begaafden en hun individuele mobiliteit, gegeven het bestaande maatschappelijke systeem.

In de probleemstelling komt het woord "schoolgeschikt" drie keer voor en dat vraagt enige nadere aandacht. We nemen een voorschotje op de beschrijving van de variabelen in de volgende paragraaf en behandelen schoolgeschiktheid nu alvast, omdat dit begrip in het Talentenproject centraal staat en in de discussie naderhand het meeste stof heeft doen opwaaien.

Eerst definieert van Heek het begrip "schoolstudietalent" als: "de re-

delijke kans een bepaalde fase van het onderwijs (b.v.: l.o., v.h.m.o. of h.o.) volledig te doorlopen en het desbetreffende diploma te verkrijgen" (p. 31). Dat berust, zegt hij, op een zekere mate van schoolgeschiktheid. Om dat te meten is het nodig een testbatterij te gebruiken die over een voorspellende waarde met betrekking tot een bepaald concreet schooltype beschikt. Op die manier wordt kans op schoolsucces vastgesteld en hij haast zich om toe te voegen, dat dat natuurlijk nog wat anders is dan talent in algemene zin.

Het begrip "talent" komt dan verder niet meer terug. Na schoolgeschiktheid op deze manier te hebben ingeperkt heeft van Heek een korte beschrijving van de totstandkoming van de geschiktheids-score: die is "vastgesteld met behulp van een testbatterij (waarin zowel intelligentie- als schoolvorderingentests zijn opgenomen), de schoolcijfers en het oordeel van de onderwijzer over de leerling" (p. 31). De testbatterij is de Nederlandse Onderwijs-Differentiatie Testserie (NDT) die op dat moment geijkt was voor de schoolresultaten in de 1e klas van het v.h.m.o.. Er is dus geen criterium voor het behalen van het einddiploma v.h.m.o., maar het zonder doubleren doorlopen van de 1e klas fungeert als plaatsvervanger. Zonder nadere precisering wordt opgemerkt, dat er "een vrij sterke samenhang" tussen de resultaten in de eerste klas en het behalen van het einddiploma bestaat.

Op basis van de bovengenoemde combinatie van variabelen wordt een predictiescore berekend die de kans van slagen in het v.h.m.o. zou weergeven en in 4 klassen wordt ingedeeld: A = meer dan 75% kans van slagen bij het v.h.m.o., B = 45-75%, C = 15-45% en D = minder dan 15% kans van slagen.

Merkwaardig is nu dat van Heek bij zijn beschrijving van de samenstelling van de predictiescore het beroep van de vader niet noemt. Dat komt pas in de rapportage van van Weeren ter sprake die een wat gedetailleerder beschrijving van de elementen, waaruit de voorspellingsscore is opgebouwd, geeft. Bijna terloops wordt het beroep van de vader aan het eind van de rij van samenstellende variabelen genoemd (p. 127). Dat is ook de enige keer in het hele boek dat blijkt, dat het social milieu deel uitmaakt van de kansscore.

Deze wat "stiekeme" introductie van het ouderlijk milieu als onderdeel van "talent" is echter niet aan de aandacht van de oplettende critici van het "Talentenproject" ontsnapt.

Tot die oplettende lezers behoorde van Kemenade die in 1969 in de Sociologische Gids de discussie opent. Van Kemenade's kritiek richt zich vooral op twee punten: de algemene conclusie uit het project dat bij de overgang l.o.-v.h.m.o. geen talent verloren gaat en, ten tweede, dat het beroep van de vader deel uitmaakt van predictiescore. Op de algemene conclusie was ook al door van Calcar, Soutendijk en Tellegen (1968; ook opgenomen in "Het verborgen talent") kritiek geïoefend. De resultaten van de eerste fase van het Talentenproject steunen slechts schijnbaar de opvatting dat de geringe deelname van arbeiderskinderen aan hogere vormen van onderwijs het gevolg kan zijn van hun geringere begaafdheid, stellen zij, omdat het project uitgang van het bestaande sterk selectieve onderwijssysteem in Nederland. In vergelijking met het buitenland komen hier maar heel weinig kinderen in het hoger onderwijs terecht. Dergelijke hoge drempels werken het meest in het nadeel van de lagere sociale klassen. Vergelijking met het buitenland leert ook dat bij ruimere toelating tot het voortgezet onderwijs de gemiddelde prestaties niet lager, maar integendeel, hoger zijn. Het bestaande onderwijssysteem zelf dient aan de orde te worden gesteld<sup>2)</sup> (Van Calcar et al. 1968, p. 257/258). Van Kemenade (1969) onderschrijft dat van harte en prijst deze benaderingswijze, maar waarschuwt ook tegen activeringsprogramma's, zoals die waar van Calcar c.s. voor pleiten, omdat ze een "spoedcursus middle-class cultuur" zijn (v. Kemenade, 1969, p. 56). Maar nu terug naar van Kemenade's tweede kritiekpunt. In een later artikel in de Sociologische Gids doet hij samen met Kropman (v. Kemenade en Kropman, 1972) verslag van een heranalyse van de van Jaar tot Jaar-data (Kropman en Collaris, 1974), waarbij leeftijd van de leerling en het beroep van de vader niet in de predictiescore waren opgenomen. Zij concluderen dan dat het resultaat van het Talentenproject "tot op zekere hoogte een artefact is van de wijze waarop "schoolgeschiktheid in het betreffende onderzoek is geoperationaliseerd". Het milieu is wel degelijk van invloed op de doorstroming naar het v.h.m.o. Soortgelijke kritiek is gekomen van Boon van Ostade (1972) die voor het Talentenproject en twee andere onderzoeken deelnameratios (feitelijke/verwachte aantallen leerlingen) berekent en tot de conclusie komt dat de gevolgtrekking uit het Talentenproject, dat de differentiële deelname aan het v.h.m.o. niet door het ouderlijk milieu maar door het verschil in prestatie wordt bepaald, in tegenspraak is met zijn bevin-

