

Boekbesprekingen

Chris Baerveldt, **De school: broedplaats of broeinest? Een vergelijkend onderzoek naar de rol van de school bij de bestrijding en verspreiding van kleine criminaliteit van leerlingen** (proefschrift). Gouda Quint, Arnhem 1990, 171 p.

Het proefschrift van Baerveldt betreft theorie en onderzoek naar de relatie tussen kenmerken van (het onderwijs in) scholen en 'kleine criminaliteit' van leerlingen. Baerveldt ontwikkelt een 'voorwaardelijke bindingstheorie': hij specificeert kenmerken van en voorwaarden voor het groter dan wel kleiner worden van de binding van leerlingen aan school. Een grotere binding hangt samen met een geringere ontwikkeling van kleine criminaliteit. In dit verband worden in totaal 19 hypothesen geformuleerd.

In de onderzoeksopzet wordt gebruik gemaakt van eerder verricht onderzoek. Uit 52 scholengemeenschappen worden er 14 geselecteerd die verwacht worden te variëren in kleine criminaliteit. Het onderzoek wordt gericht op derde klassen mavo (873 leerlingen van 15 tot 17 jaar). Verschillende vormen van dataverzameling worden gehanteerd: interviews, observaties, vragenlijsten. In de analyse van data wordt onder andere gebruik gemaakt van schaalanalyse, sociometrische analyse en multiniveau-analyse.

Bij de presentatie van resultaten blijkt dat een hogere integratie in school inderdaad samengaat met een kleinere mate van kleine criminaliteit. Ook leerlingnetwerken zijn van belang, evenals sekse (*jongens zijn in hogere mate crimineel*). Voor het overige worden de hypothesen, met name die omtrent de relevantie van school- en leskenmerken, weinig of niet gesteund. Baerveldt wijt dit vooral aan het ontbreken van pedagogisch relevante verschillen tussen scholen. Hij bepleit de ontwikkeling van gericht criminaliteitspreventiebeleid in scholen, alsmede een uitbouw van de relatie tussen leerlingbegeleiding en externe professionals op het gebied van de jeugdhulpverlening. De overheid dient het eigen beleid van een school in deze te ondersteunen.

Het onderzoek van Baerveldt staat in de traditie van het onderzoek naar schooleffecten, maar is vooral gericht op sociale effecten. In dit opzicht biedt het proefschrift een welkome aanvulling op hetgeen gebruikelijk is ten aanzien van schooleffectenonderzoek. Opvallend is de sterke benadrukking en ontwikkeling van relevante theorie en hypothesen op het terrein van de criminaliteitspreventie. Op dit gebied zijn theorieontwikkeling en theoriegestuurd onderzoek nog zwak vertegenwoordigd. Ook in methodologisch opzicht is het werk van Baerveldt 'uitstekend'.

Toch heb ik enkele kanttekeningen. De kans om op schoolniveau significante resultaten vast te stellen is, bij 14 scholen, natuurlijk klein. Een verdubbeling van het aantal deelnemende scholen zou waarschijnlijk tot 'meer' (significante) resultaten hebben gevoerd. Dit geldt natuurlijk ook voor het klasniveau. Gezien de opzet van het onderzoek, zou aan dit aspect (meer) aandacht besteed hebben moeten worden. Het zou zeker de moeite waard geweest

zijn! Ten tweede dient geconstateerd te worden dat Baerveldt de leerlingen slechts eenmalig gemeten heeft. Een herhaalde meting, en verklaring van de ontwikkeling bij leerlingen vanuit school-, klas- en leerlingkenmerken, zou mogelijk ook meer aan het licht hebben gebracht. Ten derde is het de vraag of de bepleite inschakeling van leerlingbegeleiding en externe professionals niet vaak ook al te laat en te weinig effectief is. Zou aandacht voor sociale aspecten in het proces van het leren zelf niet het meest effectief zijn om de 'sociale integratie' in school te vergroten?

Ton Mooij

R.K.W. van der Velden, **Sociale herkomst en schoolsucces. Het effect van culturele en sociale hulpbronnen op de schoolloopbaan** (proefschrift). RION, Groningen 1991, 269 p.

Uit verschillende onderzoeken naar sociale ongelijkheid kan men afleiden dat sociale herkomst de schoolloopbaan van leerlingen determineert. Longitudinaal onderzoek dat door het RION in Groningen is uitgevoerd, toont aan dat ook schoolkenmerken invloed hebben op het verloop van individuele schoolloopbanen. Deze dissertatie richt zich met name op de relatie tussen kenmerken die de gezinssocialisatie en de schoolcarrière tijdens het basis- en voortgezet onderwijs karakteriseren. De sociale herkomst van het gezin blijkt een goede voorspeller te zijn van dergelijke culturele gezinskenmerken.

Er worden primaire en secundaire effecten van sociale herkomst onderzocht. Een primair effect betreft de rechtstreekse invloed van sociale herkomst op de prestaties van leerlingen, en bij een secundair effect loopt die invloed via het keuzegedrag van leerlingen voor bijvoorbeeld een bepaald vakkenpakket. Dit onderscheid is van belang vanwege de verklaringsmechanismen die ermee verbonden zijn: gezinsculturele kenmerken en sociale hulpbronnen uit de sociale omgeving van het gezin. Het onderscheid wordt in hoofdstuk 2 verduidelijkt met de theoretische inzichten van onder anderen Bourdieu (cultureel kapitaal) en Boudon (sociaal kapitaal).

In hoofdstuk 3 worden de gegevens beschreven van het Groningencohort dat vanaf 1972 (het moment waarop de deelnemende leerlingen op de basisschool begonnen) een aantal keren is ondervraagd. In 1984 werden ouders en leerlingen, voor zover nog te achterhalen, door de onderzoeker nogmaals geïnterviewd om de onderwijspositie van de leerlingen vast te stellen. Hoewel de representativiteit van het resterende cohort in 1984 te wensen overlaat, neemt de onderzoeker aan dat de non-respons de bevindingen van de causale analyse niet ernstig zal verstoren (p. 55). Om deze opmerking te kunnen verifiëren miste ik een analyse van de non-respons.

In dit proefschrift zijn schoolsucces en schoolloopbaan nauw met elkaar verweven, want schoolsucces wordt geoperationaliseerd als het bereikte (of voor sommige leerlingen het verwachte) onderwijsniveau op 17-jarige leeftijd. Dit is de afhankelijke variabele. Door het reconstrueren van de individuele schoolloopbanen worden de effecten op de bereikte eindpositie vastgesteld van de voorschoolse periode, de basisschoolperiode en de tijd doorgebracht in het voortgezet onderwijs.

In hoofdstuk 5 wordt de invloed gemeten die sociale herkomst heeft op de schoolloopbaan. Eerst wordt berekend wat het totale effect van sociale herkomst op de bereikte onderwijspositie is en in welke schoolloopbaanperiode dit effect optreedt. De resultaten laten zien dat in elke onderscheiden loopbaanperiode milieuspecifieke selectie voorkomt. Wanneer de schoolloopbaan als leerroute wordt bekeken, dan is hiervan eveneens een effect op het be-

reikte onderwijsniveau merkbaar. De keuze van een hoger gewaardeerde leerroute biedt een leerling meer mogelijkheden voor een betere positie op 17-jarige leeftijd (de operationalisering van schoolsucces). Op (vermeende) verschillen tussen jongens en meisjes wat betreft de invloed van loopbaan op de bereikte positie op 17-jarige leeftijd en de effecten van intelligentie en sociale herkomst wordt nader ingegaan. Met name blijkt dan het belang van sociale herkomst voor de bereikte eindpositie. De auteur vermoedt dat dit te maken heeft met de verschillende mogelijkheden waarover leerlingen beschikken, namelijk gezinsculturele kenmerken en sociale hulpbronnen.

De gezinsculturele kenmerken worden in hoofdstuk 6 met schoolsucces in verband gebracht, waarbij het verband gedifferentieerd naar schoolloopbaanperiode wordt geanalyseerd. De relatie tussen sociale herkomst en schoolsucces blijkt slechts voor een gedeelte verklaard te worden door gezinsculturele kenmerken. De totale verklaarde variantie in schoolsucces is namelijk groter dan het gedeelte dat het cultureel kapitaal (van het gezin) voor zijn rekening neemt.

Terecht vraagt Van der Velden zich af of de restvariantie niet verklaard zou kunnen worden door contextuele kenmerken afkomstig van familie, kennissen of buurt. Deze contextuele kenmerken worden gedefinieerd en geoperationaliseerd als sociale hulpbronnen van het gezin, zoals de mate waarin een gezin toegang heeft tot sociale netwerken of woont in een buurt waar het vanzelfsprekend is om een specifieke schoolloopbaan te doorlopen, of waar presteren erg belangrijk wordt gevonden.

Om het afzonderlijke effect van dergelijke sociale hulpbronnen te kunnen meten gebruikt Van der Velden een multilevel-analysemodel, genaamd VARCL. Drie niveaus worden er onderscheiden om schoolsucces te verklaren: het individuele niveau, het gezinsniveau en het buurtniveau met respectievelijk individuele kenmerken, kenmerken van het gezin en kenmerken van de buurt waarin het gezin woont. Deze laatste twee soorten vormen de contextuele kenmerken waarvan het effect op de bereikte positie op 17-jarige leeftijd (schoolsucces) wordt gemeten.

Een contextueel effect van de buurt op het schoolsucces blijkt er nauwelijks te zijn, hetgeen mijns inziens te wijten is aan de gedateerde operationalisatie van het contextuele buurtkenmerk 'sociale achterstand'. Dit is namelijk afgeleid uit volkstellinggegevens van 1971.

Er is wel een gezinskenmerk met een contextuele invloed op het schoolsucces, namelijk de mate waarin een gezin kan putten uit een omringend sociaal netwerk. Of zoals op pagina 201 geconcludeerd wordt: 'het sociale netwerk van het gezin heeft een gunstige invloed op de schoolloopbaan van de kinderen, naarmate het netwerk een meer open karakter heeft, en de status van de erbij betrokken personen hoger is'. Het grootste effect blijft echter afkomstig van het individuele niveaukenmerk sociale herkomst en de daarmee verbonden gezinsculturele waarden en opvattingen als contextuele kenmerken. De sociale hulpbronnen komen dus op de tweede plaats.

Van der Veldens proefschrift is een vlot geschreven onderzoeksverslag waarin de theoretische en procedurele aspecten van de probleemstelling kort en overzichtelijk worden uitgelegd. Jammer is dat de benaming van de afhankelijke variabele nogal eens wisselt. Wanneer men leest over schoolsucces, schoolloopbaan, onderwijsloopbaan en schoolcarrière, dan wordt daarmee bijna altijd aangegeven het bereikte onderwijsniveau op 17-jarige leeftijd.

