

STRUCTURELE MODELLEN VOOR MULTINIVEAU DATA.¹**Joop J. Hox²****Samenvatting**

Veel gegevens in sociaalwetenschappelijk en economisch onderzoek kunnen worden opgevat als gegevens met een hiërarchische structuur, waarbij steekproefgegevens verzameld worden op zowel het hogere (groeps-, organisatie-) niveau, als het lagere (individuele) niveau. Hiërarchische regressiemodellen zijn voor dit type gegevens in de literatuur inmiddels beschreven in verscheidene artikelen en boeken, en efficiënte computerprogramma's (zoals HLM, ML3, VARCL) zijn al enige tijd beschikbaar. Analysemodellen voor een structureel model met latente variabelen bij gegevens met een hiërarchische structuur zijn in de literatuur wel beschreven, maar het ontbreekt tot op heden aan efficiënte rekenprocedures en gespecialiseerde programmatuur. Dit artikel beschrijft een benaderende oplossing van het probleem, en geeft een voorbeeld van een multiniveau factoranalyse en een multiniveau padanalyse. Hoewel een gespecialiseerd computerprogramma voordelen zou hebben, is het ook mogelijk conventionele software te gebruiken. Het artikel bespreekt aan het slot hoe het model in Lisrel geïmplementeerd kan worden.

¹Met dank aan Timo Bechger, Edith de Leeuw en een anonieme referent voor hun commentaar op een eerdere versie.

²Faculteit Pedagogische Wetenschappen, Universiteit van Amsterdam. IJsbaanpad 9, 1076 CV Amsterdam. Tel. 020-6223438 Internet A716HOX @ HASARA11

Inleiding

In sociaalwetenschappelijk onderzoek worden onderzoekers regelmatig geconfronteerd met gegevens die een hiërarchische structuur vertonen. Het schoolsysteem is een goed voorbeeld: leerlingen bevinden zich in klassen, die weer deel uitmaken van scholen. Op elk van deze drie niveaus kunnen gegevens verzameld worden, en het onderzoeksprobleem betreft ook dikwijls expliciet de vraag welke verbanden er zijn tussen school-, klas- en leerlingvariabelen. De data weerspiegelen de hiërarchische structuur van het schoolsysteem, en voor een adequate analyse van deze gegevens dienen dan ook hiërarchische ofwel *multiniveau* modellen toegepast te worden.

Ook andere sociale systemen worden gekenmerkt door een hiërarchische structuur. Individuen zijn op verschillende wijzen georganiseerd in groepen, die weer onderdeel zijn van grotere (bijvoorbeeld geografische) eenheden. Nu er met het oog op de eenwording van Europa steeds meer onderzoek wordt uitgevoerd in verschillende landen tegelijk, zouden multiniveau modellen gebruikt kunnen worden om systematische verschillen tussen survey-organisaties en landen te analyseren, en tot op zekere hoogte te controleren (vgl. Hox et al., 1991a; Hox, 1992). Verder zijn er specifieke onderzoekopzetten die leiden tot multiniveau gegevens. Een voorbeeld is panel onderzoek, waar sprake is van een reeks van herhaalde waarnemingen genest binnen individuele respondenten (vgl. Goldstein, 1986, 1989; Bryk & Raudenbush, 1987), en vignetten-onderzoek (Hox et al. 1991b), waar de waarnemingen immers ook genest zijn binnen de respondenten. Een ander voorbeeld is gegevens van survey interviews, waarbij respondenten genest zijn binnen interviewers, en interviewers binnen survey organisaties (cf. Hox, 1994).

Na het verschijnen van een aantal basis-publicaties (De Leeuw & Kreft, 1986; Raudenbush & Bryk, 1986; Goldstein, 1987; Longford, 1987), en vooral van efficiënte computerprogramma's (Bryk et al., 1988; Longford, 1990; Prosser et al., 1990) heeft er een ware explosie plaatsgevonden van toepassingen van multiniveau-analyse, die nog het meest doet denken aan de explosieve toename van het gebruik van structurele modellen na de publicaties van Jöreskog en het verschijnen van LISREL. Vrijwel alle toepassingen betreffen echter modellen en programma's die in essentie multiniveau versies zijn van het gebruikelijke multi-pele regressie model. Dit is niet zo'n grote beperking als het wellicht lijkt; multi-pele regressie analyse is een zeer flexibele analyse methode die op zeer uiteenlopende onderzoekproblemen toegepast kan worden (Cohen & Cohen, 1983). Desondanks is het aantrekkelijk om te beschikken over een multiniveau versie van een ruimere klasse van modellen, in het bijzonder multiniveau analyse van structurele modellen (i.e. pad analyse met latente variabelen).

Hieronder wordt het gebruik van structurele modellen beschreven voor gegevens met twee niveaus. Dat wil zeggen: gegevens die kunnen worden gezien als een tweetrapssteekproef uit twee verschillende hiërarchische waarnemingsniveaus, waarbij de variabelen op elk van deze niveaus gemeten kunnen zijn. In principe willen we voor dit soort gegevens een omvattend model kunnen specificeren, met de mogelijkheid om op elk van

beide niveaus een meetmodel en een padmodel te specificeren. De nadruk ligt op een model dat is voorgesteld door Muthén (1989, 1990, 1994), dat niet volledig algemeen is, maar het voordeel heeft dat schattingen van de populatie parameters berekend kunnen worden met behulp van conventionele software, zoals Lisrel, Liscomp of EQS.

Het decompositiemodel voor hiërarchische gegevens

Stel dat we beschikken over de gegevens van N leerlingen, verdeeld over G groepen (scholen, klassen). De leerling-gegevens zijn verzameld in een p -variate vector y_{ig} (subscript i voor individuen, $i=1..N$; subscript g voor groepen, $g=1..G$). Voor dit type gegevens hebben Cronbach en Webb (1975) voorgesteld de geobserveerde individuele scores y_{ig} op te splitsen in een between-groups component $y_B = y_g$ en een within-groups component $y_W = y_{ig} - y_g$. Met andere woorden, bij ieder individu wordt de geobserveerde totale score $y_T = y_{ig}$ vervangen door zijn samenstellende delen: de groeps-component y_B (het naar het individuele niveau gedissaggregeerde groeps-gemiddelde) en de individuele component y_W (de individuele deviatie van het bijbehorende groeps-gemiddelde). Deze beide componenten hebben de aantrekkelijke eigenschap dat zij additief en orthogonaal zijn (Cronbach en Webb, 1975):

$$y_T = y_B + y_W \quad (1).$$

Deze decompositie is de basis voor het berekenen van een between-groups covariantiematrix (de covariantiematrix van de gedissaggregeerde groeps-gemiddelden y_B) en een within-groups covariantiematrix (de covariantiematrix van de y_W). Net als de geobserveerde scores zelf zijn deze covariantiematrixen additief en orthogonaal:

$$S_T = S_B + S_W \quad (2).$$

Multiniveau analyse van structurele modellen veronderstelt dat er sprake is van een populatie van individuen, verdeeld in groepen. Wanneer de individuele scores in de populatie worden opgesplitst volgens Cronbach en Webb's procedure, dan geldt voor de covariantiematrixen in de populatie:

$$\Sigma_T = \Sigma_B + \Sigma_W \quad (3).$$

De twee populatie covariantiematrixen Σ_B en Σ_W leiden tot gecompliceerde vergelijkingen voor de likelihood functie van de geobserveerde gegevens. Het algemene model is uitgewerkt en beschreven door onder andere Goldstein & McDonald (1988), McDonald & Goldstein (1989), en Muthén & Satorra (1989). Muthén (1989, 1990) laat zien dat het model sterk vereenvoudigd kan worden, en zodat het in principe met conventionele software geschat kan worden.

Hieronder wordt het model uitgewerkt als een veralgemenisering van het vergelijkbare schattingsprobleem in de variantieanalyse (voor de algemene uitwerking in termen van de likelihood zie de hierboven aangehaalde literatuur).