dingen in de beide andere onderzoeken (Beroepskeuzeonderzoek ITS en het Beslissingsproject Helmond). Bovendien vindt Boon de beperking tot de best-presterende leerlingen niet realistisch. Naar aanleiding van deze hernieuwde belangstelling voor het Talentenproject vecht ook Oosterbaan (1973) de belangrijkste conclusies aan. Hij analyseerde de gegevens van de Amsterdamse schooltoets 1969. Ook hij neemt een 'schone' prestatie maat (de score op de schooltoets) voor "schoolgeschiktheid". De vergelijkbaarheid van variabelen (voor het sociaal milieu b.v. is zijn keuze beperkt tot een indicator van de buurt waar de kinderen wonen) en de toetsingsmethode (Wilcoxon-toetsen) zijn voor kritiek vatbaar. Oosterbaan's conclusies m.b.t. de invloed van het sociaal milieu op doorstroming naar het v.h.m.o. en het advies van die onderwijzer staan diametraal tegenover die van Van Heek c.s.. In beide gevallen toont hij een milieuinvloed, onafhankelijk van de prestatiescores aan.

Onder veel voorbehoud toetst Oosterbaan dan vervolgens nog de uitspraak dat er in de lagere sociale milieu's onzekerheid ten opzichte van de hogere vormen van voortgezet onderwijs bestaat. Hij vindt dat in de handarbeidersmilieu's men zich ofwel strakker aan het gegeven advies houdt, ofwel naar een lagere vorm van v.o. doorstroomt. In de hogere milieugroepen daarentegen worden vaker hogere vormen van v.o. gekozen dan het advies aangaf.

In hun uiteindelijke verweer zijn van Heek, Vervoort en Van Weeren (1973) met de kritiek van Boon van Ostade gauw klaar. Als Boon van Ostade alle leerlingen die succesvol het lager onderwijs hebben doorlopen als geschikt voor het v.h.m.o. wil beschouwen, accoord, maar dat veronderstelt wel een ander v.h.m.o. dan ten tijde van het Talentenproject. Zijn bezwaren tegen de doorstromingsconclusie behandelen ze tegelijk met de kritiek van van Kemenade en Kropman, waar ze wat dieper op in gaan. Alles draait om de definitie van "schoolgeschiktheid" geven ze toe. In hun publikatie hebben ze onderscheid gemaakt tussen 'parate' en 'potentiële' schoolgeschiktheid<sup>3)</sup>. Parate schoolgeschiktheid is uitgedrukt in de kans een volgende onderwijsfase met succes te doorlopen. Potentiële geschiktheid (geschiktheid onder optimale pedagogische en didactische omstandigheden) is niet te meten. Hun conclusies beperken zich tot parate geschiktheid. "Verdergaande conclusies blijven voor rekening van degenen die ze trekken" (p. 312). Natuurlijk zouden zij, net als van Kemenade en Krop-

man, de operationele definitie van schoolgeschiktheid uitsluitend hebben kunnen baseren op prestatiegegevens. Het nadeel is echter "dat prestatiescore sec niet op zo'n goede voorspeller is van de kans op succes op het v.h.m.o.". Niettemin zijn ze van Kemenade en Kropman "dankbaar voor de waardevolle aanvulling, die zij op de resultaten van het Talentenproject geleverd hebben" (313). Op de kritiek van Oosterbaan gaan ze niet in.

In de afsluitende repliek van Kropman geeft hij aan dat van Kemenade en hijzelf hun kritiek overeind houden en dat milieufactoren, ten nadele van de lagere milieu's, van invloed zijn op lagere schoolprestaties, keuze van het type voortgezet onderwijs en de prestaties in het v.o.. Wat bitter sluit hij af met de opmerking dat onder "talent" (of schoolgeschiktheid), uitgaande van het bestaande onderwijssysteem, helaas ook het behoren tot een hoger milieu moet worden begrepen (Kropman, 1973). Later hebben Lammers en Nieman (1976) de data van Van Heek en van Kemenade en Kropman nog eens opnieuw geanalyseerd. Het ging hen er niet zozeer om de discussie nog eens opnieuw op te rakelen, als wel om een vergelijking van de meer gebruikelijke additieve analyse-techniek met de multiprecatieve volgens Goodman. Met beide technieken bleek dat de schoolprestaties de belangrijkste invloed op de doorstroming uitoefenen, maar dat los daarvan milieuverschillen nog een aanzienlijke invloed blijven behouden.

In het voorafgaande hebben we gezien, dat van Heek in zijn definitie van schoolgeschiktheid ten minste twee van zijn drie criteria voor maatschappelijk succes heeft verwerkt: aanleg en prestatie. Voor gelijkheid gaat dat, althans t.a.v. sexe-verschillen, niet op<sup>4</sup>). Meisjes worden op basis van een lagere predictiescore dan jongens aan de kansgroepen toegekend. Omdat ze in de groepen A, B en C zijn ondervertegenwoordigd, vermoedt van Weeren dat de grenzen van de kansgroepen voor meisjes te hoog zijn gelegd. Dat zou volgens hem leiden tot een onderschatting van eventuele reserve bij meisjes voor het v.h.m.o. (van Weeren, 1968, p. 140). Hij besteedt daar verder geen aandacht meer aan. Het is echter goed dit in gedachten te houden, waar onze heranalyses laten zien, dat de doorstroming van meisjes in de hogere kansgroepen en m.n. in de categorie "lagere employees", toch al geringer is dan die van jongens. We hopen met het voorafgaande een indruk gegeven te hebben van de "ideologische plaats" en het centrale begrip "schoolgeschiktheid" (en de kritiek daarop)

van het Talentenproject en gaan over tot de bespreking van de door ons verrichte heranalyses.

### 3. Heranalyses

We zijn geïnteresseerd in de relaties tussen begaafdheid/intelligentie, sociaal milieu en de schoolloopbaan. De schoolloopbaan-variabelen die we belangrijk vinden zijn: de prestaties in de 6e klas l.o., het advies van de onderwijzer voor het v.o., de eerste keuze v.o. en (voorzover beschikbaar) het uiteindelijk bereikt niveau in het voortgezet onderwijs.