Deze studie kan men lezen als een pleidooi voor het toepassen van multilevel-analysemodellen in het onderwijssociologisch onderzoek naar leerprestaties. Het is prettig te constateren dat Van der Velden aan het einde van zijn nabeschuiving (op p. 220) hiermee instemt.

Harry Hüttner

Uulkje de Jong, De loopbaan doorlopen. Keuze en selectie tijdens de schoolloopbaan van basisonderwijs tot open universiteit (proefschrift). Universiteit van Amsterdam, Amsterdam 1992, 224 p.

Het proefschrift van De Jong bestaat uit een verzameling (bewerkingen van) reeds eerder verschenen artikelen of congresbijdragen met onderzoek naar schoolloopbanen. In het algemeen is samen met anderen aan het onderzoek gewerkt. Het geheel wordt voorafgegaan door een inleiding waarin De Jong enkele sociologische begrippen op een rij zet (bijvoorbeeld 'organische solidariteit' van Durkheim, 'geinstitutionaliseerd individualisme' van Parsons). Ook het studieloopbaanmodel van Tinto wordt vermeld. Vervolgens worden enkele statistische aspecten van schoolloopbaanonderzoek behandeld. Ten slotte worden de gepresenteerde artikelen en bijdragen geïntroduceerd.

Het eerste onderzoek heeft betrekking op verschillen in keuze van het type voortgezet onderwijs vanwege geslacht en milieu in respectievelijk 1965 en 1977. In 1977 is er minder verschil vanwege geslacht en milieu dan in 1965, hetgeen mede kan samenhangen met wijzigingen in de structuur van het onderwijs. Ten tweede blijkt dat Amsterdamse leerlingen verschillen van leerlingen uit de rest van Nederland wat betreft loopbaan in het voortgezet onderwijs. In Amsterdam wordt een relatief hoger type voortgezet onderwijs gekozen, maar na vijf jaar is deze voorsprong veranderd in een relatieve achterstand, die met name wordt veroorzaakt door een hogere schooluitval. Ten derde: ten aanzien van de vakkenpakketkeuze van leerlingen bestaan er tussen schoolvakken duidelijke verschillen in invloed van aanvangsprestatie, geslacht en milieu. Leerkrachtoordelen voegen weinig toe aan de informatie over het tot stand komen van vakkenpakketten. Ten vierde gaat De Jong na of de structurele verschillen tussen categoriale scholen en scholengemeenschappen kunnen leiden tot verschillen in loopbanen van leerlingen. Er zijn in het algemeen slechts kleine verschillen (een positief effect voor het categoriale vwo/gymnasium). Ten vijfde komt de vraag naar de verschillen vanwege denominatie van de school (openbaar, katholiek, of protestants-christlijk zijn) aan bod. Na analyse van gegevens uit Amsterdam blijkt dat katholieke scholen voor hogere onderwijstypen (havo, vwo) de beste resultaten realiseren bij de leerlingen. De Jong brengt in haar bespreking hiervan naar voren dat in ander onderzoek, in een 'katholiek gebied', een omgekeerd resultaat werd gevonden. Het belangrijkste zou kunnen zijn of de school en ouders ervoor kiezen 'goed onderwijs' te maken. Ten zesde wordt de aandacht gericht op studie-uitval en vertraging in het wetenschappelijk onderwijs. Tussen de studierichtingen worden aanzienlijke verschillen geconstateerd, die mede te maken hebben met de verschillen in kwaliteit van de selectie, registratie en administratie binnen de universiteiten. Dit aspect keert nadrukkelijker terug in de zevende bijdrage, waarin De Jong zich bezighoudt met studietempo als onderdeel van de rendementsdiscussie in het hoger onderwijs (wo en hbo). De belangrijkste factor hier is het 'no-show' percentage (het percentage studenten dat zich wel inschrijft, maar nooit op komt dagen, en dus bestandsvervuilend is). Ten achtste wordt de rol van geslacht bij de studieloopbanen aan de Open Universiteit onderzocht. Deze rol blijkt zeer klein.

De bundel van De Jong staat in de traditie van het schoolloopbanenonderzoek, waarvan een aantal methodologische aspecten en varianten uiteengezet wordt. Dit maakt de bundel interessant en veelzijdig. Anderzijds wordt de theoretische onderbouwing vóórf te weinig uitgewerkt, hetgeen zich mede uit in de soms zwakke operationalisatie. Gegeven het feit dat De Jong vaak werkt met door anderen verzamelde data, is dit wel overkomelijk.

Een ander punt is dat sommige onderzoeken nu reeds verouderd zijn vanwege de komst van de basisvorming en de fusiegolven in het basis- en voortgezet onderwijs. Ook wat betreft het doen van onderzoek markeert het proefschrift waarschijnlijk een afsluiting van een be-

paalde periode. De Jong geeft, zo goed mogelijk, een soort lappendeken van verschillende thematieken, onderzoeken en databestanden. Verwacht mag worden dat de nu lopende cohortonderzoeken in het basis- en voortgezet onderwijs meer systematiek zullen brengen in het loopbanenonderzoek. Hierbij is echter de vraag actueel die ook bij het werk van De Jong klemmend is: zijn de werkelijk werkzame variabelen in onderwijsprocessen wel geïdentificeerd en gemeten? Bij De Jong blijft de 'black box' toch nog erg dicht.

Ton Mooij

A.M. de Vries, Hoe breder hoe beter? De effecten van heteroog groeperen in het voortgezet onderwijs in Nederland (proefschrift). RION, Groningen 1992, 255 p.

Met dit proefschrift wordt beoogd de verhitte discussie over de wenselijkheid van geïntegreerd voortgezet onderwijs een zekere empirische basis te geven: is het voor de betrokken leerlingen nu wel of niet voordelig als zij in een klas terechtkomen met leerlingen met aanzienlijk uiteenlopende cognitieve niveaus? En is het antwoord op die vraag voor 'sterke' en 'zwakke' leerlingen hetzelfde?

Om die vraag te beantwoorden heeft Annemieke de Vries twee bestanden met gegevens omtrent leerlingen met elkaar vergeleken. Het ene, het heterogene bestand, betrof 2500 leerlingen die tussen 1983 en 1985 het geïntegreerd voortgezet onderwijs instroomden. Het andere, het homogene bestand, is een selectie uit het door het CBS samengestelde SMVO-bestand, en wel van leerlingen die in 1977 naar het categoriaal voortgezet onderwijs waren gegaan.

Bij het vergelijken van die bestanden stuit de auteur op alle denkbare problemen: het homogene bestand bestaat uit leerlingen van wie het gros al van school was toen de kinderen uit het heterogene bestand nog moesten beginnen; er is duidelijk sprake van selectieve intrede en uitrede bij deze onderwijsvormen, de doorsnede van de variabelen die in beide bestanden gemeten waren, is opvallend klein, enzovoort.

Desondanks durft de auteur het aan om de vergelijking te maken, waarbij ze een grote vindrijkheid toont. De eerste resultaten van die vergelijking zijn echter niet bemoedigend: 'Het onderwijssysteem verklaart ... weinig tot niets van de variantie in de leerlingresultaten' (p. 133).

Maar dat nulresultaat zou het gevolg kunnen zijn van 'uitmiddelen'. Vandaar dat de onderzoekster nagaat via welke 'mechanismen' de mate van heterogeniteit van de klassamenstelling van invloed zou kunnen zijn op de eindscore van de individuele leerling. Is er sprake van een 'contexteffect' (van het oorspronkelijke klasgemiddelde op de ingangstoets), of van een 'frog-pond effect' (van de afstand tussen de individuele score op de ingangstoets en het klasgemiddelde daarvan) en/of van een 'heterogeniteitseffect' (van de spreiding van de ingangstoets) op de eindscore? Bij een multilevel-analyse blijkt dat in de heteroog samengestelde klassen alleen het contexteffect van enig belang is, en dan nog alleen voor de 'zwakste klassen'. 'Heterogeniteitseffecten worden nergens gevonden. De breedte van de klassamenstelling wat betreft het cognitieve niveau blijkt dus op termijn geen invloed te hebben op de scores van de leerlingen' (p. 180).

Vanaf pagina 175 worden beide bestanden in één analyse betrokken, waarbij alleen nog de waarde op een 'dummy-variabele' het afkomstig zijn uit het ene of uit het andere bestand aanduidt. De manier waarop de auteur ontbrekende gegevens omtrent de afhankelijke variabele in het homogene bestand 'imputeert' en de resultaten van deze gecombineerde multilevel-analyse interpreteert, is een tikje te stoutmoedig voor deze recensent.

Uiteraard levert de analyse een verzameling coëfficiënten op, waarvan sommige ook nog wel het significantieniveau van 5% willen halen. Maar het beeld wordt niet wezenlijk anders: de door de voorstanders van het geïntegreerd onderwijs genoemde positieve effecten van een heterogene klassamenstelling kunnen niet of nauwelijks zichtbaar worden gemaakt. De bij de samenvoeging opnieuw geconstateerde contexteffecten zijn met name 'een gevolg van de invloed van het homogeen cohort; een leerling met een zelfde startniveau bleek immers op een categoriaal schooltype hoger te scoren' (p. 181).

De auteur meent dat de vraag uit de titel 'Hoe breder hoe beter?' uiteindelijk met 'ja' beantwoord kan worden. Want de klassamenstelling mag dan wel geen duidelijk positief effect hebben, negatief is het ook niet; en het geïntegreerd onderwijs heeft daarnaast een aantal andere voordelen boven het categoriaal onderwijs, zoals de mogelijkheid om de schoolkeuze uit te stellen en het zich beter kunnen voorbereiden op die schoolkeuze.

Met zeer veel inventiviteit zijn onwillige databestanden aan een statistische analyse onderworpen. Het zou me zeer verbazen als de fragiele uitkomsten van dit onderzoek het debat tussen voor- en tegenstanders van het heterogeen onderwijs zouden doen verstommen. Wat die uitkomsten ons in ieder geval leren, is dat het verstandig kan zijn om, alvorens allerlei diepzinnige theorieën te ontwerpen ter verklaring van klasse-effecten (en paragraaf 2.3 biedt daar een fraai overzicht van), eerst eens na te gaan of men wel bezig is met de verklaring van een bestaand verschijnsel.

J. van der Zouwen

P. van den Eeden, U. de Jong, P. Koopman en J. Roeleveld (met een inleiding van G. Meijnen), **Schoolloopbanen in Amsterdam. Leerlingen op Amsterdamse scholen voor algemeen voortgezet onderwijs in de jaren zeventig**. Swets & Zeitlinger, Amsterdam/Lisse 1993, 181 p.