Wanneer we uit een enkele univariaat-normaal verdeelde populatie met gemiddelde μ en variantie σ^2 G onafhankelijke groepen trekken, elk ter grootte van n ($\sum n = N$), dan is de verwachte waarde voor de within-groups varianties $\sum s_g^2$ ($g=1..G$) gelijk aan de populatievariantie σ^2 . Hieruit is af te leiden dat de gemiddelde variantie binnen de groepen, algemeen aangeduid als de *pooled within-groups* variantie s_{pw}^2 , een zuivere schatter is van de populatievariantie σ^2 .

Tevens geldt dat de groepsgemiddelden \bar{y}_g een normale verdeling volgen met gemiddelde μ en variantie σ^2/n (Hays, 1972). De variantie van de groepsgemiddelden \bar{y}_g in de steekproef, algemeen aangeduid als de *between-groups* variantie s_B^2 , is dus een schatter van σ^2/n , en hieruit is af te leiden dat ns_B^2 eveneens een zuivere schatter is van de populatievariantie σ^2 .

Wanneer de populatie bestaat uit verschillende groepen met een gemeenschappelijke (gelijke) variantie maar ongelijke gemiddelden, dan kan in de populatie onderscheid worden gemaakt tussen een between-groups variantie σ_B^2 en een within-groups variantie σ_w^2 (gebaseerd op de variantie van de groepsgemiddelden respectievelijk de variantie van de individuele deviaties van de groepsgemiddelden). In dit geval is de pooled within-groups variantie s_{pw}^2 nog steeds een zuivere schatter van de gemeenschappelijke within-groups variantie σ_w^2 , maar ns_B^2 is nu een schatter van $(\sigma_w^2 + n\sigma_B^2)$. In variantieanalyse wordt hiervan gebruik gemaakt door de nulhypothese dat alle groepsgemiddelden in de populatie gelijk zijn, met andere woorden dat $\sigma_B^2=0$, te toetsen met de F-ratio: $F=(ns_B^2)/(s_{pw}^2)$.

In het multivariate geval is er sprake van een vector y van multivariaat-normaal verdeelde variabelen, met de gemeenschappelijke within-groups covariantiematrix Σ_w , en de between-groups covariantiematrix Σ_B . Volstrekt analoog aan het univariate geval kan in de steekproef de *pooled within-groups covariantiematrix* S_{pw} bepaald worden:

$$S_{pw} = \frac{\sum_{g=1}^G \sum_{i=1}^n (Y_{ig} - \bar{Y}_g)(Y_{ig} - \bar{Y}_g)'}{N - G} \quad (4).$$

Tevens kan in de steekproef bepaald worden de *reguliere between-groups covariantiematrix* S_B (de covariantiematrix van de gedisaggregeerde groepsgemiddelden), gedefinieerd als n maal de covariantiematrix van de groepsgemiddelden $S_{\bar{y}}$:

$$S_B = \frac{n \sum_{g=1}^G (\bar{Y} - \bar{Y}_g)(\bar{Y} - \bar{Y}_g)'}{G} \quad (5).$$

De pooled within-groups covariantiematrix S_{PW} is een zuivere schatter van de gemeenschappelijke within-groups covariantiematrix Σ_w , terwijl de reguliere between-groups covariantiematrix S_B een zuivere schatter is van $(\Sigma_w + n\Sigma_B)$. In de multivariate variantieanalyse wordt hiervan gebruik gemaakt bij de toetsing van de nulhypothese dat de gemiddelden van de groepen niet verschillend zijn; de toetsingsgrootte hiervoor is Wilk's lambda: $\Lambda = |S_{PW}| / (|S_{PW}| + |S_B|)$.³

Muthén (1989, 1990) constateert dat

$$S_{PW} = \widehat{\Sigma_w} \quad (6)$$

en

$$S_B = \widehat{\Sigma_w + n\Sigma_B}, \quad (7)$$

en stelt voor de multigroep optie van conventionele programma's voor structurele vergelijkings-modellen te gebruiken, met als 'groepen' S_{PW} en S_B (gebaseerd op N-G respectievelijk G waarnemingen). Het model voor Σ_w wordt geschat op basis van zowel S_{PW} als S_B , met gelijkheidsrestricties tussen de beide 'groepen', en het model voor Σ_B wordt alleen geschat op basis van S_B , met een schaafactor n . De (latente) variabelen en residuen in het 'within' model zijn daarbij ongecorrleerd met die in het 'between' model.

In deze afleidingen is steeds uitgegaan van groepen met een gelijke groepsgrootte n . Wanneer de groepsgrootte n_g verschilt over de groepen, dan geldt dat

$$S_{PW} = \widehat{\Sigma_w}, \quad (8)$$

met andere woorden, S_{PW} is nog steeds een zuivere schatter voor Σ_w . Voor S_B geldt echter voor iedere verzameling d van groepen met gelijke groepsgrootte n_g dat

$$S_{Bd} = \widehat{\Sigma_w + n_{dg}\Sigma_B}, \quad (9)$$

³De determinant $|X|$ is de gegeneraliseerde variantie van de grootte X . De toetsingsgrootte Λ is een inverse maat; hoe groter de variantie van de groepsgemiddelden, des te kleiner is Λ . De overeenkomst met de univariate F-ratio is duidelijker wanneer we beschouwen dat: $1/\Lambda = |I - S_{PW}^{-1}S_B|$ (Tatsuoka, 1971, p164).

met andere woorden, voor iedere verzameling groepen met gelijke groeps grootte dient een verschillende schaal factor n_{dg} gebruikt te worden. In principe kan hiervoor wederom de multigroep optie van conventionele programma's gebruikt worden, met als 'groepen' S_{pw} en voor iedere onderscheidbare groeps grootte een aparte 'groep' S_{bd} . Het model voor Σ_w wordt geschat op basis van zowel S_{pw} als de S_{bd} , met gelijkheids restricties tussen alle 'groepen', en het model voor Σ_b wordt geschat op basis van de S_{bd} , met gelijkheids restricties tussen alle 'groepen' en voor elke 'groep' een andere schaal factor n_{dg} gebaseerd op de betreffende groeps grootte. Op deze wijze wordt alle informatie correct gemodelleerd, en is er sprake van een Full Information Maximum Likelihood oplossing (FIML, cf. Muthén, 1989).⁴

Bij een factoranalyse zal doorgaans sprake zijn van variabelen die op het individuele niveau gemeten zijn, waarbij het doel is de factorstructuur op beide niveaus te onderzoeken en te vergelijken. Bij een pad-analyse zal daarentegen dikwijls ook sprake zijn van groepsvariabelen, die alléén op het groepsniveau gemeten zijn. De multigroep optie van een programma als Lisrel vereist echter dat de matrixen voor beide groepen hetzelfde aantal variabelen bevatten. Het eenvoudig disaggregeren van de variabelen die op groepsniveau gemeten zijn levert geen oplossing. De gedisaggregeerde variabelen zijn immers voor alle individuen in een groep gelijk, zodat de within-groups covarianties van de gedisaggregeerde variabelen met de overige individuele deviatiescores niet gedefinieerd zijn. De oplossing is daarom de variabelen die alleen op groepsniveau gemeten zijn op het individuele niveau te behandelen als systematisch ontbrekende waarden (Allison, 1987; Bollen, 1989; de Lisrel-implementatie wordt uitvoerig beschreven in Jöreskog & Sörbom, 1989).⁵

Een praktische vereenvoudiging van het hiërarchische model

Het gebruik van conventionele software voor het volledige FIML model brengt twee belangrijke bezwaren met zich mee. Het eerste bezwaar is de onoverzichtelijkheid van het modelleren; er is sprake van een groot aantal 'groepen' met een diversiteit aan gelijkheids restricties, hetgeen leidt tot een groot en complex structurele-vergelijkingen model.⁶ Het tweede bezwaar is belangrijker; in veel gevallen zal er tenminste één onderscheidbare groeps grootte zijn waar slechts een of enkele groepen in vallen, zodat voor die groeps grootte S_{bd} niet berekend kan worden dan wel singulier is. Conventionele software kan dergelijke matrixen doorgaans niet analyseren.