Voor onze heranalyses van het Talentenproject en andere prominente schoolloopbaan-onderzoeken (Matthijssen en Sonnemans, 1958; Peschar, 1975) kiezen we dus steeds die variabelen die zoveel mogelijk vergelijkbaar zijn met de hierboven opgesomde.

Omdat we onze heranalyse beperken tot gepubliceerde gegevens moet uit de gepresenteerde resultaten minimaal een ruwe kruistabel te herleiden zijn. Voor wat betreft het Talentenproject komen dan de tabellen in de deelrapportage van van Weeren in aanmerking. We vinden daar per regio een kruistabel (gepercenteerd) van milieu, sexe, kansgroep en wel of niet doorstromen naar het v.h.m.o. (Van Heek e.a., 1968, p. 131-135), een kruistabel van milieu, sexe en het advies van de onderwijzer (idem, p. 146-147) en een tabel van milieu, sexe, advies en keuze v.o. (idem, p. 149). Deze tabellen tesamen vormen het basismateriaal voor onze bewerkingen.

Voor details omtrent de regionale verzameling van gegevens, ophoging van de aantallen voor Twente en het als v.h.m.o. beschouwen van de brugschool in Friesland verwijzen we naar van Weeren, de categorisering van de variabelen bespreken we apart.

#### 3.1. De variabelen

Voor het sociaal milieu wordt het beroep van de vader als indicator genomen. In het Talentenproject heeft deze variabele oorspronkelijk 9 categorieën:



- 1 = academische vrije beroepen
- 2 = leidinggevende beroepen bij overheid en bedrijfsleven
- 3 = middenkader, niet-academische vrije beroepen en daarmee gelijk te stellen functies
- 4 = uitvoerende hoofdarbeid en daarmee gelijk te stellen functies
- 5 = geschoolde arbeid
- 6 = ongeschoolde arbeid
- 7 = landarbeid
- 8 = zelfstandige middenstand
- 9 = boeren en tuinders
- + een restcategorie 'onbekend'

De categorieën 1, 2 en 3 en 5, 6 en 7 worden door van Weeren samengenomen, zodat 5 klassen overblijven. We benoemen de klassen analoog aan de in het SMVO-onderzoek gehanteerde categorieën en herorderen volgens de door de Leeuw en Stoop (1978) gevonden optimale categorieënvolgorde.

We krijgen dan:

- 1 = arbeiders
- 2 = boeren
- 3 = middenstand
- 4 = lagere employees
- 5 = hogere en middelbare employees

De categorie 'milieu onbekend' laten we vallen. Bij de door ons gehanteerde technieken wordt deze categorie, na optimale transformatie toch in of dichtbij het nulpunt geplaatst. De meeste kinderen waarvan de sociale achtergrond niet bekend is vallen in de laagste kansgroepklasse die niet naar het v.h.m.o. doorstroomt.

We voegen de variabele sexe samen met de milieuvariabele tot één gecombineerde milieu-geslacht variabele met 10 categorieën, niet alleen omdat we het handig vinden om met een zo klein mogelijk aantal variabelen te werken, maar vooral omdat we veronderstellen dat sexe-verschillen binnen de onderscheiden milieuklassen verschillende betekenis kunnen hebben<sup>5)</sup>. Eventuele interactie-effecten tussen milieu en sexe houden we op deze manier onder controle.

De variabele kansgroep die intelligentie, prestatie en, niet te vergeten, helaas ook het milieu, advies v.o. en leeftijd combineert, hebben

we in de vorige paragraaf al uitgebreid besproken. We handhaven de oorspronkelijke categorieën A, B, C en D.

Het advies van de onderwijzer en de keuze van het voortgezet onderwijs hebben in de analyses van van Weeren niet steeds dezelfde categorieën. We zullen de klasse-indeling van deze variabelen bij de afzonderlijke analyses nader toe lichten.

### 3.2. Verschillen in doorstroming naar het v.h.m.o.

We volgen in onze heranalyses de systematiek van van Weeren's rapportage en bekijken eerst de verschillen in doorstroming naar sociaal milieu en kansgroep voor de vier regio's Friesland, Twente, Brabant en Amsterdam apart. De variabele 'keuze v.o.' heeft in deze tabellen twee categorieën: wèl doorstromen naar het v.h.m.o. (WD) en niet doorstromen (ND). Als v.h.m.o. wordt in het Talentenproject beschouwd: gymnasium, h.b.s., lyceum en voor de jongens ook seminarie.

We passen op deze tabellen correspondentie-analyse toe. In het algemeen is correspondentie-analyse geschikt voor analyseren van 2-dimensionale kruistabellen, maar als we de datamatrix herordenen, zodanig dat hij uit 2 rijen (of 2 kolommen, dat maakt niet uit) bestaat, wordt de afstand tussen deze rijen (WD en ND) gemaximaliseerd en worden de kolom-variabelen (kansgroep, sociaal milieu en geslacht) allemaal op één dimensie getransformeerd.

M.a.w. de milieu-geslachtsgroepen, uitgesplitst naar hun kansgroepscore, worden steeds vergeleken naar wel en niet doorstromen. En wel zodanig dat de milieugroepen worden verdeeld in het percentage WD en ND. Dit is een bijkomend effect van deze toepassing van de techniek en dat wordt in Appendix B nader uiteengezet.

Het resultaat staat in fig. 1 t/m 4. Naarmate een groep hoger in het plaatje ligt, gaat een groter percentage leerlingen uit zo'n groep naar het v.h.m.o.