Met de invoering van de basisvorming wordt beoogd om naar prestatieniveau heterogeen te groeperen, waardoor leerlingen uit lagere statusgroepen, etnische minderheden en meisjes gunstiger schoolloopbanen zouden krijgen. Dit zou bereikt moeten worden door de schoolkeuze uit te stellen en een grotere flexibiliteit in leerroutes te bereiken. In het boek wordt dan ook de interessante vraag gesteld of heterogeen groeperen de gewenste effecten heeft.

De auteurs onderscheiden drie verklaringen voor het ontstaan van verschillen in schoolsucces: door individuele achtergrondkenmerken van leerlingen, zoals hun geslacht en milieu van herkomst (sociale selectie); door de voorafgaande leerprestaties (meritocratische selectie); en door de school en de klas (systeem-immanente selectie). Een van de kenmerken die daarbij een rol spelen, is de wijze van groeperen. Dit kan homogeen of juist heterogeen zijn naar prestatieniveau. Ook andere school- en klaskenmerken kunnen echter een rol spelen, zoals de klasgrootte of denominatie. De studie concentreert zich voornamelijk op deze drie vormen van selectie en de wijze waarop deze onderling samenhangen. De vragen worden beantwoord met behulp van gegevens van Amsterdamse leerlingen die in 1969, 1973, 1974 en 1975 in het algemeen voortgezet onderwijs zijn begonnen.

Na beschrijving van de gegevens wordt vastgesteld dat in de onderzochte periode weinig is veranderd in het soort leerlingen dat de Amsterdamse scholen bevolkt en de resultaten die zij behalen. Meer dan de helft van de verschillen in de schoolkeuze kan worden verklaard door de eerdere schoolprestaties en in mindere mate door het sociaal milieu van herkomst.

De volgende stap is om de invloed van de scholen te schatten. Als de effectiviteit van scholen van hetzelfde type wordt vergeleken, moet worden geconcludeerd dat de schoolloopba-

nen van leerlingen met gelijke achtergrondkenmerken niet sterk verschillen tussen scholen van een gelijk type: minder dan tien procent van de variantie in het eindniveau wordt hiermee verklaard. Voorts blijkt dat deze verschillen in effectiviteit niet constant zijn in de tijd. Voor het ene cohort leerlingen is een school redelijk effectief, voor een ander cohort is dezelfde school veel minder effectief. Dit laat zien dat conclusies uit eerdere schooleffectiviteitsonderzoeken gerelativeerd moeten worden, omdat daarin meestal sprake was van een onderzoek naar één cohort.

De hamvraag is of de wijze van groeperen van belang is voor het schoolsucces. Daarvoor vergelijken de auteurs de schoolloopbanen van leerlingen met een gelijk advies op categorale scholen en scholengemeenschappen. De gedachte hierachter is dat categorale scholen een naar prestatieniveau en sociaal milieu homogener populatie leerlingen heeft dan scholengemeenschappen. Als de effecten hiervan worden onderzocht blijken deze nogal tegen te vallen. Na controle voor instroomkenmerken zijn de leerlingen op categorale scholen vrijwel even goed af als leerlingen op scholengemeenschappen.

Op de gekozen analysemethoden en de wijze waarop deze zijn gebruikt valt weinig aan te merken. Wel zou voor het bereiken van een groter publiek meer aandacht besteed moeten worden aan de leesbaarheid. Een deel van de beschrijvende resultaten zetten ongeofende lezers al snel op het verkeerde been, zodat vooroordelen gemakkelijk bevestigd kunnen worden. Dit vinden we bijvoorbeeld terug in de resultaten die gepresenteerd worden over de effectiviteit van scholen, zonder dat rekening gehouden wordt met de verschillen in de samenstelling van de leerlingpopulaties.

De vraag naar de effecten van groeperen kan men rustig als pretentius omschrijven. In feite wordt vooruitgelopen op de evaluatie van de basisvorming, waarin het heterogeen groeperen een van de elementen is. Het blijkt dan ook dat het antwoord dat wordt gegeven niet erg bevredigend is. Wellicht kan ook niet worden verwacht dat op basis van deze gegevens een antwoord op de vraag gegeven kan worden. De gegevens zijn immers sterk verouderd, lokaal beperkt en niet toegesneden op het beantwoorden van de vraag. De operationalisering van de wijze van groeperen met behulp van de tegenstelling tussen categorale schooltypen (homogeen) en scholengemeenschappen (heterogeen) is het beste voorbeeld hiervan. Zo wordt in veel scholengemeenschappen die op zichzelf een heterogene categorie leerlingen bevatten, reeds in het brugjaar intern gedifferentieerd naar bijvoorbeeld het advies of de prestatiescores. Zo kunnen deze scholengemeenschappen homogene groepen herbergen die binnen de gegevens niet zichtbaar worden.

Zoals de auteurs terecht opmerken, is de context waarbinnen leerlingen opereren van belang voor hun schoolloopbanen. Voor scholen geldt dat ook hun context van belang is voor de effectiviteit, hetgeen zijn weerslag heeft op de schoolloopbanen van hun leerlingen. De context waarbinnen zij opereren wordt bijvoorbeeld bepaald door de onderlinge concurrentieverhoudingen en de wet- en regelgeving. De context van scholen in Amsterdam in de jaren zeventig verschilt dermate sterk van die van Nederlandse scholen in de jaren negentig, dat ik de onderzoeksresultaten niet als richtinggevend voor de discussie rond de basisvorming beschouw. Bovendien laat de basisvorming zoals deze uiteindelijk is ingevoerd, de mogelijkheid aan scholen voor een eigen systeem van groeperen; dat betekent dat het welslagen van de basisvorming voornamelijk af zal hangen van de wijze van invulling door de scholen zelf.

Bart Bakker

R.J. Bijleveld, **Numeriek rendement en studiestaking. Een theoretische analyse van factoren die samenhangen met rendement en studiestaking in het wetenschappelijk onderwijs** (proefschrift). Lemma, Utrecht 1993, 281 p.

Wat is het verband tussen het onderwijsproces van studenten en de effectiviteit van het wetenschappelijk onderwijs? Bijleveld stelt zich in zijn proefschrift tot doel vanuit het multilevel-perspectief daarover een theorie te ontwikkelen. Na de relatie tussen onderwijsproces en onderwijsrendement in de bestaande onderzoeksliteratuur te hebben beschreven past hij het resultaat toe op het rendement van studierichtingen. Het zijn immers de studierichtingen die een belangrijke rol spelen in de inrichting van het huidige wetenschappelijke onderwijs in Nederland.

Bijleveld vat het rendement van een studierichting eenvoudigweg op als het aandeel studenten dat zijn studie voortzet. Persoonlijke capaciteiten van studenten of investeringen van de studierichting spelen voor hem geen rol. Hij richt zich niet op het niveau van de studierichting om te zien hoe de kenmerken daarvan met rendement samenhangen. Integendeel, hij probeert die samenhang begrijpelijk te maken door te wijzen op het studieproces van de studenten. De uitkomst is een theoretisch model waarin dan ook student en studierichting als afzonderlijke niveaus voorkomen. De kern van Bijlevelds model is dat geringe inspanning en intelligentie van studenten tot lage studieresultaten leiden. Die hebben tot gevolg dat studenten hun studie staken. De inspanning van studenten hangt weer af van interesse en studiedruk, die elk blootstaan aan de invloed van kenmerken van studierichtingen. De interesse van studenten gaat gepaard met het aandeel probleemgestuurd onderwijs en interdisciplinaire werkvormen in het curriculum. De studiedruk van studenten wordt groter met de hoeveelheid studiebegeleiding, het aandeel actieve werkvormen, enzovoort. Ook de intellectuele bekwaamheid van studenten houdt in Bijlevelds model verband met kenmerken van hun studierichting, met name met de hoeveelheid wiskunde in de opleiding en het aantal mogelijkheden om onvoldoenden te compenseren.

Dit theoretisch model gaat ervan uit dat de kenmerken van een studierichting rechtstreeks inwerken op de diverse fasen van het studieproces die voorafgaan aan de beslissing van studenten met hun studie op te houden danwel die voort te zetten. Maar, ligt het niet eerder voor de hand aan te nemen dat het studieproces zelf afhangt van de voorwaarden die de gekozen studierichting biedt? Het is toch op grond hiervan dat de studierichtingen de studieprocessen van studenten feitelijk trachten te sturen? Uitgaande van deze aanname is daar inmiddels uitgebreid empirisch onderzoek naar gedaan. Dat onderzoek wordt overigens vrijwel niet besproken in Bijlevelds literatuursamenvatting. De beschreven literatuur had misschien tot het inzicht van het voorwaardenscheppende karakter van de studierichting kunnen leiden, als zij correct was samengevat (bijvoorbeeld door gebruik te maken van de mogelijkheden van meta-analyse) en als zij aan een kritische evaluatie onderworpen was geweest. De vraag hoe kenmerken van een studierichting en het rendement precies samenhangen, blijft nu onbeantwoord. Welke zijn de theoretische mechanismen die ervoor zorgen dat de kenmerken van een studierichting bijdragen aan de beslissing van studenten om al dan niet met hun studie op te houden? Op welke algemeen-theoretische uitgangspunten steunen die? Bijlevelds model maakt daarmee een willekeurige indruk, temeer daar hij geen moeite heeft gedaan zijn model aan het risico van empirische toetsing bloot te stellen.

Dan was gebleken dat Bijlevelds theoretisch model te kort schiet voor gebruik in empirisch onderzoek naar rendementen van studierichtingen. Zo komen de intermediaire niveaus van de jaargangen en de gegeven vakken niet aan bod. Deze tekorten zijn opmerkelijk, omdat er een studie van De Jong, Koopman en Roeleveld voorhanden is (*Snelwegen en slingerpaden in en om het Hoger Onderwijs*, Den Haag: Ministerie van Onderwijs en Wetenschappen,

1991), waarin eenzelfde vraagstelling een meer genuanceerd model heeft opgeleverd, dat bovendien wel empirisch is getoetst. Het was vruchtbaarder geweest wanneer Bijleveld zijn theoretisch model op een dergelijke studie had gebaseerd en zelf enige moeite had gedaan zijn resultaat op empirische houdbaarheid te onderzoeken.

Pieter van den Eeden

J.J. Hox, **Applied multilevel analysis**. TT-Publikaties, Amsterdam 1994, X+112 p.

'Applied multilevel analysis' is een niet-technische inleiding in de multilevel-analyse en is bedoeld voor onderzoekers die het model willen toepassen in de praktijk. Het geeft een overzicht van de verschillende technieken en van toepassingen van die technieken. De aanpak van verschillende soorten vraagstukken wordt helder uiteengezet en van methodologisch commentaar voorzien. De stappen in de analyse worden helder aangegeven. Kortom, het is een nuttig en goed leesbaar boek. Bij het boek wordt een schijfje geleverd met voorbereerde datasets in een format waarin ze ingelezen kunnen worden in een van de multilevel-programma's HLM, VARCL of ML3. Hulpfiles zijn van tevoren klaargemaakt.