⁴Voor een FIML oplossing dienen bovendien voor de between 'groepen' de factorgemiddelden geschat te worden met factorgemiddelden geschat te worden, met gelijkheids restricties tussen alle groepen (Muthén, 1989).

⁵Het programma EQS staat groepsvergelijkingen toe met een verschillend aantal variabelen in de groepen, waardoor in dit geval een aantal ingewikkeldheden vermeden wordt.

⁶Een bijkomend probleem is dat veel modellen door een programma als Lisrel als inadmissible worden gekenmerkt, hetgeen als gevolg heeft dat Lisrel niet zelf startwaarden kan berekenen.

Bij ongelijke groepsgroottes kan het model vereenvoudigd worden door uit te gaan van één enkele between groups covariantiematrix, berekend over alle groepen, gewogen naar groeps grootte. De formule voor de reguliere between-groups covariantiematrix S_B , gedefinieerd als n_g maal de covariantiematrix van de groeps gemiddelden $S_{\bar{Y}}$ wordt dan:

$$S_B = \frac{\sum_{g=1}^G n_g (\bar{Y} - \bar{Y}_g) (\bar{Y} - \bar{Y}_g)'}{G} \quad (10).^7$$

De schaalparameter wordt in dit geval gegeven door de grootte c , waarbij

$$C = \frac{N^2 - \sum_{g=1}^G n_g^2}{N(G-1)} \quad (11).$$

Bij gelijke groepsgroottes is c gelijk aan de gemeenschappelijke groeps grootte n , bij ongelijke groeps groottes is C bij benadering gelijk aan de gemiddelde groeps grootte. Overeenkomstig de procedure bij gelijke groeps groottes wordt het model voor Σ_w geschat op basis van zowel S_{Pw} als S_B , met gelijkheidsrestricties tussen de beide 'groepen', en wordt het model voor Σ_B geschat op basis van één enkele between groups matrix S_B , met een schaalfactor c gegeven door formule 11. Muthén (1990) laat zien dat S_B zoals berekend in formule 10 een consistente schatter is van Σ_B . Aangezien S_B geen maximum likelihood schatter is, levert de analyse nu slechts een benaderende oplossing. Wanneer de groeps groottes niet al te zeer van elkaar verschillen, kan de quasi likelihood procedure chikwadraat waarden en standaardfouten opleveren die een goede benadering zijn van de FIML schattingen.

Muthén (1990, 1994) heeft de resultaten van de quasi likelihood procedure bij een aantal datasets en modellen vergeleken met de FIML resultaten. Zijn conclusie is dat de verschillen minimaal zijn. McDonald (1994) heeft dezelfde vergelijking uitgevoerd op een synthetische dataset; ook hij concludeert dat de resultaten van de quasi likelihood procedure vrijwel identiek zijn aan die van de FIML procedure. Hox (1993) onderzoekt bij een synthetische dataset de invloed van de grootte van de intraclass correlatie en van de steekproefgrootte op beide niveaus op de nauwkeurigheid van de quasi likelihood oplossing; hij concludeert dat bij een steekproefgrootte van 50 op het hogere niveau de resultaten acceptabel zijn, en dat bij

⁷Dit is een eenvoudige formule voor de covariantiematrix van de gedissaggregeerde groeps gemiddelden.

een steekproefgrootte van 100 de resultaten zeer nauwkeurig zijn. In al deze gevallen lopen de groepsgroottes sterk uiteen.

Voorbeeld van een multiniveau factoranalyse

De gegevens van dit voorbeeld zijn afkomstig uit een onderzoek van Van Peet (1992). Het betreft de scores van 187 kinderen, afkomstig uit 37 grote gezinnen (gemiddeld vijf kinderen per gezin), op zes subtests van de Groninger Intelligentie Test (de GIT, zie Snijders & Verhage, 1962). De subtests zijn: woordlijst, kaartleggen, matrixen, verborgen figuren, dieren benoemen, en beroepen benoemen. De gegevens hebben een hiërarchische structuur, met kinderen genest binnen gezinnen. We nemen aan dat intelligentie sterk beïnvloed wordt door de gemeenschappelijke erfelijke en omgevingsfactoren die binnen de gezinnen door de kinderen gedeeld worden, en we verwachten dus tamelijk sterke effecten op het gezinsniveau (between groups).

De scores op de zes subtests zijn allereerst uitgesplitst in gezinsgemiddelden en individuele afwijkingsscores, volgens Cronbach en Webb (1975, vgl. formule 1 en 2). Tabel 1 hieronder toont de gemiddelden en de varianties van de totaalscores, de varianties van de uitgesplitste scores, en de Intra Class Correlatie (ICC), welke een schatting is van de proportie variantie die op gezinsniveau ligt.

Tabel 1. Gemiddelden, varianties en ICC voor gezinsdata.

Subtest	Total		Family		Indiv.
	Mean	Var.	Var.	Var.	ICC
woordlijst	29.80	15.21	7.48	7.73	.37
kaarten	32.68	28.47	13.65	14.82	.35
matrixen	31.73	16.38	5.24	11.14	.15
figuren	27.11	21.23	6.84	14.38	.16
dieren	28.65	22.82	8.46	14.36	.22
beroepen	28.28	21.42	9.11	12.31	.28

De resultaten in Tabel 1 geven aan dat een belangrijk deel van de variantie op het gezinsniveau ligt. Om de factorstructuur van de zes subtests te onderzoeken is een binnen gezinnen covariantiematrix S_{pw} en een tussen gezinnen covariantiematrix S_b berekend volgens formules 4 en 10.

De eerste stap is het specificeren van een model op het individuele niveau, op basis van een analyse van de binnen gezinnen covariantiematrix S_{pw} . Een exploratieve factoranalyse suggereert een twee factor model, met de eerste drie subtests op de eerste factor, en de andere drie subtests op de tweede factor. Een confirmatieve factoranalyse levert voor dit model een bevredigende passing op: het twee-factor model is acceptabel (chi-kwadraat=7.21, df=8, $p=.51$), terwijl het eenvoudiger model met slechts één (algemene) factor wordt verworpen (chi-kwadraat=44.87, df=9, $p=.00$).

De volgende stap is het specificeren van een model voor de tussen gezinnen covariantiematrix S_B . Daartoe worden de beide matrixen S_{pw} en S_B simultaan geanalyseerd volgens de multi-groep procedure, waarbij in beide groepen hetzelfde twee-factor model voor de Σ_w wordt gespecificeerd, met gelijkheidsrestricties voor alle overeenkomstige parameters. Wat betreft het model voor de between groups matrix Σ_B is het zinvol om twee grensgevallen te analyseren. Het ene grensgeval is het minimale model, dat op gezinsniveau géén model specificeert voor Σ_B ; hierbij wordt er in feite van uitgegaan dat er in de populatie geen gezinseffecten zijn. Het andere grensgeval is het maximale model, dat op gezinsniveau een volledige covariantiematrix specificeert; hierbij wordt op gezinsniveau dus geen enkele restrictie opgelegd (voor het gezinsniveau geldt immers een verzadigd model). Interessante modellen op gezinsniveau zullen zich, wat betreft aantal vrijheidsgraden en passing, ergens tussen het minimale en het maximale model bevinden. De passing van het minimale en het maximale model kan gebruikt worden om voor het gezinsmodel goodness-of-fit maten te berekenen zoals bijvoorbeeld Bentler en Bonet's delta (Bentler & Bonet, 1980; Bollen, 1989).⁸

In eerste instantie zijn op gezinsniveau twee modellen onderzocht: een model met één algemene factor, en een model met dezelfde twee factoren als op individueel niveau gevonden zijn. Tabel 2 hieronder geeft de resultaten van de modeltoets voor deze twee modellen plus het minimale en het maximale model.

