

Reactie op Van Kolverschoten en Kerstholt

Harry B.G. Ganzeboom en Paul M. de Graaf

In de bijdragen van Van Kolverschoten en Kerstholt worden we geconfronteerd met kritiek op enige recente publicaties van (voornamelijk) onze hand. De kritiek is nogal opgewonden geformuleerd en voor een groot deel onzakelijk van inhoud. Zij berust helaas in beide gevallen op slecht lezen en elementair onbegrip van gangbare onderzoeksprocedures. We zullen slechts op de zakelijke punten van beide stukken ingaan. De lezer van *Mens en Maatschappij* oordele zelf over de rest.

Van Kolverschoten heeft nog het meest concrete punt van kritiek in handen, namelijk de op het eerste gezicht merkwaardige observatie dat in de in 'Leefstijlen in Nederland' (Ganzeboom, 1988) gebruikte steekproef uit de bevolking van de stad Utrecht de 'beroepsstatus' hoger zou correleren met de opleiding van de partner van de respondent dan met de eigen opleiding. Om zijn observatie te verklaren bouwt hij een voor ons ondoorzichtige redenering op waarin verondersteld wordt dat de ondervraagden uitsluitend uit niet-werkenden (huisvrouwen en 'huismannen') zouden bestaan.

De betreffende gegevens zijn in 1986 verzameld door De Graaf in het kader van zijn promotie-onderzoek (De Graaf, 1987). Bij de betreffende steekproefprocedure werd de eerst aangetroffen persoon op een adres ondervraagd. Deze procedure brengt met zich mee dat de steekproef als zodanig niet representatief is voor de bevolking van de stad Utrecht. De Graaf vond in zijn steekproef van 342 respondenten tussen de 25 en 64 jaar 217 (66%) vrouwen. Om dit te corrigeren paste hij – zoals bij dergelijke steekproefprocedures gebruikelijk is – een dusdanige weging toe, dat er zich in de geanalyseerde steekproef weer 50% mannen en 50% vrouwen bevonden. Van Kolverschoten meent dat hiermee een grote fout is gemaakt. De mannen in de geanalyseerde steekproef zouden volgens hem uitsluitend uit niet-werkenden bestaan en door de herweging nog eens extra oververtegenwoordigd worden.

Hoe anders zijn echter de feiten. Enige elementaire cijfers uit het betreffende databestand (wij hadden Van Kolverschoten die gaarne op aanvraag verstrekt) leert dat het met de veronderstelde oververtegenwoordiging van niet-werkende mannen nogal meevalt. In het onderzoek is gevraagd of de respondent werkt of niet. Het blijkt dat 64% van de ondervraagde mannen en 39% van de ondervraagde vrouwen betaalde beroepsarbeid

uitoefent. Verder bleek 11% van de mannen werkloos, tegen 3% van de vrouwen. In totaal blijkt dus 75% van de mannen in de steekproef te behoren tot de beroepsbevolking. De overige 25% van de mannen was studierend, arbeidsongeschikt of vervroegd gepensioneerd. Volgens berekeningen naar het Statistisch Zakboek van 1986 (p. 60 en 158) behoorde in dat jaar 87% van de Nederlandse mannen tot de beroepsbevolking. Helaas hebben we niet de beschikking over Utrechtse statistieken op dit gebied en we weten dus niet in hoeverre de 75% in de steekproef afwijkt van de Utrechtse populatie. Hoe dan ook, de verschillen zijn niet dramatisch.

Cruciaal voor het begrijpen van het merkwaardige correlatiepatroon is echter een niet door Van Kolverschoten (maar wel door Ganzeboom) genoemd feit, namelijk dat in 'Leefstijlen in Nederland' de betreffende beroepsstatus niet betrekking heeft op de *respondent*, maar op het *huishouden* waarvan de respondent deel uitmaakt. Ingeval van niet-werkende respondenten werd de score van beroepsstatus ontleend aan de partner (Ganzeboom, 1988, p. 80). Deze procedure is ingegeven door het feit dat in de betreffende analyse van leefstijlkenmerken (deels huishoudkenmerken) het beroepsniveau van het huishouden nu eenmaal het relevantere gegeven is. Het is deze keuze die ervoor zorgt dat de correlatie met de opleiding van partner van de respondent wat hoger is dan met de opleiding van de respondent zelf. Definieert men het beroep uitsluitend als het beroep van de respondent, dan is de correlatie met de opleiding van de respondent inderdaad de hoogste.