Het multilevel-model wordt beschreven als een multilevel-variant van het multipelere-regressiemodel, met vergelijkingen op meerdere niveaus. Drie klassen van multilevel-modellen worden besproken: de standaard multilevel-regressiemodellen, gegeneraliseerde lineaire modellen voor multilevel-analyse van proporties en multilevel-modellen voor covariantiestructuren (multilevel-factoranalyse en padanalyse). Bekende toepassingen zijn die voor longitudinale data, groeicurves, interview-surveydata en meta-analyse.

Hoofdstuk 1 is een inleiding in de multilevel-analyse en een definitie van het model, het hiërarchische lineaire model: '... a hierarchical system of individuals and groups, with individuals and groups defined at separate levels of this hierarchical system (...) This leads to interaction between variables that describe the individuals and variables that describe the social groups'. Andere benamingen voor het multilevel-model zijn 'random coefficient model' en 'variance component model'. Deze benamingen refereren aan het feit dat het intercept en de regressiecoëfficiënten een random parameter krijgen in de vergelijking, ter verklaring van de wisselende invloed die variabelen op een hoger niveau uitoefenen op variabelen op het lagere niveau.

Hoofdstuk 2 ('Multilevel Regression Models') beschrijft het basismodel, twee schattingsmethoden, FML (Full Maximum Likelihood) en RML (Restricted Maximum Likelihood), de verschillen daartussen en de implicaties voor het schattingsproces met behulp van een specifiek pakket. Met RML kunnen uitsluitend verschillen in de random variantiecomponenten getoetst worden, met FML ook in de fixed componenten. Deze informatie is gekoppeld aan al even nuttige informatie over de gebruikte methodes in de verschillende pakketten. Het iteratieve schattingsproces wordt besproken, met de nodige informatie van een ervaren gebruiker. Er wordt een exploratieve methode besproken voor de opbouw van de modelvergelijking door te beginnen met het simpelste model, het 'intercept only model' zonder verklarende variabelen, dan het model voor het laagste niveau, en als laatste wordt het model voor alle niveaus bepaald, met de cross-level-interacties.

Hoofdstuk 3 is een bespreking van de drie pakketten HLM, VARCL en ML3 en bevat zeer veel nuttige informatie en praktische tips: een 'must' voor de beginner, niet in de laatste plaats vanwege de afwijkende notatie in de output van de pakketten en de verschillende mogelijkheden die de pakketten bieden. De opbouw van een model met behulp van de drie pakketten wordt beschreven en uitgewerkt aan de hand van een voorbeeld.

In hoofdstuk 4 wordt een aantal toepassingen van het multilevel-regressiemodel besproken. Hier komen de *gegeneraliseerde* multilevel-modellen aan de orde. Zeer aardig is de bespreking van de relatie en de verschillen tussen het (algemene) multilevel-model en het *gegeneraliseerde* multilevel-model. De overdispersiefactor dient wel degelijk geschat te worden. Het effect hiervan is dat parameterschattingen over het algemeen minder significant worden.

In hoofdstuk 5 worden meer geavanceerde technieken besproken; multilevel-technieken voor covariantie-structuuranalyse. Eén voorbeeld is multilevel (confirmatorische) factoranalyse, een ander voorbeeld betreft multilevel-padanalyse. Deze modellen worden geanalyseerd met behulp van LISREL, EQS of LISCOMP. Ondanks het hoge abstractieniveau worden de modellen begrijpelijk en interessant beschreven.

De bespreking van de voorbeelden is doorspekt met methodologische adviezen van een ervaren gebruiker. In dit onderzoeksgebied is dat bepaald niet overbodig. Voorbeelden zijn het centreren van variabelen, het standaardiseren van regressiecoëfficiënten, de opbouw van een model, de voor- en nadelen van onderspecificatie respectievelijk overspecificatie van een model tijdens de schattingsprocedure, de interpretatie van interacties (en de verplichte aanwezigheid van de bijbehorende hoofdeffecten), de codering van variabelen (0,1 of 1,2), het definiëren van contrastvectoren en het effect van (ontbreken van) conditionele onafhankelijkheid. De kennis van de ervaren gebruiker wordt helder en efficiënt overgebracht. Misschien was het overzichtelijker geweest om een aantal methodologische aanwijzingen in één speciale paragraaf te bespreken: ze overstijgen het geanalyseerde voorbeeld.

Schoonheidsfoutjes zijn er ook: verwisselde indices in formules (2.9, 3.11) en de definitie van de variantie/covariantie van de residuele errors u_{0j} en u_{1j} (p. 13) wordt niet consequent doorgevoerd (p. 34, 49, 59 en verder), maar daar leest men wel overheen.

Rian Vogelesang

Henk Guldemon, Van de kikker en de vijver. Groepseffecten op individuele leerprestaties (proefschrift). GARANT, Leuven-Apeldoorn 1994, 133 p.

In het multiniveau-onderzoek worden relaties onderzocht tussen de kenmerken van eenheden van verschillend niveau. Bijvoorbeeld, voor de verklaring van ziekteverzuim bij arbeiders worden niet alleen persoonskenmerken (functie, leeftijd, ...), maar ook bedrijfskenmerken (sector, personeelsbeleid, ...) opgenomen. In de sociologie neemt dit onderzoekstype overwegend de vorm aan van contextueel onderzoek. Het is immers de gedachtengang van sociologen dat het gedrag van individuen niet alleen door individuele kenmerken, maar ook en vooral door kenmerken van de sociale context moet worden verklaard. In de onderwijssociologie wordt deze context gevormd door de school of de schoolklas. Een contextueel onderzoek bevat dan een afhankelijke variabele op microniveau (studieprestaties van leerlingen), verklarende variabelen op microniveau (sekse, etniciteit en intelligentie van leerlingen) en verklarende variabelen op macroniveau (gemiddelde studieprestatie van de schoolklas, onderwijsaanbod van de school).

Op dit onderzoekstype is felle kritiek uitgeoefend door Erbring en Young. Zij hekelen het erin vervatte structuurbegrip. Want ten eerste, de sociale structuur wordt er weergegeven door een simpele maat van groepsamenstelling (een gemiddelde score) en niet door de processen van sociale interacties tussen individuen. En ten tweede, in het contextueel onderzoek is er sprake van een magisch verschijnsel van sociale telepathie. Want er wordt bijvoorbeeld gesteld dat de gemiddelde socio-economische status van de ouders van leerlingen van een schoolklas invloed heeft op de studieprestaties van die leerlingen. Maar dat zou betekenen

dat elke leerling contact heeft met de ouders van alle andere klasgenoten. En daar deze hypothese implausibel is, berust het model op sociale telepathie, op actie van een afstand, hetgeen een gekend principe is uit de magie, maar niet uit de wetenschap.

Om van de sociale telepathie af te geraken, formuleren Erbring en Young nieuwe voorstellen. Hun belangrijkste alternatief model is dat van sociale besmetting, ook wel endogeen feedbackmodel genoemd. Het contextueel model wordt hier geïnterpreteerd als het resultaat van sociale interacties tussen individuen. De gedachte is dat het presteren van leerling j (Y_{ij}) een impact heeft op en zelf ook wordt beïnvloed door de prestaties van de relevante andere leerlingen j' ($Y_{ij'}$), met andere woorden: presteren steekt aan. De statistische analyse is nogal gecompliceerd en legt een groot beslag op het beschikbare geheugen van de computer. Belangrijk in deze analyse is de opname van een matrix W , die de structuur van de interacties tussen de leerlingen weergeeft. Deze vierkante matrix bevat nullen op de hoofddiagonaal, omdat individuen niet geacht worden met zichzelf te interacteren. Als op andere plaatsen in de matrix overal hetzelfde getalletje staat, dan wordt ervan uitgegaan dat de interacties tussen paren van individuen alle even intens zijn. Op basis van sociometrische informatie kunnen deze gewichten worden gewijzigd.

Henk Guldemond is bij mijn weten de eerste die dit Erbring-Young-model heeft toegepast in zijn op 6 december 1994 verdedigde dissertatie 'Van de kikker en de vijver'. Net als Erbring en Young, maar in een minder retorische stijl, verdedigt hij de gedachte dat de leerling-leerlinginteracties van cruciaal belang zijn in het contextueel onderzoek.

De wijze waarop hij dit verdedigt is een verdediging in de positieve zin. Men zou hier kunnen vergelijken met een pianist die een nocturne van Chopin 'verdedigt'. Iedereen die wel eens Chopin heeft proberen te vertolken, weet hoe moeilijk de linkerhand is, want deze moet dynamisch ingetogen en in een vloeiend legato worden gebracht, zonder terecht te komen in de valstrik van het aangedikte sentiment door een verkeerd pedaalgebruik of een opdringerig rubato. Dit terwijl de rechterhand een vlotte beweging maakt over het toetsenbord. De linkerhand legt het fluwelen tapijt neer, waarop de rechterhand dartel binnentreedt. Het fluwelen tapijt is hier de brede discussie over contextueel onderzoek in onderwijssituaties. De rechterhand is de toepassing van het Erbring-Young-model. De auteur heeft de nocturne voortreffelijk verdedigd.

In de eerste vier hoofdstukken voert hij de theoretische en methodologische discussie van het contextueel onderzoek. In de laatste vier hoofdstukken presenteert hij zijn eigen empirisch onderzoek. De theoretische discussie draait rond twee vormen van interactie in de schoolklas, de leerling-leerlinginteractie en de leerkracht-leerlinginteractie. Wat deze laatste betreft worden drie onderzoeksstromingen besproken: deze der effectieve instructie, het onderzoek naar de invloed van homogeen of heterogeen groeperen in schoolklassen en het onderzoek naar het (labeling) effect van de leerkrachtverwachtingen. De reden voor de inconsistente onderzoeksbevindingen van met name het effectiviteitsonderzoek wordt gelegd in de eenzijdigheid van de modellen van effectieve instructie. Eenzijdig in de zin dat vrijwel alleen de technisch-instrumentele aspecten van het onderwijzen aan bod komen en de sociaal-psychologische processen onbelicht blijven.