De milieu-groepen zijn (jongens en meisjes apart) onderling met elkaar verbonden als ze tot dezelfde kansgroep (A, B, C, D) behoren. De figuren zijn alle vier op dezelfde schaal getekend, zodat ze a.h.w. over elkaar heen gelegd kunnen worden om de regio's onderling naar doorstroming te vergelijken. Het eerste wat opvalt is dat er nogal

grote regionale verschillen in doorstroming lijken te bestaan. We moeten daarbij wel bedenken dat, naarmate een kansgroep minder leerlingen bevat, de relatieve verschillen tussen de milieugroepen binnen zo'n kansgroep groot kunnen worden en ze in het plaatje verder uit elkaar worden gelegd. Het omgekeerde zien we duidelijk bij kansgroep D die het grootste gedeelte van de leerlingen bevat. De relatieve vulling van de milieugroepen verschilt binnen deze groep weinig en de lijnen voor deze groep hebben dus een veel 'vlakker' verloop dan voor de andere kansgroepen. Hoe belangrijk zijn nu de regionale verschillen? Van Weeren beschouwt de regionale verschillen van beperkt belang en volgens Ruiter, Eisma en Passenier hangen de regionale verschillen in deelneming aan het v.h.m.o. grotendeels samen met regionale verschillen in de samenstelling van de bevolking naar beroepsgroepen (Ruiter et al., 1968).

Om enig inzicht te krijgen in het belang van de variabele regio, hebben we de totale datamatrix die een samenvoeging is van de 4 regionale matrices, nog eens opnieuw geanalyseerd. We hebben de variabelen optimaal getransformeerd en correlaties tussen deze kwantificaties berekend. In fig. 5 zijn de kwantificaties geplott. Voor deze kwantificaties zijn de regressies vrijwel lineair. De correlaties zullen daarom de samenhang tussen elk paar variabelen goed weergeven. De correlaties staan in tabel 1.

Tabel 1: Correlaties tussen gekwantificeerde variabelen

	1	2	3	4
1. doorstroming	1.00			
2. kansgroep	.67	1.00		
3. milieu-sexe	.32	.37	1.00	
4. regio	.06	.08	.13	1.00

De correlatie van regio met de andere variabele is zo laag, dat we het gerechtvaardigd vinden om regio verder buiten beschouwing te laten en alleen nog optellingen over de vier regio's te analyseren. Van de correlaties met de andere variabelen, is die met het sociaal milieu nog het hoogst, wat consistent is met de bevindingen van Ruiter et al.

De gekwadraterde multiële correlatiecoëfficiënt tussen doorstroming en de overige variabelen is .46 voor deze variabelen. Het aandeel van

het verband tussen kansgroep en doorstroming daarin is  $(.67)^2 = .45$ . Doorstroming wordt dus vrijwel volledig door de predictiescore voorspeld. Dat is niet verwonderlijk, gezien het opnemen van sociaal milieu in de kansgroepvariabele en correlaties van milieu met de andere twee.

De verbanden met de milieu-sexe variabele zijn overigens niet rechtstreeks te interpreteren, vanwege het mogelijk interactie-effect van milieu en sexe. Om dat nader te bezien en te verduidelijken, hebben we fig. 6 gemaakt. Deze grafiek is een nettere versie van de milieu-sexe plotjes uit fig. 5. De getransformeerde scores van de milieu-sexe variabele (met inachtneming van kansgroep, doorstroming en regio) zijn voor jongens en meisjes apart uitgezet. Uit het vrijwel parallel verlopen van de curves valt af te leiden, dat het interactie-effect van milieu en sexe heel gering is. We hadden in dit geval milieu en sexe dus ook wel als afzonderlijke variabelen kunnen behandelen en de correlaties van de andere variabelen met milieu alleen zullen niet veel verschillen van die met de gecombineerde variabele.

Verder laat de figuur nog zien, dat als categorieën 'middenstand' en 'lagere employees' zouden worden omgewisseld, de milieu-variabele monotoon zou worden getransformeerd. De 'ideale' volgorde komt overeen met de volgorde die van Weeren vond, als hij de milieugroepen naar doorstromingspercentage en gemiddelde predictiescore rangschikte.

Keren we nu terug naar onze correspondentie-analyse. We menen nu de regionale verschillen te kunnen verwaarlozen en hebben de data-matrix over de regio's opgeteld. De doorstromingsverschillen zijn fig. 7 gevisualiseerd. Dit plaatje is weer op dezelfde schaal getekend als de afzonderlijke regionale grafieken ( en kan daar dus mee vergeleken worden) en laat de verschillen in doorstroming goed zien. Naarmate de predictiescore (kansgroep) hoger is, stromen de leerlingen meer door naar het v.h.m.o.. In de onderste groep gaan de kinderen van hogere en middelbare employees meer naar het v.h.m.o. dan de andere milieugroepen. Dochters van lagere employees blijven in alle kansgroepen achter qua doorstroming. De dochters van hogere en middelbare employees stromen in alle kansgroepen meer door. Alleen bij de meest begaafde, best presterende jongens is er weinig verschil tussen de milieugroepen. De kansgroep gaf de kans van slagen in het v.h.m.o. in percentages weer. Omdat de plaats van een milieugroep in het plaatje het percentage leerlingen aangeeft dat naar het v.h.m.o.

doorstroomt (zie p. 9 en appendix B) hebben we in de figuur horizontale lijnen getrokken voor de kansgroepgrenzen. De leerlingen stromen dus vrijwel precies in die proporties door als ze volgens de predictiescore kans van slagen hebben. Daarop berust de conclusie van van Heek c.s. dat er geen "talentenreserve" bestaat. We beschouwen dit als het gevolg van het meerekenen van het advies van de onderwijzer en de sociaal milieu-indicator in de predictiescore.

### 3.3 Het sociaal milieu en het advies van de onderwijzer

Uitgangspunt voor onze herbewerking is de tabel op p. 146-147 van Van Weeren. We reconstrueren een kruistabel van milieu-sexe en het advies van de onderwijzer. Op grond van de in de vorige paragraaf gerapporteerde bevindingen tellen we over de regio's op. Van de advies-variabele nemen we 'lagere land- en tuinbouwschool<sup>6)</sup> samen met 'lager nijverheids onderwijs' tot één categorie LBO. Omdat de categorie 'm.m.s.' bij jongens uiteraard niet voorkomt, voegen we deze samen met v.h.m.o.. De categorie 'ander advies' is weinig gevuld en kunnen we niet interpreteren; die laten we weg. Advies heeft op die manier uiteindelijk 4 categorieën: LBO, VGLO, ULO, VHMO.