Voor de duidelijkheid geven we nog twee punten aan waarop het gebrekkig lezen van de zijde van Van Kolverschoten tot misverstanden zou kunnen leiden. Ten eerste worden – anders dan Van Kolverschoten meldt – in 'Cultuurdeelname in Nederland' (Ganzeboom, 1989) deze zelfde gegevens *niet* gebruikt, het gaat daar om een ander, eerder verzameld databestand. Ten tweede wordt in het proefschrift van De Graaf weliswaar wel hetzelfde bestand gebruikt maar heeft de analyse daar in het geheel geen betrekking op de beroepsgegevens. Het betreffende hoofdstuk gaat over de relatie tussen het culturele en materiële klimaat in het milieu van herkomst en het uiteindelijk behaalde opleidingsniveau.

We menen dat Van Kolverschotens opmerking dat 'de dissertatie van De Graaf (1987) op ondeugdelijke data berust' kwaadaardig overdreven is. Het is niet alleen zo dat De Graaf in het geheel geen kenmerken analyseert die tot stand zijn gekomen na het beëindigen van de schoolloopbaan, maar ook berusten zijn analyses zeker niet uitsluitend op de Utrechtse gegevens. In de dissertatie wordt een aantal verschillende databronnen aangewend, waaronder gegevens uit het door het CBS Leefsituatieonderzoek van 1977 en de uit Groningen afkomstige longitudinale gegevens over schoolloopbaan van leerlingen in het basisonderwijs. Het is aan Van Kolverschoten te argumenteren dat de bevinding van De Graaf – dat het niet de financiële situatie maar de culturele dispositie van de ouders is die de schoolloopbaan van hun kinderen stuurt – kan worden ontzenuwd door de procedure die gevolgd is bij het verzamelen van de Utrechtse gegevens.

Heeft Van Kolverschoten nog een concreet punt in handen, Kerstholt legt zich uitsluitend toe op ongefundeerde verdachtmakingen. Omdat zijn voorstelling van zaken tot misverstanden leidt, vatten we aard en bevindingen van de gewraakte drie artikelen nog

even samen. In alle drie de gevallen gaat het om een analyse van op individueel niveau samengevoegde nationaal representatieve databestanden (in aantal variërend tussen 10 en 14), waarin de coderingen van de betreffende kenmerken (beroepen en opleidingen van respondenten en hun vader) zorgvuldig met elkaar vergelijkbaar werden gemaakt. In elk van de drie artikelen wordt – in verband met beschikbaarheid van variabelen – een verschillende deelverzameling van deze databestanden gebruikt. Ganzeboom e.a. (1987) analyseren de intergenerationele beroepsmobiliteit van mannen voor 10 surveys gehouden tussen 1970 en 1985. Zij constateren een vermindering van de relevante associatieparameters met ca. 25% in de bestudeerde periode van 15 jaar. Ganzeboom en De Graaf (1989a) analyseren de intergenerationele opleidingsmobiliteit van mannen en vrouwen volgens 14 surveys en betrekken hun historische vergelijking op geboortecohorten. De voordelen van cohortvergelijkingen boven surveyvergelijkingen en de betekenis van opleidingsmobiliteit boven beroepsmobiliteit (waaronder het feit dat bij opleidingsmobiliteit cohortvergelijking mogelijk is en bij beroepsmobiliteit niet) worden in dit artikel uitvoerig uit de doeken gedaan. De belangrijkste bevinding is dat de relevante associatieparameters met ca. 40% teruglopen in de geanalyseerde periode van 60 jaar. In Ganzeboom en De Graaf (1989b) wordt hetzelfde probleem met 10 databestanden en via cohortvergelijking aangepakt, maar nu multivariaat, door namelijk de invloed van vaders beroep en vaders opleiding op het uiteindelijk behaalde opleidingsniveau te analyseren. Ook dan luidt de conclusie dat de intergenerationele associaties in de geanalyseerde periode aanzienlijk zijn teruggelopen.