Een vergelijkbare diagnose wordt gemaakt voor het onderwijsonderzoek aangaande leerling-leerlinginteracties. De referentietheorieën van Festinger, Merton, Homans, Kelley en Davis dienen hier als uitgangspunt. Van Festinger komt de sociale-vergelijkingstheorie. Van Merton komt, in navolging van Stouffer, de referentiegroepstheorie en het erin vervatte begrip relatieve deprivatie in al zijn verschijningsvormen. Bij zowel Festinger als Merton zijn de referenties comparatief, maar de vergelijkingsanderen fungeren hier toch slechts als een passieve sociale meetlat, waaraan het individu zijn positie kan aflezen. Heel anders ligt dit in de normatieve referentiegroepstheorie van Homans. Groepen worden hier opgevat als norm-

stellend, en omdat het aanzien van een persoon stijgt al naar gelang zijn gedrag meer in overeenstemming is met de normen, ontstaat er een hiërarchische ordening.

In navolging van Kelley analyseert de auteur zowel de comparatieve als de normatieve functie van referentiegroepen in zijn eigen onderzoek. Ook de frog-pond-theorie van Davis is voor hem belangrijk. In 'The Campus as a Frog Pond' deed Davis een beroep op Mertons relatieve deprivatietheorie alsook op Kelleys normatieve referentiefunctie en kwam hij tot de conclusie dat *It is better to be a big frog in a small pond than a small frog in a big pond*, met andere woorden: in het land der blinden is eenoog koning.

Met deze theorieën voor ogen stelt Guldemond vast dat al het onderwijsonderzoek naar contextuele effecten aan hetzelfde euvel lijdt: de indirecte wijze waarop de context wordt gemeten, bijvoorbeeld door middel van geaggregeerde groepskenmerken, en het ontbreken van de sociaal-psychologische vergelijkingsprocessen in de leerling-leerlinginteractie. Hetzelfde geldt volgens hem voor de methodologie van het gangbare multiniveau-onderzoek. Hijzelf daarentegen betreft sociometrische data in zijn onderzoek en maakt gebruik van het endogeen feedbackmodel van Erbring en Young. Voor het bepalen van de gewichten in de matrix *W* doet hij beroep op de grafenanalyse. Door gebrek aan ruimte en vanwege het technische karakter kunnen de details van zijn onderzoek, bij een steekproef van 176 uit 8300 Nederlandse basisscholen (met 186 groepen en in totaal 3648 kinderen), hier niet worden weergegeven. Maar ik verzeker de lezer dat het met kennis van zaken is uitgevoerd. De negatieve kritieken die ik wou geven – de steekproefproblemen, de missing-data-problemen, de beperkte operationalisering van de sociometrische data, het benodigd computergeheugen en het gevaar voor Hineininterpretierung – worden door de auteur zelf gesignaleerd en besproken. Een gave rechterhand dus in de uitvoering van Chopins nocturne.

Er werden geen frog-pond-effecten gevonden. In die zin brengt de titel van de dissertatie de lezer op een verkeerd spoor. De belangrijkste conclusie van het onderzoek is dat het duidelijk aantoonbare effect van de context hoofdzakelijk via de sociaal-emotionele dimensie (vriendschappen) en niet via de taakgerichte dimensie verloopt: in klassen waarin populaire kinderen goede leerprestaties leveren, ontstaat door de voorbeeldfunctie een positieve impuls, en omgekeerd. Verder is duidelijk gemaakt dat homogeen groeperen niet wenselijk is, omdat dit de lage niveaugroepen in een neerwaartse spiraal brengt.

Terugblikkend met een zekere distantie kan worden gezegd dat deze dissertatie slechts is wat zij is: een pleidooi voor een andere aanpak in het contextueel onderzoek. Maar, net als Chopin het instinct of de wijsheid bezat zich bij zijn eigen wereld te houden, dat koninkrijk waar hij absoluut alleenheerser was: de piano, zo ook geldt voor Henk Guldemond: *In der Beschränkung zeigt sich der Meister*.

Jacques Tacq

Jaap Roeleveld, **Verschillen tussen scholen. Kenmerken, effectiviteit en stabiliteit van onderwijsinstellingen in Nederland** (proefschrift). SCO-Kohnstamm-Instituut, Amsterdam 1994, 238 p.

Promoveren op eerder gepubliceerde artikelen. Dat is de formule voor Roelevelds proefschrift, een niet onverstandige keuze voor opdrachtonderzoekers die alleen marginaal ruimte krijgen om aan hun werk de academische status te geven die het verdient. Zeven artikelen heeft Roeleveld op rij gezet en voorzien van een methodisch-theoretisch inkaderende introductie.

Anders dan de titel doet vermoeden, biedt het proefschrift geen overzicht van verschillen

tussen scholen of van kenmerken en stabiliteit van onderwijsinstellingen. Output-effectiviteit is het sleutelbegrip: leerprestaties, gecorrigeerd voor instroomverschillen. Een actueel topic in de onderwijssociologie, die langzaam maar zeker haar aandacht richt op wat altijd al haar formele object was: het onderwijs, en niet langer vooral de sociale achtergrond van leerlingen die in zo hoge mate resistent lijkt voor de vele pretenties van onderwijsmensen.

In zijn introductie – en daarna in vrijwel elk hoofdstuk – geeft Roeleveld duidelijk weer te kiezen voor een wel erg conservatieve interpretatie van het empirisch materiaal: zijn de verschillen in output tussen scholen niet betrekkelijk? Het is toch niet zo dat alle leerlingen op de ene school een heel andere prestatie neerzetten dan alle leerlingen van een andere school!? En dat geldt al evenmin wanneer je de achtergrond van leerlingen verdisconteert, die zoals bekend een niet onbelangrijke voorspeller vormt van schoolloopbanen.

Dat schooleffectiviteit behalve mathematisch ook inhoudelijk een problematisch begrip is, onderkent Roeleveld expliciet. Het is een hypothetisch construct, een optelsom van de effectiviteit van diverse leerkrachten, liefst nog gedifferentieerd naar diverse klassefactoren of 'peergroups'. De determinanten van die effectiviteit lijken elk afzonderlijk weinig indrukwekkend en bovendien contextgebonden. Alles bij elkaar 'a nice mess' voor onderwijssociologen die graag als volwassen wetenschappers een bijdrage willen leveren aan schoolverbetering.

Roeleveld maakt het zichzelf voorwaar niet makkelijk. Zo wordt in het eerste inhoudelijke hoofdstuk getracht om maar liefst per school voor voortgezet onderwijs een model op te stellen voor de determinanten van schoolloopbanen. Dat (b)lijkt van school tot school te verschillen met als resultaat een weinig overtuigende en vage conclusie. Met dat probleem wordt de lezer vaker opgescheept. Dat de tussenschoolse variantie groter is dan de binnenschoolse is aardig om te weten, maar alleen verwonderlijk voor wie naïef start: geldt het niet voor het gros van alle relevante ongelijkheidsindelingen? Dat scholen relatief stabiel zijn in hun output, terwijl de directies het leven ondergaan als een voortdurend bombardement van beleidsinterventies, ook dat is niet verwonderlijk voor wie de pedagogische provincie kent, met haar voorkeur voor schijnwerelden en haar wisselende waan van de dag.

Ronduit interessant is de bijdrage die ingaat op de stabiliteit van schooleffectiviteit over meerdere jaren. Die is in de gepresenteerde resultaten ongecorrigeerd-voor-instroom constanter dan gecorrigeerd. Helaas raakt de auteur daardoor in twijfel over het nut van het effectiviteitsbegrip, in plaats van door te gaan op de these dat scholen net als ziekenhuizen altijd hun bedden wel vol krijgen maar ook nooit meer dan dat. Zo'n bevinding is toch bepaald van groot belang voor de betrokken leerlingen!

Misschien is onderwijseffectiviteit juist daarom zo moeilijk te 'pakken', omdat ze als de uitkomst wordt gezien van een planmatig en professioneel gebeuren, terwijl ze in werkelijkheid de uitkomst is van de verborgen intern-organisatorische logica van scholen, waarvan veel onderwijskundigen zich in hun theorievorming zo wars tonen, en van de schrijnend geringe mate waarin het onderwijsvak daadwerkelijk geprofessionaliseerd is, dat wil zeggen: in zijn beroepsmatig handelen voorspelbaar, ongeacht de toevallige persoonlijke voorkeuren van de bekleder van de functie. Die twee factoren samen maken dat leerlingen behalve door de sociale factoren die hun zijn meegegeven bij geboorte, in hun onderwijskansen nog wees extra ongelijkheid oplopen door het 'toeval' van de effectiviteit van hun school.

Paul Jungbluth

Harry J.M. Hüttner en Pieter van den Eeden, **The multilevel design: a guide with an annotated bibliography 1980-1993**. Greenwood Press, Westport 1995, 276 p.

Reeds gedurende meerdere decennia ken ik de methodologen Harry Hüttner en Pieter van den Eeden als de pleitbezorgers van multiniveau-onderzoek. Zij laten geen gelegenheid voorbijgaan om een boek of artikel te publiceren, een congres of colloquium te organiseren, een lezing te verzorgen, een debat te voeren of een workshop te leiden over dit onderwerp. Zo ook dit Engelstalige boek met uitgebreide bibliografie. Ongeveer 200 van de 250 pagina's bevatten een opsomming van 589 verschenen publikaties over multiniveau-onderzoek in de periode 1980-1993, voorzien van een abstract. Ongeveer 20% van de auteurs is van Nederlandse origine en ongeveer 10% van de publikaties is in het Nederlands gesteld. Het moet een monnikenwerk geweest zijn om dit met zijn tweeën allemaal te verzamelen en samen te vatten. Er kan misschien worden aanbevolen dit bestand te integreren in het Social Research Methodology (SRM) bestand, welke database zowel op cd-rom als via on-line verbinding kan worden geraadpleegd, alsook in de informatietijdschriften die door SRM worden uitgegeven.

De auteurs laten deze geannoteerde bibliografie voorafgaan door 53 pagina's bespreking van het multiniveau-design. Dit is de 'Guide', waar de titel naar verwijst. Bij deze tekst bekruipt de lezer een dubbel gevoel. Aan de ene kant worden heel wat onderwerpen behandeld die belangrijk zijn en alle aandacht verdienen. Aan de andere kant is de tekst niet scherp en krachtig geschreven, waardoor de concentratie van de lezer gaandeweg afneemt.