Het resultaat van de correspondentie-analyse laat zien dat kinderen van boeren en arbeiders vaak het advies LBO of VGLO krijgen. Voor kinderen van middenstanders en lagere employees is naar het oordeel van de onderwijzer de ULO de meest geschikte opleiding. VHMO-adviezen komen het vaakst voor in de categorie 'hogere en middelbare employees'. In fig. 8 en 9 zijn de optimale transformaties van advies en sociaal milieu in beeld gebracht. Het advies van de onderwijzer wordt vrijwel lineair getransformeerd. Vergelijking van fig. 8 en 9 geeft een indruk van het verband tussen het ouderlijk milieu en het advies van de onderwijzer. Hoe hoger het milieu, hoe hoger het advies voor het voortgezet onderwijs. Fig. 9 laat ook zien dat het interactie-effect van milieu en sexe vooral gezocht moet worden in de 'lagere' klassen. De afstand tussen de punten voor jongens en meisjes is steeds ongeveer gelijk, behalve in de laagste milieu's, waar jongens en meisjes dicht bij elkaar liggen.

We kunnen de resultaten van de analyse nog op een andere manier in beeld brengen. In fig. 10 hebben we langs de X-as de optimale trans-

formaties van het onderwijzersadvies geplaatst en langs de Y-as de gekwantificeerde categorieën van de onderscheiden milieu-sexegroepen. De getallen in de figuur zijn de oorspronkelijke celfrequenties uit de kruistabel en de lijnen geven de regressies van milieu op advies en van advies op milieu weer. De hoek die de regressielijnen maken, geeft de sterkte van het verband aan. (De correlatie tussen de getransformeerde variabelen is .36). Op deze manier is alle informatie die deze data bieden in één plaatje samengevat. In appendix A wordt deze toepassing toegelicht.

### 3.4 Het advies van de onderwijzer en de schoolkeuze in de diverse sociale milieu's

Uitgangspunt is de tabel op p. 149. Analoog aan de werkwijze in de vorige paragrafen, tellen we weer op over de vier regio's. De variabele 'keuze voortgezet onderwijs' moeten we dan ook hergroeperen. De keuzen m.m.s. en h.a.v.o. komen alleen in Amsterdam en Brabant voor. Deze categorieën voegen we samen met v.h.m.o. Door de percentages voor keuze ulo en keuze vhmo op te tellen en van 100% af te trekken krijgen we een derde categorie: 'andere keuze'. De variabele 'advies onderwijzer' heeft alleen de categorieën ulo en vhmo in deze tabel. We combineren beide variabelen tot één met 6 (3x2) categorieën die overeenstemming of discrepantie tussen advies en keuze aangeven: ulo-advies/ulo-keuze, ulo-advies/vhmo-keuze, ulo-advies/andere keuze, vhmo-advies/vhmo-keuze, vhmo-advies/ulo-keuze en vhmo-advies/andere keuze. We bereken de oorspronkelijke celfrequenties terug en analyseren de aldus verkregen tabel.

Op de data van het CBS-SMVO cohortonderzoek heeft de Glopper (1981) een vergelijkbare analyse gedaan. Met behulp van log-lineaire modellen analyseert hij het verband tussen sociaal milieu, het advies van de onderwijzer en discrepantie/overeenstemming van advies en keuze v.o. De laatstgenoemde variabele heeft bij hem 3 categorieën: advies lager dan plaatsing, overeenstemming advies en plaatsing en advies hoger dan plaatsing. Daardoor gaat, volgens ons, te veel nuancering van de discrepanties verloren en daarom kiezen we voor de boven aangeduide categorie-indeling. In fig. 11 en 12 staan de categorie-kwantificaties voor respectievelijk de milieu-sexe en de advies-keu-

ze variabele. Fig. 11 laat zien dat het verschil tussen jongens en meisjes in de lagere milieu's groter is dan in de andere milieugroepen. Vergelijking van beide plaatjes toont ons, dat in de hogere en middelbare milieu's de vhmoadvies het vaakst worden opgevolgd en dat bovendien na een uloadvies toch vaak een vhmokeuze wordt gemaakt. De uloadvies worden in de lagere sociale milieu's het meest opgevolgd en in de hogere het minst. In de lagere milieu's wordt na een uloadvies nogal eens een andere opleiding gekozen. We nemen aan dat dat dan het lager beroepsonderwijs (LTS, Huishoudschool) betreft. In het algemeen geldt, dat hoe hoger een advies/keuze-combinatie in de figuur ligt, hoe vaker een dergelijke combinatie bij de hogere milieugroepen voorkomt en omgekeerd.

De combinatie vhmoadvies/ulo-keuze ligt in de figuur ongeveer in het midden. Deze combinatie is dus niet expliciet met één van de milieugroepen te associëren. Het is echter een combinatie die maar weinig voorkomt.

Ter aanvulling en vergelijking voegen we tabel 2 toe, waaruit we de gestandaardiseerde afwijkingen ( $\chi$ ) van de verwachte celfrequenties kunnen lezen ( $\chi^2$  is 850.17 bij  $df=45$ , wat zeer significant is). De tabel geeft nog eens in detail weer wat we globaal al in de plaatjes hadden gezien.