Tabel 2 Vergelijking van vier gezinsniveau modellen

Gezinsmodel	Chi-kwadraat	df	p
Minimaal model	125.41	29	.00
Eenfactor model	21.28	17	.21
Tweefactor model	20.06	16	.22
Maximaal model	7.21	8	.51

⁸Als gevolg van de verplichte gelijkheidsrestricties tussen de groepen zijn de goodness-of-fit indices die het programma uitrekt soms misleidend. Bovendien is het wenselijk om de goodness-of-fit apart te kunnen onderzoeken voor het individuele en het gezinsmodel.

Het minimale model wordt verworpen. Het maximale model, dat geen restricties plaats op de between groups structuur, en de within groups structuur schat door gebruik te maken van de informatie zowel in S_{PW} als in S_B , levert een goede passing op, met parameterschattingen die nauwelijks afwijken van de parameterschattingen die verkregen worden bij het analyseren van alleen S_{PW} .⁹ De passing van het twee-factor model is in dit geval niet beter dan die van het eenfactor model; op gezinsniveau is één algemene factor kennelijk voldoende.

Tabel 3 toont de gestandaardiseerde parameterschattingen voor het model (standaardisatie naar een gemeenschappelijke metriek voor beide groepen).

Tabel 3 Individueel en gezinsmodel, gestandaardiseerde factorladingen

	Individueel	Gezin
woordlijst	.30*	.84*
kaarten	.52	.78
matrixen	.70	1.02
figuren	.30	.58
dieren	.70	.86
beroepen	.48*	.33 ^{ns}

Correlatie tussen beide individuele factoren: 0.22^{ns}

* = ongestandaardiseerde parameter gefixeerd op 1; ns = niet significant

Tabel 3 suggereert dat op gezinsniveau, waar de effecten van de door alle kinderen gedeelde erfelijkheids- en omgevingskenmerken zichtbaar zijn, één algemene (g) factor voldoende is om de covarianties te verklaren. Op individueel niveau, waar de effecten van ideosyncratische kenmerken zichtbaar zijn, zijn twee factoren nodig. De eerste factor kan geïnterpreteerd worden als een 'reasoning' factor, en de tweede als 'fluency'. Deze resultaten sporen goed met bijvoorbeeld de intelligentietheorie van Cattell (1971), die stelt dat als gevolg van individuele ervaringen (opvoeding, scholing) de algemene g-factor 'uitkristalliseert' in de vorm van specifieke individuele vaardigheden.

⁹In principe kan het maximale model eveneens gebruikt worden om het individueel model te exploreren. Omdat S_{PW} doorgaans op aanzienlijk meer waarnemingen is gebaseerd dan S_B gaat er weinig informatie verloren wanneer alleen S_{PW} geanalyseerd wordt, terwijl de setups aanzienlijk eenvoudiger zijn en de rekenprocedure sneller convergeert.

Voorbeeld van een multiniveau padanalyse

De gegevens uit dit voorbeeld zijn afkomstig uit een onderzoek van Schijf en Dronkers (1991). Het betreft hier gegevens van in totaal 1379 lagere school-leerlingen verdeeld over 58 lagere scholen.¹⁰ Van de leerlingen zijn op individueel niveau de volgende variabelen bekend: beroepsstatus vader (FOCC), opleiding vader (FEDUC), opleiding moeder (MEDUC), gezinsgrootte (FSIZE), sexe (SEX), aantal malen dat de leerling in het verleden was blijven zitten op school (REPEAT), de score op de GALO schoolprestatietest (GALO), en het advies van de leerkracht voor het type middelbare school (ADVICE). Op schoolniveau is er één variabele: de denominatie van de school (DENOM). Deze is hier gecodeerd als 1=protestant, 2=openbaar, 3=katholiek (scoring gebaseerd op optimale schaling). De vraagstelling in dit onderzoek was of de denominatie van de school invloed heeft op de prestatiescore en het advies van de onderwijzer, rekening houdend met de sexe en het sociaal milieu van de leerling, en de milieusamenstelling van de school.

In de analyses van Schijf en Dronkers werden daartoe multiniveau regressiemodellen geanalyseerd, met als predictoren op het individuele (leerling-)niveau de individuele leerlingvariabelen, en als predictoren op het schoolniveau de geaggregeerde individuele predictoren als indicator van de milieusamenstelling van de school als geheel, aangevuld met de denominatie van de school. Zij voerden twee aparte multiniveau regressie analyses uit: een voor de afhankelijke variabele 'GALO prestatiescore' en een voor 'advies leerkracht'.

Het voordeel van een padmodel boven een serie aparte regressieanalyses is dat het in een padmodel mogelijk is om een beschrijving te geven van de onderlinge invloeden in een verzameling van onafhankelijke, intervenierende, en afhankelijke variabelen. De vraagstelling van Schijf en Dronkers (zie ook Dronkers en Schijf, 1985) heeft echter een duidelijk multiniveau karakter, en vereist daarom ook een multiniveau padanalyse, waarin de geaggregeerde variabelen en de pure schoolvariabele 'denominatie' op statistisch correcte wijze wordt opgenomen.

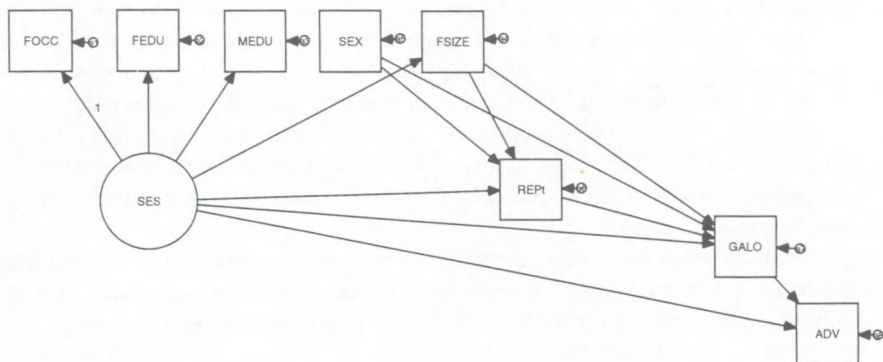
Zoals hierboven uiteengezet is, treedt bij een variabele als 'denominatie', die uitsluitend op schoolniveau bepaald is, het probleem op dat voor deze variabele geen tegenhanger heeft op het individuele niveau. Dit wordt opgelost door 'denominatie' op het individuele niveau te behandelen als een variabele met systematische ontbrekende waarden (Allisin, 1987; Bollen, 1989). Dit komt erop neer dat de schoolvariabele 'denominatie' in S_b op de normale wijze wordt opgenomen en in het between-model wordt gemodelleerd. In de matrix S_{pw} worden de (co)varianties van 'denominatie' vervangen door pseudo-waarden (vgl. Jöreskog, 1971): de variantie wordt gesteld op 1 en de covarianties worden gesteld op nul, en in het within-model wordt voor deze variabele alleen de residuele varianties geschat. Omdat het programma de pseudo-waarden als 'gewone' waarnemingen beschouwt, wordt het aantal vrijheidsgraden niet correct berekend. Dit moet handmatig worden aangepast (Jöreskog & Sörbom, 1989).

¹⁰De gegevens zijn verzameld in 1971. De hier geanalyseerde dataset bestaat uit alle leerlingen en scholen met volledige gegevens op de gebruikte variabelen. In het oorspronkelijke onderzoek werd ook nog een wijkniveau onderscheiden. Dit verklaarde echter geen variantie, en is daarom in deze analyse niet opgenomen.

De eerste stap in de analyse is weer het bepalen van een goed passend model voor de pooled within matrix S_{PW} . Hiervoor is een individueel schoolloopbaanmodel opgesteld, ontleend aan eerdere padanalyses zonder multiniveauspecificatie (Dronkers & Schijf, 1986; Schijf en Dronkers, 1990, 1991). Figuur 1 op de volgende pagina bevat een grafische weergave van het individuele model.