Na de term 'multilevel' te hebben omschreven als [verwijzend naar] hiërarchisch gestructureerde eenheden, en na te hebben vastgesteld dat er een groeiende belangstelling is voor het multiniveau-onderzoek, worden vier zaken onderscheiden: de benadering, de theorie, het model en de procedure. De benadering (approach) verwijst naar de methodologische visie op de samenleving als hiërarchisch gestructureerd, waarbij de niveaus relatief autonoom, relatief interdependent of causaal verbonden zijn in de ene of de andere richting. Een theorie wordt omschreven als een verzameling van uitspraken die relaties tussen feiten in de werkelijkheid verklaren. Zo'n theorie kan monolevel zijn op het hogere niveau (de informatie prestatie-theorie), monolevel op het lagere niveau (de prestatie motivatie-theorie) of multilevel. In de bespreking van de 'multilevel theory' wordt de prestatie motivatie-theorie (op-niveau-leerling) uitgebreid met een conditionele variabele van hoger niveau (op-niveau-schoolklas). Een model wordt opgevat als een geformaliseerde uitdrukking van de theorie, gewoonlijk in de vorm van wiskundige vergelijkingen. Het model van de tweestapsprocedure en het 'random coefficient model' worden als voorbeelden gegeven. De procedure is de analysemethode voor het schatten en testen van modellen, zoals de Maximum Likelihood-schatting en de gewogen kleinste-kwadratenschatting.

De auteurs hopen met deze specificatie van benadering, theorie, model en procedure: 'to have offered clarity in the confusion of tongues'. Maar nog afgezien van het feit dat deze termen in de moderne wetenschapsfilosofie andere betekenissen krijgen, is het jammer dat zij in hun eigen tekst weinig gebruik maken van deze termen en overigens andere termen invoeren, zoals multiniveau-analyse en multiniveau-design, zodat de aandacht voor het onderscheid verloren gaat.

Vervolgens wordt ingegaan op multiniveau-theorieën waarin contextuele effecten worden verklaard. Vier mechanismen worden onderscheiden en per mechanisme wordt een aantal theorieën besproken: het zich conformeren (normative reference group theory, community-identification theory, differential exposure model theory), keuzegedrag (comparative reference group theory, relative deprivation theory, frog pond effect theory, peer-group theory, balance theory, opportunity structures, self-selection), ontmoetingskansen (contagion theo-

ry) en adaptatie (*presence of resources, differential impact theory*). De auteurs zijn van oordeel dat de meeste van deze mechanismen terug te vinden zijn in de rationele-keuzetheorie en beschouwen deze als een vruchtbaar uitgangspunt voor theoretische vernieuwing in het multiniveau-onderzoek, een apologie waar niet iedereen het mee eens zal zijn. Met deze aanbeveling treden zij in het voetspoor van James Coleman, die in zijn lijvig werk 'Foundations of Social Theory' (1990) de rationele-keuzetheorie nader uitwerkt, een werk dat in deze bibliografie merkwaardig genoeg niet is opgenomen.

In een volgend hoofdstuk wordt het multiniveau-design besproken. Ook hier wordt weer beloftevol gestart met de vaststelling dat multiniveau-onderzoek essentieel gebaseerd is op het multistage sample design, iets waar de geïnteresseerde lezer graag meer over zou vernemen, maar sample designs blijven verder onbesproken.

Na het aanduiden van diverse soorten fallacies als verwarringen tussen niveaus, variabelen en relaties (*category-as-a-unit fallacy, variable-as-a-unit fallacy, fallacy of the actual population*) worden weer andere fallacies gerelateerd aan de verschillende fasen van de empirische cyclus: onderzoeksprobleem, conceptueel model, operationalisatie, typologieconstructie, analyse en interpretatie. De analysefase krijgt de meeste aandacht (tweestapsmodel, hiërarchisch lineair model, macro model, micro model, contextueel model) en een aantal computerprogramma's (ML3, HLM3, VARCL3) wordt aangeduid. Ook bij deze tekst bekruipt de lezer een dubbel gevoel. Enerzijds bevat hij belangrijke zaken, zoals de waarschuwing voor de gevaren bij het gebruik van geaggregeerde data, en heldere uiteenzettingen, zoals de opbouw van het tweestapsmodel. Maar anderzijds bevat hij onheldere gedeeltes en ook stellingnamen die ik als ingewijde begrip en bestrijd, maar waar de leek niets aan heeft (er is geen direct macro-micro effect, maar het effect is van het cross-level interaction type: p. 24; exploreren is de type-I fout vermijden, testen is de type-II fout vermijden: p. 26; de afhankelijke variabele van het hiërarchisch lineaire model kan op beide niveaus gesitueerd zijn: p. 27).

De laatste drie hoofdstukken staan een beetje apart. Eerst wordt ingegaan op geaggregeerde groepskenmerken. Dan volgt een aparte behandeling van het multiniveau-design in evaluatieonderzoek. En ten slotte worden onder de titel 'Trends and Prospects' de nieuwste ontwikkelingen geschetst. Zoals Harry Hüttner in zijn dissertatie reeds constateerde, worden groepskenmerken in de meeste studies gemeten als analytische (lees: geaggregeerde) groepskenmerken. Globale en structurele groepskenmerken worden (te) zelden benut. Het gebruik van analytische groepskenmerken (zoals de gemiddelde socio-economische status) schept echter grote validiteitsproblemen, omdat men niet altijd weet wat men meet. Het kenmerk kan immers zowel verwijzen naar verschillende aspecten van de context (sociaal klimaat, materiële cultuur, netwerkstructuur), alsook naar verschillende contexten (*frog-ponds, subgroups, peer groups of zelfs: de sociaal-hiërarchische structuur van de maatschappij buiten de school*). Dit validiteitsprobleem kan niet genoeg worden beklemtoond.

Met de bespreking van evaluatieonderzoek verwijzen de auteurs eigenlijk naar het experimenteel design met voor- en nameting bij een experimentele groep en een controlegroep. Volgens hen is zo'n evaluatieonderzoek doordrongen van multiniveau-problemen. Bijvoorbeeld, alleen al het feit dat de experimentele stimuli aan groepen worden toegekend en dat de metingen bij individuen binnen groepen worden verricht, waarna gemiddelden van individuele scores tussen groepen worden vergeleken, houdt allerlei niveausprongen in. Bij deze stellingname zijn heel wat kritische kanttekeningen te plaatsen. Bijvoorbeeld, als men dit doordenkt, kan elk onderzoek waarin categorieën van variabelen worden opgenomen (seks: een eerste groep van vrouwen en een tweede groep van mannen) als multiniveau-onderzoek worden gezien en is er dus geen verschil meer tussen het (ondoordachte) invoeren van een categorische variabele en de (doordachte) vorming van groepen. Ik ben het ook niet met de auteurs eens dat bij de vergelijking van groepsgemiddelden in een experimenteel design de

binnengroepsvariaties buiten schot blijven, want die zijn wel degelijk in de formules van de variantieanalyses vervat. Het is ook jammer dat ter illustratie van het evaluatieonderzoek een voorbeeld van een tweestapsmodel wordt gegeven waarin de experimentele logica ver te zoeken is, zodat het argument niet echt kracht wordt bijgezet. Maar anderzijds is het wel moedig dat de auteurs van de gangbare behandelingen van het multiniveau-design hebben durven afwijken en een controversieel onderwerp hebben durven aansnijden.

Wat de toepassingen van multiniveau-onderzoek betreft, waarmee de auteurs hun inleidende tekst beëindigen, valt te constateren dat de meeste voorbeelden nog steeds uit de onderwijsresearch afkomstig zijn. Maar er is hoop. Langzamerhand vindt het multiniveau-verklaringsmodel zijn weg in nieuwe probleemgebieden, zoals stemgedrag, minderhedenonderzoek, gezondheidszorg, criminaliteit, druggebruik, alcoholisme, communicatieonderzoek en de validiteit van metingen in interviewsituaties. De aanbevelingen van de auteurs, dat meer gebruik moet worden gemaakt van structurele groepskenmerken, waarin sociometrische informatie wordt geïntegreerd, en dat er nood is aan geïntegreerde theorie, zijn ook de mijne.

Jacques Tacq

Nicholas T. Longford, **Random coefficient models**. Oxford University Press, Oxford 1993, XIV + 270 p.

In een jaar of vijftien is een geheel van methoden ontwikkeld voor een regressieachtige aanpak van gegevens met een geneste structuur, zoals individuen binnen groepen, dat in de sociale wetenschappen de naam *multiniveau-analyse* heeft gekregen, en dat wiskundig-statistisch gezien berust op regressiemodellen met stochastische coëfficiënten. Dit boek geeft in dit onderwerp een statistisch georiënteerde inleiding, met zeer veel aandacht voor verstandige modellering en concrete toepassingen.

De eerste zes hoofdstukken bevatten de onderdelen die langzamerhand zijn uitgekristalliseerd tot de kern van deze methodologie. In een eerste hoofdstuk geeft de auteur een uitstekende opwarmer (voor de onvoorbereide lezer wellicht een oprisser) over statistiek, matrixrekening, en geclusterde (of geneste) datastructuren. Wie wel matrixrekening en statistiek heeft bestudeerd maar dit niet meer allemaal op een rijtje heeft, leest hier hoe het ook al weer zat en leert ook over matrixdifferentiatie en het numeriek berekenen van Maximum Likelihood schatters. Ook bevat dit hoofdstuk verstandige basisprincipes aangaande de analyse van geclusterde waarnemingen. Dit hoofdstuk zet de toon voor het hele boek: het gaat er wiskundig tamelijk fors aan toe (de lezer moet niet bang zijn voor log-likelihoods of informatiematrices), maar wel vanuit een benadering die volstrekt gericht is op het verkrijgen van inzicht. Vervolgens worden in twee theoriehoofdstukken en twee toepassingshoofdstukken eerst modellen met een stochastisch intercept, en daarna modellen met stochastische regressiecoëfficiënten ('random slopes') behandeld. Het 'kern' deel van het boek wordt afgerond met een hoofdstuk over situaties met drie of meer niveaus.

Het boek besluit met twee hoofdstukken over speciale onderwerpen: het ene over factoranalyse en structurele vergelijkingen, het andere over modellen voor niet-normaal verdeelde afhankelijk variabelen, met dichotome variabelen als belangrijkste speciaal geval.

De nadruk van het boek ligt ten eerste op de statistische modellering aan de hand van theorie en veel uitgebreide voorbeelden, ten tweede op het numeriek berekenen van de statistische schatters. Wat betreft beide aspecten is het het tegengestelde van een kookboek en wordt van

de lezer de nodige wiskundige rijpheid verwacht (alles gaat in matrixnotatie). Het 'hoe doe je het nu' is gericht op inzicht en principes, niet op programma's. Op pagina 189 zegt de auteur dat het beter is ('gives the qualified analyst a distinct advantage') deze methoden zelf te programmeren in nieuwe generatie computertalen dan gebruik te maken van bestaande software, inclusief het door Longford zelf ontwikkelde VARCL (Longford 1993). Aan programmatuur wordt verder dan ook geen aandacht besteed.