Tabel 2. Gestandaardiseerde afwijkingen ( $\chi$ ) voor milieu-sexe en advies-keuze.

	ulo/ulo	ulo/vhmo	ulo/and	vhmo/ulo	vhmo/vhmo	vhmo/and.	totaal
<b>jongens</b>							
Arbeiders	1.7	-4.4	5.1	0.3	-3.3	0.0	-0.6
Boeren	-0.0	0.6	1.3	1.0	-1.6	-0.2	1.1
Middenst.	-1.5	1.4	-1.1	0.1	2.0	0.2	1.1
L.empl.	1.8	-1.0	-2.1	-0.8	-0.3	-0.6	-2.9
H./M. empl.	-10.2	8.1	-4.9	1.0	12.4	1.9	8.4
<b>meisjes</b>							
Arbeiders	7.4	-5.8	5.3	-0.7	-10.4	-1.1	-5.3
Boeren	2.3	-3.2	4.5	1.0	-4.8	0.1	-0.1
Middenst.	2.8	-0.4	-0.5	0.6	-3.6	-1.0	-2.0
L.empl.	3.5	-2.7	-0.1	0.5	-3.5	-0.6	-2.9
H/M empl.	-2.5	3.7	-5.1	-2.0	5.5	0.2	-0.1
<b>Totaal</b>	<b>5.4</b>	<b>-3.7</b>	<b>2.4</b>	<b>1.0</b>	<b>-7.5</b>	<b>-1.1</b>	<b>-3.5</b>

Ook na vmo-advies wordt nog wel eens een 'andere' keuze gemaakt. Het vaakst door jongens uit de hogere milieu's. Interpretatie hiervan is uiteraard speculatief. We vermoeden dat het hier ook om een technische opleiding gaat. Overigens komt deze combinatie van advies en keuze slechts bij een zeer klein gedeelte van de leerlingen voor.

### 3.5 De invloed van het sociaal milieu en het geslacht van de leerlingen op advies voor en doorstroming naar het voortgezet onderwijs

De in de vorige paragrafen beschreven analyses laten zien dat de sociale achtergrond van de onderzochte leerlingen een belangrijke invloed uitoefent bij de overgang van het lager naar het voortgezet onderwijs. We kunnen ons afvragen of het geslacht van de leerlingen er ook nog iets toe doet of dat er nog interactie tussen milieu en geslacht is. Anders gezegd of de milieu-invloed voor jongens en meisjes anders werkt. We vergelijken daartoe de figuren 6, 9 en 11 nog eens met elkaar. Het eerste wat opvalt is dat de kwantificaties van het sociaal milieu afhankelijk zijn van het stelsel van variabelen dat wordt geanalyseerd. Ten tweede worden meisjes vrijwel steeds lager geplaatst dan jongens. Meisjes stromen dus minder door en krijgen lagere adviezen dan jongens. Van Weeren stelde (p. 140) dat de reserve's bij meisjes waarschijnlijk enigszins zouden worden onderschat. Los daarvan blijken ze dus toch al minder door te stromen. Wat doorstroming betreft (fig. 6) zien we verder geen effect van milieu en geslacht samen (de lijnen lopen vrijwel parallel). Bij de adviezen van de onderwijzer zien we in de lagere sociale milieu's wel een samenhang: als een kind uit een arbeidersmilieu afkomstig is maakt het de onderwijzer weinig of hij met een jongen of een meisje te doen heeft. Dat jongens uit de categorie 'boeren' zo laag geplaatst worden heeft er mee te maken dat deze jongens van de onderwijzer vaak het advies krijgen naar de lagere land- en tuinbouwschool te gaan. Het advies van de onderwijzer zal in dit geval ook wel enigszins beïnvloed zijn door de wens van de ouders. In het agrarische milieu leek in 1968 toch weinig belangstelling voor een algemeen vormende opleiding te bestaan. Dat vinden we terug in fig. 11. De keuze 'anders' (waarvan we veronderstellen dat het de keuze voor LBO of VGLO is) na een ulo-advies komt in de categorie 'boeren' relatief vaak voor, zeker als het de meisjes aangaat.



Wanneer we dus uitsluitend naar de doorstroming kijken en de betreffende tabellen, waarin het grootste gedeelte van de variabele is vercalculeerd, analyseren, constateren we dat meisjes minder doorstromen dan jongens maar vinden we verder geen specifieke werking van geslacht binnen de milieugroepen. Ten aanzien van het advies van de onderwijzer, alleen of in samenhang met de feitelijke keuze v.o. is er wel een dergelijke interactie in de "lagere" milieu's (arbeiders en boeren).

#### 4. Vergelijking van resultaten

Globaal gesproken is er een grote overeenstemming tussen onze bevindingen en die van van Weeren:

- Er is een nauwe samenhang tussen de doorstroming van het vhm0 en de sociale achtergrond van de leerlingen.
- De voorspellingsscore voor het vhm0 heeft een positief verband met het beroepsmilieu.
- Er bestaat een sterk verband tussen de voorspellingsscore en de doorstroming naar het vhm0.
- Het oordeel van de onderwijzer, het beroepsmilieu en de feitelijke schoolkeuze hangen met elkaar samen.
- De regionale verschillen zijn van (zeer) geringe betekenis.

We menen echter dat de door ons uitgevoerde analyses een belangrijke aanvulling op en detaillering van van Weeren's resultaten leveren. Dat komt op de eerste plaats door de door ons toegepaste technieken. Doordat in onze analyses de variabelen optimaal herschaald worden, krijgen we meer inzicht in de structuur van de variabelen. De ordening b.v. van de milieuvariabele die van Weeren vindt op basis van vergelijking van percentages en gemiddelden krijgt op deze manier een heldere onderbouwing en tevens weten we iets over het onderscheidingsvermogen van de categorieën van de variabelen. Als blijkt dat de structuur van een paar variabelen dat toelaat, kan het verband zonder bezwaar in een getal (de correlatie) worden uitgedrukt. Onze analyses tonen op overtuigende wijze de ondergeschikte betekenis van de verschillen tussen de regio's aan. Dat maakt het mogelijk meer ge-

aggregeerde tabellen te analyseren en dat leidt tot betrouwbaarder resultaten.