Figuur 1 toont een padmodel met één latente variabele voor Sociaal-Economische Status (SES). De chi-kwadraat voor dit model is 32.8 (df=15, p=.01), en alle modificatie-indexen zijn klein. Gegeven de steekproefgrootte van 1321 onafhankelijke (within schools) observaties wordt het model daarom acceptabel geacht. De goodness-of-fit indices zijn: GFI=.99 en AGFI=.99. De grafische weergave in Figuur 2 toont de gestandaardiseerde parameterschattingen.

Figuur 1: Padmodel voor 'Pooled Within' GALO data.



Inspectie van de geschatte intraclass correlaties levert op dat de geschatte variantie op school-niveau van de variabele 'sexe' gelijk is aan nul; dat wil zeggen dat het percentage meisjes in alle 58 (lagere) scholen nagenoeg hetzelfde is. Op schoolniveau is deze variabele daarom uit het model verwijderd, door hem te behandelen als ontbrekende waarneming op dezelfde manier als de variabele 'denominatie' op het individuele niveau behandeld wordt.

De volgende stap is het specificeren van een model voor de tussen scholen covariantiematrix S_B . Daartoe worden de beide matrixen S_{pw} en S_B simultaan geanalyseerd volgens de multi-groep procedure, waarbij in beide groepen het hierboven model wordt gespecificeerd voor de Σ_w , met gelijkheidsrestricties voor alle overeenkomstige parameters. Voor de between groups matrix Σ_B worden in eerste instantie drie vergelijkingsmodellen gespecificeerd. Twee vergelijkingsmodellen zijn de grensgevallen: het minimale model, dat op schoolniveau géén model specificeert voor Σ_B , en het maximale model, dat op schoolniveau de volledige covariantiematrix specificeert en daarmee geen enkele restrictie oplegt.

Interessante modellen op schoolniveau zullen zich, wat betreft aantal vrijheidsgraden en passing, ergens tussen het minimale en het maximale model bevinden.¹¹ Naast het minimale en het maximale model is een derde vergelijkingsmodel gespecificeerd, dat alleen de between schools varianties verklaard. Dit model postuleert dat Σ_B een diagonale matrix is, met andere woorden: op schoolniveau zijn alleen varianties aanwezig, maar geen covariantiestructuur. Wanneer het minimale model acceptabel is, kan van het opstellen van een schoolmodel worden afgezien; een analyse van de pooled within matrix in plaats van de ruwe scores voor het individuele model is echter wel aan te raden omdat daarin wordt gecorrigeerd voor de variantiecomponenten op schoolniveau als gevolg van het trekken van een cluster steekproef. (Voor een vergelijking van het multiniveau perspectief met dat van meertraps steekproeftrekking zie Muthén en Satorra, 1991, en Muthén, 1994). Analyse van de drie vergelijkingsmodellen levert de volgende toetsingsresultaten op:

Tabel 4 Vergelijkingsmodellen schoolniveau

Model	Chi-kwadraat	df	p
Minimaal	623	51	.00
Alleen varianties	392	44	.00
Maximaal model	39	23	.02

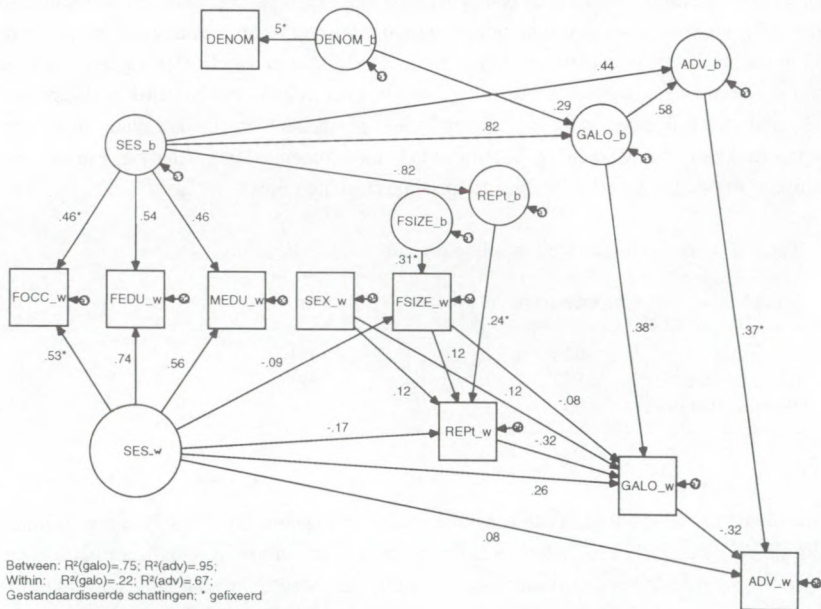
Het maximale model produceert een schatting van de covariantiematrix Σ_B . In eerste instantie is voor de geaggregeerde variabelen op het schoolniveau een model gespecificeerd gelijk aan het model op het individuele niveau (met weglating van sexe). Op schoolniveau blijken de variabelen beroepsstatus vader (FOCC), opleiding vader (FEDUC) en opleiding moeder (MEDUC) zeer hoog te correleren (met geschatte correlaties groter dan 0.98). Op het individuele niveau indiceren deze variabelen een latente variabele SES. Op schoolniveau leidt

¹¹De passing van het minimale en het maximale model kan gebruikt worden om voor het schoolmodel apart goodness-of-fit maten te berekenen. Omdat het within model in dit voorbeeld op zich statistisch significant is, zal de chi-kwadraattoets van het multiniveau model vermoedelijk ook een significante toetsingsgrootheid zal opleveren. Analyse van de passing door goodness-of-fit maten is dan een zinvol alternatief voor de statistische toets.

een factor-model voor deze drie variabelen tot niet convergerende modellen, en wordt daarom voor deze variabelen een component-model gehanteerd waarin deze drie variabelen gecombineerd zijn tot een indicator. Op schoolniveau is daarmee sprake van een padmodel zonder latente variabelen. Verder blijken vele paden op schoolniveau niet significant te zijn; deze zijn daarom uit het schoolmodel verwijderd. Het model dat aldus ontstaat heeft een goede passing ($\chi^2=58$, $df=40$, $p=.03$). Het model past daarbij niet significant slechter dan het maximale model, en er zijn geen grote modificatieindexen voor de gefixeerde parameters.

In een laatste stap is de schoolvariabele 'denominatie' toegevoegd. Deze blijkt uitsluitend een significant effect te hebben op de GALO-score (het betreffende model heeft een $\chi^2=64$, $df=47$, $p=.05$). Het uiteindelijk model, met de gestandaardiseerde parameterschattingen, staat hieronder in figuur 2:

Figuur 2: Uiteindelijk padmodel voor GALO data.