De meeste voorbeelden worden in diverse hoofdstukken steeds verder uitgewerkt, met veel aandacht voor alternatieve analysemogelijkheden en controle van modelveronderstellingen. Longfords werkzaamheden op de Amerikaanse Educational Testing Service leiden tot een uitgebreide analyse over het voorspellend karakter van 'undergraduate'-cijfers voor de resultaten in 'graduate school'. In dit voorbeeld komt de meerwaarde van multiniveau-analyse ten opzichte van gewone regressieanalyse goed tot uiting. De voorbeelden strekken zich uit tot de analyse van longitudinale en van multivariate data, waar een twee-niveaustructuur wordt verkregen door de data te beschouwen als meettijdstippen, respectievelijk variabelen, genest binnen respondenten. Dit illustreert de grondigheid en nuttigheid van het boek: deze toepassingen passen binnen het basisparadigma van het gebruik van random coëfficiëntmodellen voor multiniveau-analyse, maar de mogelijkheid ervan is nog niet bekend bij alle gebruikers van deze methoden.

Heel nuttig, en een onderscheid ten opzichte van andere leerboeken op dit terrein, is de aandacht die het boek besteedt aan de hoeveelheid informatie die de data bevatten over de diverse statistische parameters. Die hoeveelheid is soms klein, wat impliceert dat de gebruiker niet teveel moet verwachten van de nauwkeurigheid waarmee 'moeilijke' parameters zoals varianties en covarianties van 'random slopes' kunnen worden geschat bij onderzoeksoptellingen met een tamelijk gering aantal groepen. Soms is Longford daarbij naar mijn smaak te pessimistisch, bijvoorbeeld bij zijn suggestie op pagina 244 (zie ook pagina's 167 en 168) dat het ontdekken van 'random slopes' bij multiniveau-logistische streekproefomvang; Hox en De Leeuw (zie bijvoorbeeld Hox 1994, par. 4.2) horen bij de onderzoekers die een voorbeeld geven van het tegendeel. De opmerkingen over de hoeveelheid informatie in data worden voor het geval van modellen met stochastisch intercept (par. 2.9) uitgewerkt in praktisch bruikbare conclusies over benodigde steekproefomvang.

Een nadeel van de tamelijk wiskundige stijl is dat conclusies en belangrijke praktische implicaties soms verstopt zitten in de formules en wiskundige terminologie. Een voorbeeld is paragraaf 2.8, waarin de verschillen worden besproken tussen enerzijds multiniveau-analyse en anderzijds gewone regressieanalyse op de gedisaggregeerde gegevens. Daar wordt bijvoorbeeld gesteld dat als de design-matrices voor alle groepen gelijk zijn, de door deze twee methoden geleverde schatters identiek zijn, waarna een stukje volgt over de onzuiverheid ('bias') van de standaardfouten. De lezer moet zelf maar concluderen dat gewone regressieanalyse ook in deze 'mooie' situatie wel goede schattingen, maar misleidende standaardfouten oplevert (tenzij de intraklasse correlatiecoëfficiënt nul is). De hierop gebaseerde toetsen zijn dus volstrekt niet te vertrouwen. Een ander voorbeeld is het bespreken, zonder tam-tam, op pagina's 125 en 126 van hoe je tijdreeksanalyse kunt doen (voor betrekkelijk eenvoudige tijdreeksmodellen) simultaan op meerdere niet al te lange observaties van tijdreeksen.

De aandacht die in de hoofdstukken 7 en 8 wordt besteed aan respectievelijk structurele vergelijkingsmodellen en niet-normale afhankelijke variabelen, is zowel een kracht als een zwakte. Dit zijn belangrijke onderwerpen, waarnaar de laatste tijd veel onderzoek wordt gedaan dat vooral in zeer specialistische literatuur terecht komt. Het is handig om daarover een aantal fundamentele zaken, uitgebreid toegelicht via voorbeelden, in dit leerboek bijeen te hebben. Het nadeel van het opnemen van dergelijke niet-uitgekristalliseerde onderwerpen in een boek is echter dat de behandeling heel snel verouderd is.

Wat zijn de sterke punten van dit boek in vergelijking met andere leerboeken over multiviewanalyse (Bryk en Raudenbush 1992; Hox 1994; Goldstein 1995)? Voor wie de rekenalgoritmen zelf wil programmeren of er inzicht in wil hebben, is dit boek een must (met de kanttekening dat het zich vooral, maar niet uitsluitend, richt op Fisher scoring). Het is ook aan te raden voor wie de wiskundig statistische achtergrond van de multiviewanalyse wil kennen; wat dat betreft is het uitgebreider, en didactisch beter dan de concurrenten. Allerlei formules, bijvoorbeeld van standaardfouten van geschatte varianties, kunnen de gebruiker van pas komen en zijn moeilijk elders te vinden. Ook het verband met de meer algemene statistische literatuur wordt in de 'Bibliographical notes' aan het eind van elk hoofdstuk goed aangegeven (al mis ik de dwarsverbanden met de econometrische literatuur). De uitvoerige, zeer diverse en soms complexe voorbeelden zijn een schatkamer waarin meer aan te treffen is dan in andere boeken, die zich concentreren op een standaardbehandeling volgens de beschikbare software. De prijs van dit alles is duidelijk: voor het ontsluiten van de schatkamer is enige wiskundige rijpheid en statistische bagage nodig, of op zijn minst het vermogen om zich niet te laten afschrikken door het bos der formules.

Referenties

- Bryk, A.S., en S.W. Raudenbush (1992), *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage, Newbury Park etc.
- Goldstein, H. (1995), *Multilevel Statistical Methods* (2nd ed.). Edward Arnold, Londen.
- Hox, J.J. (1994), *Applied Multilevel Analysis*. TT Publikaties, Amsterdam.
- Longford, N.T. (1993), *VARCL: Software for variance Component Analysis of Data with Nested Random Effects*. Manual. Groningen, iec ProGAMMA.

Tom A.B. Snijders

A.S. Bryk en S.W. Raudenbush, **Hierarchical Linear Models: applications and data analysis methods**. Sage Publications, Newbury Park, Londen/New Delhi 1992, XVI + 265 p.

Bryk en Raudenbush kunnen met Goldstein en Longford gezien worden als de grondleggers van het statistisch multilevel-model. Niet omdat zij als eersten de theorie formuleerden, maar meer nog door hun werk aan het toepasbaar maken van de theorie en aan de verfijning ervan. Het boek van Bryk en Raudenbush kan beter dan Goldsteins *Multilevel Statistical Models of Longfords Random Coefficient Models* gezien worden als inleidend.

De basis van het boek bestaat na een kort historisch overzicht en een heldere motivering uit drie kernhoofdstukken: een algehele inleiding in de logica van hiërarchisch lineaire modellen (HLM) (hoofdstuk 2); schatten en toetsen bij HLM (hoofdstuk 3); en de adequaatheid van HLM (hoofdstuk 9). De overige hoofdstukken zijn verfijningen, uitbreidingen en illustraties.

De logica van HLM start zeer basaal met de principes van lineaire regressie (wat is een intercept, een helling en een residu) en drie pagina's verderop zitten we al midden in HLM, geholpen door een groot aantal grafieken. Vervolgens wordt uitgelegd hoe HLM gezien kan worden als een algemenere (namelijk met minder restricties) formulering van het lineaire model zoals dat gebruikt wordt in variantieanalyse Covariantieanalyse en regressieanalyse. Het bijzondere zit hem in het modelleren van effecten op meerdere niveaus (leerlingen en

klassen, wijkbewoners en buurten, herhaalde metingen en individuen) tegelijkertijd. De auteurs kiezen ervoor – om begrijpelijke didactische redenen – het model steeds uiteen te leggen in submodellen voor de diverse niveaus van de hiërarchie. Naar mijn smaak wordt niet expliciet genoeg gewezen op het type steekproef dat aanleiding geeft tot het gebruik van HLM (namelijk de meertrapssteekproef) en wordt het probleem van de locatie van het intercept bij niet-parallelle regressielijnen wel heel erg uitgemolken (in de paragrafen over centre-nen van variabelen) zonder dat de lezer nu meteen weet wat hij met deze kennis moet. Voor de lezer is de voornaamste opgave consequent de notatie te blijven volgen.

Hoofdstuk 3 bevat de statistische theorie betreffende schatten en toetsen bij HLM. Voor de ongeofende sociale wetenschapper bevat dit hoofdstuk de taaiste kost, namelijk de nodige matrices, inversen en wat al niet meer. Misschien schrikt dit hoofdstuk menigeen zodanig af dat ze het boek terzijde zullen leggen, maar het aardige is dat wat in dit hoofdstuk moeizaam wordt uitgelegd in de daarop volgende hoofdstukken helder wordt geïllustreerd. Maar zelfs zonder de gehele matrixnotatie te kunnen volgen zal de lezer begrijpen:

- 1) dat de twee soorten variantie hebben (steekproevenvariantie en parametervariantie);
- 2) dat het vervelend is dat het design meestal opgebalanceerd is en dat daar statistisch het nodige aan te verhapstukken valt (niet evenveel leerlingen per school bijvoorbeeld);
- 3) dat daarbij betrouwbaarheidsgewichten worden gebruikt;
- 4) dat de regressiecoëfficiënten bijbehorende standaardfouten hebben waarvan de grootte afhangt van op welk niveau het submodel gesitueerd is;
- 5) dat voor de variantiecomponenten (hoeveel variantie zit er tussen wijken?) andere toetsen nodig zijn dan voor de regressiecoëfficiënten;
- 6) dat we op moeten passen als er maar weinig scholen, buurten en dergelijke zijn.

Hoofdstuk 4 doet hoofdstukken 2 en 3 nog eens dunnetjes over door een en ander te bespreken aan de hand van een voorbeeld en de lezer die in de matrices verdwaald was geraakt, zal hier snel de draad weer kunnen oppakken.

Hoofdstuk 9 ten slotte is te ver achterin het boek geplaatst: het behandelt namelijk de assumpties van HLM, hoe je kunt nagaan of en in welke mate die geschonden worden en wat je daar dan vervolgens aan kunt doen. Bovendien worden problemen met betrekking tot misspecificatie uitvoerig behandeld. Ongelukkigerwijze wordt misspecificatie alleen opgevat in termen van vergeten variabelen en niet als: is er een belangrijk niveau in de hiërarchie over het hoofd gezien?

De hoofdstukken 5 tot en met 7 zijn vervolgens toepassingen van het twee-niveaumodel op: organisatieonderzoek (leden binnen organisaties, leerlingen binnen scholen); het meten van veranderingen (de tijdstippen waarop gemeten is, worden hier gezien als genest binnen het individu); en meta-analyse (onderzoeken in de review worden gezien als steekproefelementen uit een populatie en de primaire onderzoekseenheden worden gezien als genest binnen elk onderzoek). De lezer doet er wel verstandig aan andere boeken over het meten van veranderingen, respectievelijk meta-analyse te raadplegen alvorens met HLM in deze situaties aan de slag te gaan.