Het verband tussen milieu, doorstroming en predictiescore kon door de gevolgde werkwijze op een heldere manier in één grafiek (fig. 7) worden samengevat. Persoonlijk vinden wij deze presentatie veel inzichtelijker dan de nauwkeurige beschrijving van vele pagina's vol gepercenteerde kruistabellen. Genoemde figuur laat ook zien waar (om in termen van het Talentenproject te spreken) de "reserve" vooral moet worden gezocht. Niet zozeer bij arbeiderskinderen, zoals men bij het Talentenproject veronderstelde (hoewel zoons van arbeiders en boeren in kansgroep B bij de andere milieu's achterblijven) maar vooral bij de dochters van lagere employeés.

Van Weeren's conclusie dat de onderwijzer zich bij zijn advies niet laat leiden door milieu-invloeden, moeten we bestrijden. Bij zijn analyse van het materiaal van de Amsterdamse schooltoets van 1969 had Oosterbaan (1973) ook al een dergelijk verband gevonden. Wij vinden dat nog eens bevestigd aan de hand van de data van het Talentenproject zelf. Er is wel degelijk een samenhang tussen het sociaal milieu en het advies van de onderwijzer (zie fig. 9). We vermoeden dat van Weeren, gegeven het overwegend effect van 'kansgroep' (waarin het advies mede is opgenomen) op de doorstroming, de neiging heeft gehad alle andere verbanden in het materiaal te bagatelliseren.

Tenslotte nog eens het sterke vergand dat er tussen predictiescore en de doorstroming naar het vmo bestaat. Dat dat voornamelijk door het opnemen van het advies van de onderwijzer en het beroepsmilieu in de predictiescore wordt veroorzaakt, wordt ook in deze heranalyses nog eens aangetoond.

Dat er desondanks nog verschillen in doorstroming tussen de sociale milieu's te vinden zijn, bevestigt de juistheid van de kritiek die in dit opzicht op het Talentenproject is uitgeoefend en geeft aan wat een overwegende invloed de sociale achtergrond van de leerlingen op de doorstroming naar het voortgezet onderwijs in de jaren '60 had.

## NOTEN

1. Die overigens door van Kemenade in zijn eerste reactie (van Kemenade, 1969) gekwalificeerd wordt als "een vreemdsoortige verzameling van hypothesen, ideologieën en beleidsoplossingen".
2. Dat doet de tweede fase van het Talentenproject meteen twee plaatsen opschuiven op het continuüm van Matthijssen (vgl. Matthijssen, 1975).
3. Dat doet denken aan het onderscheiden dat in de publicaties van het Demologisch Instituut wordt gemaakt tussen "parate" en "potentiële" intelligentie (Demologisch Instituut, 1957), waaraan van Heek c.s. overigens niet refereren.
4. Terzijde zij opgemerkt, dat het onduidelijk blijft wat van Heek met gelijkheid bedoelt. Na de introductie van het begrip in zijn inleiding, waarin hij aangeeft dat het begrip op verschillende kenmerken (afkomst, erfelijke aanleg etc.) betrekking kan hebben, komt hij er eigenlijk niet meer op terug.
5. Gifi (1980, 1981) heeft verschillende analyses op vergelijkbare gegevens (het CBS-SMVO cohort en de van Jaar tot Jaar data) gedaan, waarvan de resultaten een dergelijk effect suggereren. Zie Gifi (1980, p. 257/258) en Gifi (1981, p. 366).
6. In de regio's Friesland, Twente en Brabant wordt deze categorie vrijwel uitsluitend door boerenzoons gevuld, in Amsterdam is hij geheel leeg.

## LITERATUUR

- Boon van Ostade, A.H., 'Het milieu en de overgang naar het voortgezet onderwijs in Nederland', *Sociologische Gids*, 19 (1972), p. 211-218.
- Calcar, C. van, S. Soutendijk en B. Tellegen, 'School, milieu en prestatie', in: F. van Heek e.a., 'Het verborgen talent', Meppel, 1968.
- Demologisch Instituut, 'Begaafdheidsonderzoek en intelligentiespreiding deel II', *Intelligentie en Arbeidspotentieel*, Utrecht, 1957.
- Gifi, A., 'Niet-lineaire multivariate analyse', Leiden, 1981.
- Gifi, A., 'Non-linear multivariate analysis', Leiden, 1981.
- Glopper, K. de, 'Niet-overeenstemming van onderwijzersadviezen en plaatsingen in het voorgezet onderwijs', Paper bijvak Methoden en Statiek, Amsterdam, 1981.
- Heek, F. van, 'Het verborgen talent', Meppel, 1972.
- Heek, F. van, C.E. Vervoort en P. van Weeren, 'Nogmaals: verborgen talenten?', *Sociologische Gids*, 20 (1973), p. 309-316.
- Idenburg, Ph.J., 'Gelijkheid van kansen', *Paedagogische Studien*, 1959, p. 210-224.
- Idenburg, Ph.J., 'Het ideaal van de optimale ontwikkeling der talenten en de paedagogische structuur van het schoolweten', *Paedagogische Studien*, 1962.
- Kemenade, J.A. van, 'Het verborgen talent; een programma voor onderzoek', *Sociologische Gids*, 16 (1969), p. 51-57.

- Kemenade, J.A. van en J.A. Kropman, 'Verborgen talenten? Kritische kanttekeningen bij een onjuiste interpretatie', *Sociologische Gids*, 19 (1972), p. 219-228.
- Kropman, J.H. en J.W.M. Collaris, 'Van jaar tot jaar, eerste fase', ITS, Nijmegen, 1974.
- Lammers, J.G.M. en F.H.M. Nieman, 'Additieve versus multiplicatieve modellen in de causale analyse; een heranalyse van gegevens rond "Het verborgen talent"', *Sociale Wetenschappen*, 19 (1976), p. 193-213.
- Leeuw, J. de en I. Stoop, 'Secondaire analyse 'Van jaar tot jaar' met behulp van niet-lineaire multivariate technieken', in: J.L. Peschar (red.), 'Van achteren naar voren', 's-Gravenhage, 1979.
- Matthijssen, M.A.J.M., 'Sociale ongelijkheid van onderwijskansen; vijf interpretaties van sociale werkelijkheid', *Sociologische Gids*, 22 (1975), p. 427-445.
- Matthijssen, M.A.J.M. en G.J.M. Sonnemans, 'Schoolkeuze en schoolsucces bij vmo en ulo in Noord-Brabant', *Paedagogische publicaties nr. 3*, Tilburg, 1958.
- Oosterbaan, J.W., 'De Amsterdamse schooltoetst 1969 in relatie tot enige conclusies van 'Het verborgen talent'', *Sociologische Gids*, 20 (1973), p. 88-97.
- Peschar, J., 'Milieu, school, beroep', Groningen, 1978 (diss.).
- Ruiter, R., M. Eisma en J. Passenier, 'Factoren die de regionale verschillen in deelneming aan het vmo in Nederland beïnvloeden', in: F. van Heek e.a., 'Het verborgen talent', Meppel, 1968.
- Weeren, P. van, 'Uitkomsten van het psychologisch onderzoek omtrent milieu, schoolkeuze en schoolgeschiktheid in Amsterdam, Twente, Friesland en Noord-Brabant', in: F. van Heek, 'Het verborgen talent', Meppel, 1968.