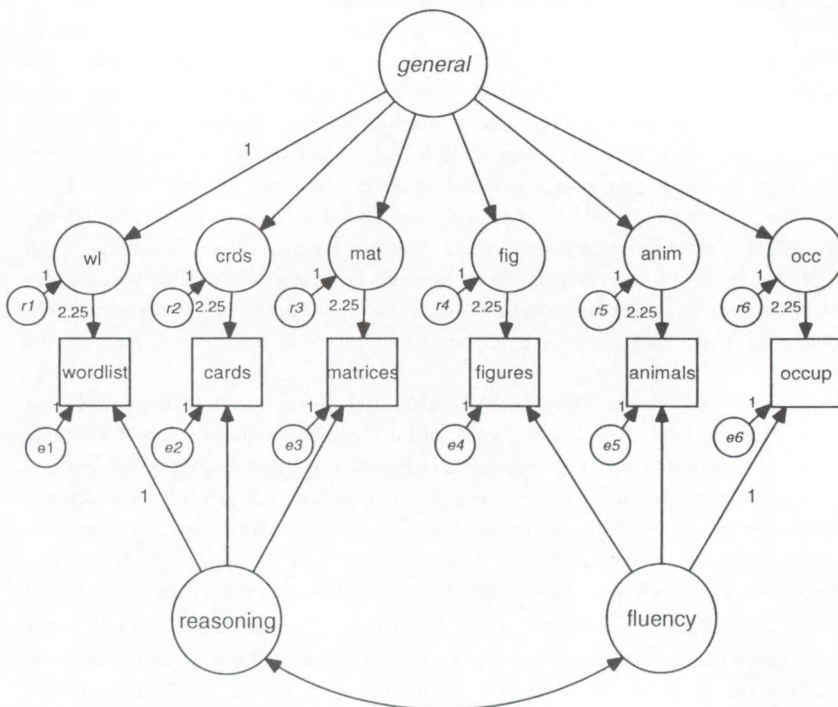
Een interessant verschil met de multiniveau regressie analyses die Schijf en Dronkers (1991) uitvoeren voor GALO-score en Advies leerkracht apart, is dat bij de afzonderlijke analyses de denominatie in beide gevallen een significante bijdrage levert, terwijl uit de multiniveau padanalyse blijkt dat het effect op het advies geheel indirect is en loopt via de GALO-score. Het voordeel van een padmodel boven een serie regressieanalyses is dat uit afzonderlijke regressie analyses een dergelijke structuur niet naar voren komt.

Figuur 2 laat zien dat SES niet alleen op individueel niveau, maar ook op schoolniveau een effect heeft op de variabelen als 'zittenblijven', GALO score, en advies leerkracht. De interpretatie van dit soort effecten in het schoolmodel is in termen van een contextueel effect van de compositie van de scholen. De interpretatie is *niet* dat we voor scholen met veel lage-SES leerlingen daardoor 'uiteraard' relatief slechte schoolprestaties mogen verwachten. Het effect van ongelijke samenstelling van de scholen wordt gemodelleerd in het individuele model dat eveneens voor de between covariantiematrix S_B wordt geschat. De effecten in het schoolmodel laten zien dat in scholen met veel lage-SES leerlingen, in aanvulling op de slechte schoolprestaties die te verwachten zijn op basis van de individueel lagere verwachtingen, de *concentratie* aan lage-SES leerlingen een extra drukkend effect heeft op de schoolcarriere.

Implementatie van de modellen in LISREL

Voor het factormodel voor de Van Peet data ziet de grafische weergave er als volgt uit:

Figuur 3: Factormodel voor Van Peet Data



De onderste helft van Figuur 3 is het model voor Σ_w , dat zowel voor S_{pw} als voor S_B gespecificeerd wordt, met gelijkheidsrestricties voor alle overeenkomstige parameters. Dit is een 'normale' Lisrel specificatie voor een factormodel: de zes geobserveerde variabelen (de zes GIT-subtests) zijn Y-variabelen, de latente factoren 'verbal' en 'fluency' zijn Eta-variabelen, met voor iedere latente factor één padcoëfficiënt op 1 gefixeerd om de schaal van de latente factor te identificeren. In Lisrel zijn de padcoëfficiënten voor de factorladingen te vinden in de matrix Lambda-X. De residuele varianties van de unieke factoren e1 tot en met e6 worden weergegeven in de diagonale matrix Theta-Delta.

De bovenste helft van de figuur is het model voor Σ_B , dat alleen voor S_B gespecificeerd wordt. In dit model wordt voor iedere geobserveerde Y-variabele een bijbehorende latente variabele gespecificeerd met de padcoëfficiënt gefixeerd op $\sqrt{c}=2.25$. Dit verdisconteert de schaafactor c (zie vergelijking 11). Bij elk van deze latente variabelen hoort een residuele variabele r1 tot en met r6. Figuur 3 laat duidelijk zien dat ten behoeve van het model op het hogere niveau bij elke \bar{Y}_g -variabele een latente variabele wordt gedefinieerd die de tussenvariantie schat. Op het groepsniveau worden deze latente variabelen vervolgens behandeld alsof het 'geobserveerde groeps-variabelen' waren, waarvoor vervolgens een factormodel wordt opgesteld. De gefixeerde padcoëfficiënten op de \bar{Y}_g -variabelen zijn daarom inhoudelijk niet interessant; de gefixeerde waarde dient alleen voor de correcte schaling (belangrijk wanneer men gelijkheidsrestricties tussen het within- en betweenmodel zou willen opstellen).

In Figuur 3 wordt voor de zes 'gezinsvariabelen' een factormodel gespecificeerd, met één enkele latente factor 'general', die ladingen heeft op alle zes 'gezins-variabelen', waarvan de eerste op 1 gefixeerd is om de schaal van de 'general' factor te identificeren. In de grafische weergave ziet dit er nog betrekkelijk 'normaal' uit. Bij nadere beschouwing blijkt echter dat deze factorladingen niet zoals gebruikelijk ladingen zijn van een latente Eta-variabele op geobserveerde variabelen, maar 'ladingen' van een latente Eta-variabelen op een aantal andere. In Lisrel zijn de padcoëfficiënten van de Eta-variabelen onderling de Beta-coëfficiënten, die te vinden zijn in de matrix Beta. De factorladingen in het gezinsmodel moeten dus als padcoëfficiënten gespecificeerd worden in de matrix Beta. De varianties van r1 tot en met r6, de residuele factoren op gezinsniveau, zijn te vinden in de matrix Psi. De covarianties tussen de latente variabelen Eta worden alle op nul gesteld, met uitzondering van de covariantie tussen Eta1 en Eta2, die de covariantie tussen de beide individuele factoren weergeeft.

In matrix-georiënteerde programma's zoals Lisrel (Jöreskog en Sörbom, 1989) en Liscomp (Muthén, 1987) leidt het bovenstaande tot een zeer complexe setup. Voor beide 'groepen' moet hetzelfde aantal geobserveerde en latente variabelen gespecificeerd worden. In de within groups specificatie leidt dit tot een ladingenmatrix LX met 9 latente factoren, waarin voor latente variabele drie tot en met negen alleen maar nullen staan, een matrix PSI die grotendeels leeg is, en een matrix Beta die geheel leeg is. In de between groups specificatie wordt voor elk van de zes geobserveerde variabelen voor de bijbehorende latente variabelen één coëfficiënt in LX gefixeerd op de wortel uit de schaafactor: \sqrt{c} . In de matrix Beta worden nu padcoëfficiënten opgenomen voor de paden van de 'general' factor op de zes andere factoren, en de diagonaal van PSI wordt nu gevuld. Al met al leidt dit voor een

eenvoudige confirmatieve factoranalyse tot een setup met een groot aantal matrixen, voor het grootste deel gevuld met structurele nullen. Opgemerkt dient nog te worden dat Lisrel7 dit model definieert als 'inadmissible' en bovendien geen startwaarden kan berekenen (Lisrel8 kan dit laatste wel). De analist moet dus in Lisrel de zogenaamde 'admissibility check' afzetten, en in sommige gevallen zelf startwaarden invoeren. Een goede reden om de analyses voor het within model uitsluitend op de covariantiematrix S_{PW} uit te voeren, waarbij deze problemen zich niet voordoen.