Kerstholt gaat niet in op onze inhoudelijke bevindingen maar heeft kennelijk op voorhand bezwaar tegen het samenvoegen van gegevens uit verschillende databestanden. Hij beroept zich daarbij op de discussie tussen De Leeuw en Dronkers in 1978. Hij heeft kennelijk niet gezien dat het in die discussie om een volstrekt andere zaak ging. Dronkers had (in navolging van Jencks) voor zijn analyse correlaties uit verschillende onderzoeken in één matrix gecombineerd en De Leeuws kritiek bevatte als belangrijk punt dat het daardoor goed mogelijk is dat de ontstane correlatie-matrix statistische anomalieën vertoont. Dronkers gebruikte zowel gepubliceerde als zelf berekende correlaties en ontleende bovendien gegevens uit steekproeven uit uitdrukkelijk verschillende populaties, zoals schoolgaanden, de volwassen bevolking en recruten (en leende in één geval zelfs een correlatie uit buitenlands onderzoek). In ons geval gaat het echter in het geheel niet om het combineren van aan andere publikaties ontleende correlaties in één matrix, maar om het bijeenbrengen en tot in detail vergelijkbaar coderen van individuele gegevens, waarbij de databestanden alle afkomstig zijn uit landelijk representatieve bevolkingssurveys. De populatie is steeds de Nederlandse (beroeps)bevolking en de kenmerken waarover het gaat, beroep en onderwijs, zijn steeds op een eendere manier gemeten en vervolgens identiek gehercodeerd.

Natuurlijk is het zo dat ook dan gegevens uit verschillende surveys onvergelijkbaarheden kunnen blijven behouden. De door Kerstholt genoemde mogelijkheid van verschillende steekproefprocedures (huishoud- versus individuensteekproef) is er een van.

Anders dan Kerstholt het voorstelt is de vraag naar de validiteit van de gevolgde procedure, niet of er op de vergelijkbaarheid van de gegevens überhaupt iets is aan te merken, maar of aan te nemen valt dat verschillen in procedures tussen de gebruikte surveys

op de een of andere manier verantwoordelijk zouden kunnen zijn voor de door ons gevonden trends. We hebben in Kerstholt's bijdrage geen enkel argument kunnen vinden dat laat zien dat de centrale bevinding, namelijk die van groeiende openheid van de Nederlandse samenleving, op zulke verschillen berust. Men kan eenvoudig de voorwaarden specificeren waaronder zulks het geval zou kunnen zijn. Vertekening van de resultaten door onvergelijkbaarheid tussen de gebruikte survey zou optreden wanneer aan alle volgende voorwaarden is voldaan:

- a. de betreffende onvergelijkbaarheid is van invloed op de hoogte van de geschatte parameters;
- b. voor wat betreft de surveyvergelijkingen: de betreffende onvergelijkbaarheid is gecorreleerd met het surveyjaar, dat wil zeggen treedt relatief vaker op in de recentere, dan wel in oudere surveys;
- c. voor wat betreft de cohortvergelijkingen: het optreden van de betreffende onvergelijkbaarheid is gecorreleerd met het geboortejaar van de respondent.