Hoofdstuk 8 bevat uitbreiding van het twee niveaumodel naar drie niveaus (leerlingen in klassen in scholen, tijdstippen binnen individuen in organisaties etcetera). De prachtige paragraaf 'A measurement model at level 1' over hoe je uit percepties van individuen variabelen op organisatieniveau kunt destilleren, had wel een belangrijkere plaats in het boek mogen krijgen.

Voor de statistische 'die-hards' is er tenslotte nog een hoofdstuk 10 met technische details.

Het boek is didactisch zeer adequaat opgezet met 'advance organizers', vele en heldere voorbeelden en samenvattingen aan het eind van elk hoofdstuk. In combinatie met het software-pakket HLM3 is het bijzonder leerzaam, daar datasets behandeld in het boek door de

gebruikers van dat pakket ook geanalyseerd kunnen worden.

Wathetboekallemaaniendoet(logistischeHLM,kruisclassificaties,multivariatetoepassing,etc.)istevindenindeboekenvanLongforden/ofGoldstein.Hetboekbehandeltechterprecieswathetsoftware-pakketHLMkan.Hetniveauvanhetboekschatikinop:debestede beste inleiding voor sociale wetenschappers die met multilevel-modellen aan de slag moeten en geschikt voor postdoctoraal onderwijs (wellicht) voorzichinM&Tspecialiserende predoctoraal studenten.

Roel J. Bosker

R. Prosser, J. Rasbash, H. Goldstein, **ML3 Software for Three-level Analysis, Users' Guide for V.2**. Institute of Education, University of London, januari 1991, 142 p. (1).

R. Prosser, J. Rasbash, H. Goldstein, **Data Analysis with ML3**, Institute of Education, University of London, juli 1991, 121 p. (2); **ML3 V2.2 Manual Supplement, New and Amended commands**. Multilevel Models Project, Institute of Education, University of London, november 1991, 10 p. (3).

G. Woodhouse (red.), **A Guide to ML3 for New Users, second edition**. Multilevel Models Project, Institute of Education, University of London, juli 1993, 124 p. (4).

J. Rasbash, **ML3-E Version 2.3 Manual Supplement**. Multilevel Models Project, Institute of Education, University of London, juli 1993, 11 p. (5).

De uitdaging om samen met Rob Eisinga een bijdrage te leveren aan dit themanummer van de Sociologische Gids over multiniveau-analyse noodzaakte me op korte termijn enig idee te krijgen van de theorie van de multiniveau-analyse, een programma te kiezen en er snel in thuis te raken. Het moest een logistische analyse worden op een groot databestand ($n = 244.050$). Na het lezen van Bryk en Raudenbush's *Hierarchical Linear Models* (1992) had ik de beschikking over de manuals van HLM, VARCL en ML3 (hierboven 1) en van het laatste programma tevens de hierboven genoemde *Guide to ML3 for New Users* (4). Hoewel gewaarschuwd dat de 'ease of learning' van de drie programma's respectievelijk 'easy, easy and slow' is (*Newsletter 4* (2)), viel de keuze toch op ML3. Met HLM zijn blijkens dezelfde *Newsletter* geen binaire data te analyseren en in de *Guide to ML3*, naar de titel doet vermoeden begaan met problemen van 'new users', was in ieder geval een voorbeeld opgenomen van een logistische analyse.

Toen het artikel er bijna lag, werd ik door de redactie verzocht mijn ervaringen met ML3 en de programmadocumentatie op schrift te stellen. Na enige aarzeling te hebben toegezegd, stond ik voor het probleem: wat hoort precies tot de programmadocumentatie? Want het multiniveau denken, zo had ik al snel door, strekt zich uit over vele gebieden (logistische analyse, multivariate respons, primaire netwerken, herhaalde metingen, crosslevel-analyse, covariantiestructuuranalyse, etc.) en ik wist niet of ik over de relevante documentatie beschikte. Het verzoek aan het Institute of Education te Londen om de programmadocumentatie van ML3 toe te zenden, werd beantwoord met toezending van de vijf bovengenoemde documenten. Ik heb de indruk dat voor bepaalde toepassingen er meer nuttige documenten zijn. Zo zou ik zelf voor de logistische analyse met binaire respons hiermee hebben kunnen volstaan, maar in een later stadium kreeg ik *A Guide to ML3 Macros, Multilevel Binary Response Logistic Models* (juni 1993) onder ogen dat wel gedeeltelijk, maar niet helemaal overlappende informatie bevat. Men moet er ook rekening mee houden dat sommige toepassingen in deze documenten niet aan bod komen. Als wij bijvoorbeeld een polytome responsvariabele hadden gehad (op de programmadiskette zijn de macro's daarvoor voorhanden), dan hadden we niet genoeg gehad aan deze documenten. In het algemeen is het voor de gebruiker lastig er-

achter te komen welke stukken voor hem of haar relevant zijn. Het beste is de *Newsletters* te raadplegen of rechtstreeks contact op te nemen met het Institute of Education in Londen. Overigens is dit ongemak begrijpelijk voor een gebied dat zo in ontwikkeling is.

Naast 1 (users' guide en manual), zijn 2 en 4 belangrijke informatiebronnen; 3 en 5 zijn slechts kleine aanvullingen. Ik beperk me tot 1, 2 en 4.

De users' guide en manual (1) begint met een inleiding, een stukje pittige abstracte theorie (voor de beginneling niet direct iets om mee te beginnen); daarna een toelichting op de commando's gegroepeerd achtereenvolgens rond data-input (input, checks), utilities (rekenen, dummyficeren, sorteren, construeren van contextvariabelen, wegschrijven), grafische mogelijkheden (stem and leaf, boxplots, plots, groepsgewijze plots), statistics (beschrijvende statistics, tabellen, correlaties), modelspecificatie (benoeming niveaus, fixed en random parameterspecificatie), runnen en output (schatting, berekenen voorspelde waarden en residuen). Daarna volgen voorbeelden: eerst een gewoon voorbeeld van een variantiecomponentenmodel en een model met random effecten, daarna voorbeelden van groeicurveanalyse, analyse met herhaalde metingen, analyse van proporties en drie-niveaugroeianalyse.

Over het algemeen vind ik achteraf dit document helder geschreven, beknopt en goed hanteerbaar tijdens het werk; goed voor degene dus die thuis is in multilevel-denken en het programma. Om er in thuis te geraken, is het naar mijn mening te beknopt. Ik zou willen pleiten voor een veel uitgebreidere toelichting van theorie, commando's en voorbeelden met daarnaast een korte handleiding. De command reference en de command index zijn bovendien niet volledig. Zo mis ik in de index de commando's histogram, summary, put, count, sum, mlsum, cumulate, setting 2. De laatste zes zijn ook niet in het manual te vinden (onder het help-commando komen de eerste vijf genoemde commando's wel voor). Het stimuleert in ieder geval een vermoedelijk commando uit te proberen en het aardige is dat het programma soms doet wat men wil. Als men eenmaal een beetje thuis is in de syntax, is het verder prettig om met ML3 te werken. Het kent wel wat gebruikersonvriendelijkheden. Zo moet men zelf de plaatsing van variabelen in kolommen bijhouden. Als men bijvoorbeeld een som wil hebben van twee variabelen, gelabeld met 'X1' en 'X2', dan lukt niet het commando: calculate 'som' = 'X1' + 'X2'. Men moet eerst opzoeken welke kolom leeg is (anders wordt een kolom zonder warning overschreven) en dan de commando's geven (zeg dat kolom 10 leeg is): calculate c10 = 'X1' + 'X2' name c10 'som' (of: add 'X1' 'X2' c10 name c10 'som'). Foutmeldingen zijn wat kort en soms onvolledig. Zo leidt het commando PRINT 'SEX' tot de mededeling 'wrong parameter', bedoelend dat de sluitquote vergeten is, terwijl het commando PRINT 'SEX' 'AGE' een print van alleen SEX geeft en verder niets zegt over AGE. Mooier dan syntactische foutmeldingen zouden natuurlijk warnings zijn omtrent wat er fout kan zijn aan het model en hints tot verbetering. Ik ben die, dacht ik, niet tegengekomen. Ook is het irritant dat bij het gebruik van veel predictoren niet alle predictoren op het setting-screen (het schermoverzicht van de modelkeuze) passen.

Van de guides 2 en 4 is 2 het oudst (1991) en daarmee soms gedateerd. Het behandelt voorbeelden van verschillende toepassingsgebieden. Ze staan niet geordend naar oplopende moeilijkheidsgraad. Ook datasets worden niet op de ML3 software-diskette geleverd. Dat neemt niet weg dat het nog alleszins de moeite waard is om het te lezen. Een inleidend voorbeeld ontbreekt evenwel. Hoewel de titel er niet toe uitnodigt, zou ik de beginneling toch aanraden te beginnen met hoofdstuk 5, 'A multilevel bivariate model', paragraaf 1. Het bevat simpele databeschrijvende en univariate analyses met het variantiecomponentenmodel en het random coëfficiëntmodel.

De meest recente guide is document 4. Het is didactisch beter dan de vorige (2). Het eerste hoofdstuk (bijna de helft van het geheel) heeft een duidelijk inleidend karakter. Daarna worden de multivariate multilevel-analyse, de logistische multilevel-analyse en de herhaalde

metingen multilevel-analyse geïllustreerd. Ook worden op de programmadiskette alle datasets bijgeleverd zodat men alles zelf nog eens kan nadoen. Het eerste inleidend hoofdstuk begint simpel, maar wordt beslist moeilijker. Ik vraag me af of in zo'n inleidend hoofdstuk niet iets gezegd moet worden over de mogelijkheden van ML3 om de data te checken en er een eerste beschrijving van te geven. Ook zou ik in een inleiding bij het weergeven van modellen liever binnen- en tussenunits vergelijkingen zien, zoals in het manual.

Samenvattend, een binnenkomer in het veld van de multilevel modeling, wellicht argeloos denkend dat het wel zal gaan om zoiets als leerling-klasschoolniveaus, zal na kennismaking met ML3 en de bijbehorende documentatie verbijsterd staan van de uitgestrektheid van het veld van mogelijke toepassingen. In lichte mate hierdoor gedesoriënteerd zal hij of zij behoefte hebben aan een simpele tekst over een simpel probleem. Die teksten zijn er wel, maar ik ben bang dat ze voor de gemiddelde onderzoeker nog niet eenvoudig genoeg zijn. Het programma en de documentatie zijn afgestemd op de technisch geschoolde onderzoeker. Die zal er een nieuwe vorm van methodologisch denken in aantreffen met een voortreffelijk programma waarmee dat allemaal op een breed vlak is toe te passen.

J. Lammers