Bij software die de modelspecificatie verwacht in de vorm van stelsels van vergelijkingen, zoals de Simplis optie in Lisrel8 (Jöreskog & Sörbom, 1993), Eqs (Bentler, 1985), en Amos (Arbuckle, 1992) is de modelspecificatie eenvoudiger. Voor iedere variabele, zowel de geobserveerde als de latente, dient nu te worden opgegeven van welke andere variabelen deze afhangt. Uitgangspunt hierbij is de grafische weergave van het gewenste model. Belangrijk hierbij is dat een multiniveau factormodel zoals weergegeven in Figuur 3 een groot aantal restricties impliceert die meestal expliciet opgegeven moeten worden, zoals de gelijkheidsrestricties en de restricties die voortvloeien uit de aanname dat (de modellen voor) Σ_w en Σ_b geheel ongecorrleerd zijn. Een andere belangrijke restrictie is dat de latente variabelen op beide niveaus ongecorrleerd moeten zijn. Veel vergelijking-georiënteerde software neemt bij default aan dat de correlaties tussen exogene variabelen alle geschat moeten worden, en voor de correlaties in Figuur 3 tussen de gezinsfactor 'general' en de individuele factoren 'reasoning' en 'fluency' moet dit nu juist niet. Het probleem van de complexe matrixinvoer bij de matrix-georiënteerde software wordt bij de vergelijking-georiënteerde software door het probleem dat alle noodzakelijke restricties expliciet gemaakt moeten worden.

Voor de multiniveau padanalyse weergegeven in Figuur 2 is de situatie nog ingewikkelder. De matrix Beta bevat nu voor het within model een aantal padcoëfficiënten die causale relaties tussen de individuele variabelen specificeren, en voor het between model de padcoëfficiënten die de plaats innemen van de factorladingen plus de padcoëfficiënten die causale relaties tussen de schoolvariabelen specificeren. Daarnaast is er nog de variabele 'denominatie' waarvan de covarianties op het individuele niveau niet gedefinieerd zijn, en waarvoor pseudo-waarden zijn ingevuld. In het within-model wordt deze variabele als een unieke variabele gemodelleerd, terwijl deze in het between model 'gewoon' in het model wordt opgenomen. Door de opneming van pseudowaarden in de within-covariantiematrix dient het aantal vrijheidsgraden te worden gecorrigeerd.

Voor multiniveau factor- en padanalyses volgens het hierboven beschreven model is het handig om over een preprocessor te beschikken die op basis van de ruwe data de benodigde covariantiematrixen S_{PW} en S_b (formules 4 en 10) en de schaafactor c uitrekent. Muthén heeft hiervoor het programma *BW* geschreven (Muthén, 1989; Nelson & Muthén, 1991), dat op aanvraag bij Muthén verkrijgbaar is. Een PC-versie van *BW* is ook te verkrijgen bij de auteur, evenals het vergelijkbare programma *SPLIT2* (Hox, 1994), dat minder algemeen is, maar ook eenvoudiger in het gebruik. De handleiding voor *SPLIT2* bevat ook Lisrel-setups voor de hier gebruikte voorbeelden.

Discussie

De vereenvoudigde quasi likelihood procedure lijkt een algemeen bruikbare methode te zijn voor de analyse van hiërarchische structurele modellen. Zowel vergelijking met bekende populatiewaarden bij synthetische data als vergelijking met Full Information Maximum Likelihood oplossingen bij empirische data geven aan dat de benadering tamelijk nauwkeurig is. In de praktijk zal bij vrijwel alle toepassingen het probleem optreden dat het aantal waarnemingen beperkt is, zodat de between-groups matrix op een betrekkelijk klein aantal waarnemingen is gebaseerd. Boomsma (1983) suggereert een aanvaardbare ondergrens van 200 observaties bij de analyse van structurele modellen; bij hiërarchische data zal deze ondergrens wat betreft de between-groups matrix dikwijls niet gehaald worden. Anderzijds leveren de hierboven weergegeven analyses van het factoranalyse model bij 37 gezinnen en het padmodel bij 58 scholen op het hogere niveau resultaten op die een plausibele indruk wekken. Wellicht dat het bij de analyse op het hogere niveau een rol speelt dat deze wordt uitgevoerd op een covariantiematrix van groepsgemiddelden. Groepsgemiddelden zijn stabielier dan de onderliggende ruwe scores, en uit de literatuur is bekend dat bij aggregatie de correlaties doorgaans hoger worden, een verschijnsel dat bekend staat als het 'Robinson effect' (naar Robinson, 1950). Het 'Robinson-effect' is in bovengenoemde voorbeelden terug te vinden in de vorm van hogere ladingen voor het factormodel op het hogere (gezins-)niveau, en in de vorm van hogere ladingen voor de padcoëfficiënten op het hogere (school-) niveau. Het verlies aan onderscheidingsvermogen door aggregatie naar een hoger niveau wordt door de grotere stabiliteit van gemiddelden en door het Robinson effect tot op zekere hoogte goedge maakt; een bevinding waarop ook door Stelzl (1982) in de context van de variantie analyse gewezen is.

Vergelijking (7) laat zien dat het model voor de geobserveerde between-groups matrix een combinatie is van het within-model en het between-model. In feite modelleert het between-model de covarianties die niet verklaard worden door het gemeenschappelijke within-groups model. De kwaliteit van het between-model is dus voor een deel afhankelijk van de kwaliteit van het within-model. In de analyses van de twee empirische voorbeelden is daarom steeds begonnen met het bepalen van het within-model. Dit kan op twee verschillende manieren: men kan eenvoudig de covariantiematrix S_{pw} analyseren, of men kan een simultane analyse uitvoeren op de covariantiematrixen S_{pw} en S_B , met een verzadigd model voor het between-model. De eerste benadering is de eenvoudigste, en om die reden hier ook gebruikt. In de meeste gevallen zal het aantal waarnemingen waarop S_{pw} gebaseerd is ($N-G$) aanmerkelijk groter zijn dan het aantal waarnemingen waarop S_B gebaseerd is (G), zodat er weinig informatie verloren gaat.

Bij het bepalen van het between-model is het zinvol om gebruik te maken van specifieke vergelijkingsmodellen, waarmee de passing van een specifiek model vergeleken kan worden. Bentler en Bonett (1980) stellen voor om gebruik te maken van een maximaal model en een minimaal model, en de passing van een specifiek model te evalueren door het te vergelijken met deze beide modellen. Voor conventionele modellen stellen zij voor dit doel

een aantal goodness-of-fit indices voor (vgl. Bollen, 1989), waarbij bij conventie voor het minimale model het onafhankelijkheidsmodel wordt genomen, en voor het maximale model het verzadigde model. Bij multiniveau analyse moet bedacht worden dat er steeds sprake is van twee modellen, één voor Σ_w en één voor Σ_b , die elk voor zich een goede of slechte passing kunnen hebben. Een voordeel van de hierboven aangegeven strategie om te beginnen met een analyse van alleen de covariantiematrix S_{pw} is dat de diverse goodness-of-fit indices door de beschikbare programmatuur kunnen worden uitgerekend. Bij de bepaling van de passing van het between-model treden echter problemen op. De conventionele goodness-of-fit indices worden immers bepaald voor het model als geheel, en de passing van dat model wordt bepaald door zowel de passing van het within-model als van het between-model. Vrijwel alle goodness-of-fit indexen komen neer op het bepalen, in de vorm van een index die bij voorkeur tussen 0 en 1 ligt, waar de passing van een model gelokaliseerd moet worden vergeleken met het onafhankelijkheidsmodel en het verzadigde model. Bij het between-model wordt de plaats van het verzadigde model ingenomen door een model met het gemeenschappelijke within-model plus een verzadigd between-model. Het onafhankelijkheidsmodel komt overeen met het gemeenschappelijke within-model plus een between-model dat alleen de varianties van de between-variabelen schat. Bij multiniveau analyse is het mogelijk om voor Σ_b een nog minimaler model te schatten, namelijk het minimale model dat specificeert dat er géén between-variantie is. Voor onderlinge vergelijking van between-groups modellen lijkt het 'alleen-varianties' model echter het meest geschikt; het is geheel analoog aan de gebruikelijke procedure bij conventionele modellen, en bovendien kan de hypothese dat er in het geheel geen between variantie is eenvoudiger getoetst worden door voor de gebruikte variabelen met een eenvoudige variantieanalyse de between-groups variantie op significantie te toetsen.