## Appendix A: Correspondentie Analyse

We geven slechts een korte inleiding. Voor gedetailleerde besprekingen van de techniek verwijzen we naar Nishisato (1980), Gifi (1981), Benzécri e.a. (1973), Benzécri e.a. (1980).

In de vorm waarin wij het gebruiken is correspondentie-analyse een techniek om de correlatie tussen twee geordende variabelen te schatten. De techniek werkt uitstekend wanneer de geobserveerde kruistabel beschouwd kan worden als een gediscretiseerde versie van een continue bivariate verdeling met lineaire regressies, bijvoorbeeld als een gediscretiseerde bivariate normaalverdeling. Aan te bevelen is om de discretisatie zo fijn mogelijk te maken, op voorwaarde dat de cellen redelijk gevuld blijven. In het algemeen kunnen we zeggen dat correspondentie-analyse een goede schatting van de correlatiecoëfficiënt geeft in die gevallen waarin het zinvol is associatie tussen twee variabelen met produkt-moment methoden te meten. Wanneer de discretiseringsen afwijken van het interval-niveau, met name wanneer er sprake is van aanzienlijke scheefheid, dan zal correspondentie analyse een schatting van de correlatiecoëfficiënt opleveren die verschilt van de gebruikelijke. De verschillen zijn over het algemeen niet groot, maar een lichte verbetering is er in de meeste gevallen wel.

Rekenkundig gezien berekent correspondentie analyse de grootst mogelijke correlatie die door scoring van de variabelen bereikt kan worden. Een bijproduct zijn de scoringen (of transformaties of kwantificaties) van de variabelen. De gevonden transformaties zijn steeds zodanig, dat ze de beide regressies exact lineariseren. We kunnen correspondentie analyse dus ook formuleren als een techniek die scores vindt die de regressies exact lineariseren. Uit de oplossingen van dit probleem wordt dan vervolgens door ons degene gekozen met de maximale correlatie.

In een plaatje kunnen we de beide regressielijnen uitzetten. De maximale correlatie (een schatting van de 'ware' correlatie) wordt weergegeven door de hoek tussen de regressielijnen. Op de assen van het plaatje staan de herschaalde variabelen. In een dergelijk plaatje kun-

nen we ook de cellen met hun celinhoud weergeven, zodat de kruistabel als het ware op een compacte manier grafisch weergegeven wordt.

Aan het plaatje is niet te zien of de associatie in de tabel op zinvolle manier met een produkt-moment maat weer te geven is. Een noodzakelijke voorwaarde is dat de categorieën van de variabelen in de 'juiste' volgorde geschaald worden, maar voldoende is deze voorwaarde zeker niet. De vraag kan beantwoord worden door naar de overige oplossingen van correspondentie analyse te kijken, die andere transformaties geven om de regressies te lineariseren. In de door ons geanalyseerde voorbeelden voegen deze overige oplossingen nauwelijks informatie toe aan de eerste en maximale oplossing. Ze kunnen daarom gevoeglijk weggelaten worden.

Appendix B: Correspondentie analyse op een tabel met twee rijen

Stel  $N = \{n_{ij}\}$  is een kruistabel met twee rijen. Dus  $i=1,2$  en  $j=1,\dots,m$ . De rijtotalen zijn  $n_+$  en  $n_-$ , ze tellen op tot  $n$ , het totale aantal observaties. De kolomtotalen noemen we  $n_j$ .

Correspondentieanalyse vindt, zoals gebruikelijk, rijsscores en kolomscores in afwijkingen van het gemiddelde die de regressies linealiseren. Omdat er maar twee rijen zijn, worden de rijsscores tot op een schaal-factor bepaald door de eis dat ze in afwijkingen van het gemiddelde staan. We vinden  $y_+ = -\alpha n_-$  en  $y_- = \alpha n_+$ , waarbij  $y_+$  en  $y_-$  de scores voor de eerste en tweede rij zijn, en waarbij  $\alpha$  de schaal-factor is. De  $m$  kolomscores vinden we door de kolomgemiddelden te berekenen.

Dus

$$x_j = (-\alpha n_- n_{1j} + \alpha n_+ n_{2j})/n_j.$$

Wanneer  $p_j = n_{1j}/n_j$ , dus  $p_j$  is de proportie positieven in kolom  $j$ , en  $q_j = 1 - p_j = n_{2j}/n_j$ , dan

$$x_j = \alpha(n_+ q_j - n_- p_j).$$

Als  $p = n_+/n$ , dus  $p$  is de proportie positieven in de gehele tabel, en  $q = 1 - p = n_-/n$ , dan

$$x_j = n\alpha(pq_j - qp_j) = n\alpha\{p(1 - p_j) - (1 - p)p_j\} = n\alpha(p - p_j) = n\alpha(q_j - q).$$

Dus: de kolomscores van correspondentieanalyse zijn een eenvoudige lineaire transformatie van de proporties positief in een kolom (dus van de rijen van een kolomsgewijs gepercenteerde versie van de kruistabel).