Zoals eerder opgemerkt leidt het gebruik van conventionele software tot complexe en vrij onoverzichtelijke setups. Bovendien lijdt de Full Information Maximum Likelihood procedure bij de meeste software aan de handicap dat alle between-groups covariantiematrixen positief definitief moeten zijn, terwijl het achterliggende model in feite alleen stipuleert dat de populatie covariantiematrix Σ_b positief definitief moet zijn. Het laatste bezwaar kan worden ondervangen door software te gebruiken die deze eis niet stelt, of door bestaande software hiertoe aan te passen. De onhandige setups met grote en spaars gevulde matrixen kunnen alleen voorkomen worden door specialistische software, waarin de benodigde manipulaties intern in het programma worden uitgevoerd. Deze programmatuur is momenteel nog niet algemeen beschikbaar, hoewel het werk van McDonald (1994) in deze richting vergevorderd en veelbelovend is.

Referenties

- Allison, P.D. (1987). Estimation of linear models with incomplete data. In: C. Clogg (Ed.) *Sociological Methodology 1987*, pp 71-103. San Francisco: Jossey-Bass.
- Arbuckle, J. (1992). AMOS 3.1 Documentation package. Philadelphia: Department of Psychology, Temple University.
- Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bentler, P.M. (1985). Theory and implementation of EQS: A structural equations program. Los Angeles: BMDP Statistical Software.
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality*. Amsterdam: Sociometric Research Foundation.
- Bryk, A.S. & Raudenbush, S.W. (1987). Applying the hierarchical linear model to measurement of change problems. *Psychological Bulletin*, 99, 388-399.
- Bryk, A.S., Raudenbush, S.W., Seltzer, M. & Congdon, R.T. (1988). An introduction to HLM. Computer program and user's guide.
- Cattell, R.B. (1971). *Abilities: their structure, growth, and action*. Boston: Houghton Mifflin.
- Cohen, J. & Cohen, P. (1983). *Applied multiple regression analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cronbach, L.J. & Webb, N. (1975). Between-class and within-class effects in a reported aptitude x treatment interaction: Re-analysis of a study by G.L. Anderson. *Journal of Educational Psychology*, 67, 717-724.
- De Leeuw, J. & Kreft, I. G.G. (1986). Random coefficient models. *Journal of Educational Statistics*, 11, 1, 55-85.
- Dronkers, J. & Schijf, H. (1986). Buurten, scholen en ongelijke onderwijskansen. In A.B. Berends, J. Dronkers, J.K.M. Gevers, J. de Jong-Gierveld en F. Lammertijn (red.) *Sociologisch Jaarboek 1985*. Deventer: Van Loghum Slaterus.
- Gifi, A. (1990). *Nonlinear multivariate analysis*. New York: Wiley.
- Goldstein, H. (1986). Efficient statistical modelling of longitudinal data. *Annals of Human Biology*, 13, 129-141.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Griffin.
- Goldstein, H. (1989). Efficient statistical modelling of longitudinal data. In: D. Bock (ed.). *Multilevel analysis of educational data*. San Diego: Academic Press.
- Goldstein, H. & McDonald, R. (1988). A general model for the analysis of multilevel data. *Psychometrika*, 53, 455-467.
- Härnqvist, K. (1978). Primary mental abilities at collective and individual levels. *Journal of Educational Psychology*, 70, 706-716.
- Hox, J.J. (1992). Modeling interviewer effects with multilevel methods. *Kwantitatieve Methoden*, 13, 59-74.
- Hox, J.J. (1993). Factor analysis of multilevel data: gauging the Muthén model. In: H. Oud & R. Voeglesang (eds.) *Advances in multivariate and longitudinal analysis*. Nijmegen: ITS.
- Hox, J.J. (1994). Multilevel covariance structure analysis: background and Lisrel-implementation. Short guide to the PC versions of BW and SPLIT2. ML/S report # 54, Department of Education, University of Amsterdam.
- Hox, J.J. (1994). Hierarchical regression models for interviewer and respondent effects. *Sociological Methods & Research*, 22, 300-318.
- Hox, J.J. & Willemse, J. (1985). De empirische samenhang van sociometrische scores; een analyse op twee niveau's. *Tijdschrift voor Onderwijsresearch*, 10, 82-85.
- Hox, J.J., Kreft, I. G.G., & Hermkens, P.L.J. (1991a). The analysis of factorial surveys. *Sociological Methods & Research*, 19, 493- 510.
- Hox, J.J., De Leeuw, E.D. & Kreft, I. G.G. (1991b). The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In: P.P. Biemer, R.M. Groves,

- L.E. Lyberg, N.A. Mathiowetz, & S. Sudman (eds.) *Measurement Errors in Surveys*. New York: Wiley, 439-462.
- Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 67, 409-426.
- Jöreskog, K.G. & Sörbom, D. (1989). *LISREL 7. A guide to the program and applications*. Chicago: SPSS Inc.
- Jöreskog, K.G. & Sörbom, D. (1993). *LISREL8: structural equation modeling with the SIMPLIS command language*. Hillsdale, NY: Erlbaum.
- Longford, N.T. (1987). A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects. *Biometrika*, 74, 817-827.
- Longford, N.T. (1990). VARCL. Software for variance component analysis of data with nested random effects (maximum likelihood). Princeton, NJ: Educational testing Service.
- Mason, W.M., Wong, G.M. & Entwisle, B. (1984). Contextual analysis through the multilevel linear model. In S. Leinhardt (Ed.). *Sociological Methodology, 1983-84*. San Francisco: Jossey-Bass.
- McDonald, R.P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods & Research*, 22, 399-413.
- McDonald, R.P. & Goldstein, H. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology*, 42, 215-232.
- Muthén, B. (1987). *LISCOMP. Analysis of linear structural equations with a comprehensive measurement model. Theoretical integration and user's guide*. Mooresville, In.: Scientific Software Inc.
- Muthén, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, 54, 557-585.
- Muthén, B. (1990). *Means and covariance structure analysis of hierarchical data*. Los Angeles: UCLA Statistics series, #62.
- Muthén, B. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement*, 28, 338-354.
- Muthén, B. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, 22, 376-398.
- Muthén, B. & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In R.D. Bock (Ed.) *Multilevel analysis of educational data*. San Diego: Academic Press.
- Muthén, B. & Satorra, A. (1991). Complex sample data in structural equation modeling. Los Angeles: University of California.
- Nelson, G. & Muthén, B. (1991). Analysis preparation steps for multilevel analysis using LISCOMP. Technical report, University of California, Los Angeles.
- Prosser, R., Rashbash, J. & Goldstein, H. (1990). ML3. Software for three level analysis. Users guide. London: Institute of Education, University of London.
- Raudenbush, S.W. & Bryk, A.S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59, 1-17.
- Robinson, W.S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.
- Schijf, H. & Dronkers, J. (1991). De invloed van richting en wijk op de loopbanen in de lagere scholen van de Stad Groningen in 1971. In: I.B.H. Abram, B.P.M. Creemers en A. van der Leij (red.). *Curriculum. Bundel Onderwijsresearchdagen 1991*. Amsterdam: SCO.
- Schijf, H. & Dronkers, J. (1990). Groningen opnieuw bezocht; buurten en schoolloopbanen tussen 1960-1971. In: C.A.C. Klaasen en P.L.M. Jungbluth (red.). *Onderwijs en Samenleving. Bundel Onderwijsresearchdagen 1990*. Nijmegen: ITS.
- Snijders, J.T. & Verhage, F. (1962). *Voorlopige handleiding bij de Groningse Intelligentietest (GIT)*. Lisse: Swets & Zeitlinger.
- Stelzl, I. (1982). *Fehler und Fallen der Statistik*. Bern: Huber.
- Tatsuoka, M.M. (1971). *Multivariate analysis*. New York: Wiley.
- Van Peet, A.A.J. (1992). *De potentieeltheorie van intelligentie*. Dissertation, University of Amsterdam, Amsterdam.