Gegeven deze drie voorwaarden valt gemakkelijk in te zien waarom het niet aannemelijk is dat onze conclusie van groei naar openheid op rekening van mogelijke onvergelijkbaarheden in de surveyprocedures kan worden geschreven. Ten eerste is het niet duidelijk waarom de door Kerstholt aangevoerde onvergelijkbaarheid (individueel- versus huishoudsteekproef) invloed op de geschatte parameters zou hebben. Ten tweede valt niet in te zien waarom de betreffende steekproefprocedure gecorreleerd zou zijn met het tijdstip van survey. Ten derde is het wel zeer onaannemelijk dat een uit het surveydesign voortkomende onvergelijkbaarheid gecorreleerd zou zijn met het geboortejaar van de respondenten, op basis waarvan de vergeleken cohorten zijn samengesteld. In elk cohort zitten immers personen uit verschillende surveys, waardoor het effect van de surveyeigenaardigheden gerandomiseerd wordt.

Bezien vanuit de kwaliteit van de gegevens heeft het bijeenbrengen van data uit verschillende bestanden grote voordelen boven het analyseren van een enkelvoudig bestand. Hierdoor overwint men immers de specifieke beperkingen die aan een enkel bestand verbonden zijn. Het grootste voordeel ligt echter in de winst aan onderscheidend vermogen dat ermee behaald wordt. In een vierde artikel uit deze reeks (De Graaf & Ganzeboom, 1990) laten we nauwkeurig zien hoeveel baat vergelijkend mobiliteitsonderzoek heeft bij steekproeftotalen zoals zij door het poolen van verschillende datasets ontstaan. Hoe cruciaal dat is wordt vooral aangetoond door berekeningen die laten zien dat de door ons gevonden verschuivingen (die toch substantieel van omvang zijn) pas bij steekproefomvang van 1500 (mannen) à 3000 (vrouwen) significant zijn. Dat zijn steekproefgrootten die de omvang van elk afzonderlijk bestand te boven gaan.

Kerstholt's opmerkingen bevatten tenslotte twee voor ons duistere suggesties. Zo wil hij toetsen of de verschillende bestanden zouden kunnen worden beschouwd als aselechte steekproeven uit dezelfde populatie. Dat zou merkwaardig zijn, want onze artikelen tonen aan dat dit *juist niet* het geval is: de centrale bevinding is dat de onderliggende populatie in de loop der tijd veranderd is. Vervolgens suggereert Kerstholt nog om dummy-variabelen voor elk survey te introduceren in de analyse. Dit zou weliswaar voor de cohortanalyse een doenlijke procedure zijn geweest, maar toch niet voor de surveyver-

gelijking: wat valt er nog te vergelijken als de surveys elk een afzonderlijke dummy hadden gekregen?

De bekritiseerde drie artikelen bevatten onzes inziens belangwekkende resultaten over historische veranderingen in de openheid van de Nederlandse samenleving. Het is jammer dat deze resultaten besmet worden met irrelevante verdachtmakingen. Wie informatieve kritiek wil leveren, is gehouden om te laten zien dat een onderzoeksbevinding het resultaat is van onjuiste procedures.

Literatuur

- Ganzeboom, Harry B.G.; Luijckx, Ruud; Dessens, Jos; Graaf, Paul M. de; Graaf, Nan Dirk de; Jansen, Wim; Ultee, Wout C. (1987). 'Intergenerationele klassenmobiliteit in Nederland tussen 1970 en 1985.' *Mens en Maatschappij* (62-1), p. 17-43.
- Ganzeboom, Harry (1988). *Leefstijlen in Nederland*. Samsom, Alphen aan de Rijn. (Cahier Sociaal en Cultureel Planbureau #60).
- Ganzeboom, Harry (1989). *Cultuurdeelname in Nederland*. Van Gorcum, Assen.
- Ganzeboom, H.B.G. & P.M. de Graaf (1989a). 'Intergenerationele opleidingsmobiliteit in Nederland van geboortecohorten 1891-1960'. *Sociale Wetenschappen* (32-4), p. 263-278.
- Ganzeboom, H.B.G. & P.M. de Graaf (1989b). 'Veranderingen van onderwijskansen in Nederland tussen 1900 en 1960.' P. 58-78 in I. Gadourek & J.L. Peschar (Red.), *De open samenleving*, Van Loghum Slaterus, Deventer. (Boekaflevering Mens en Maatschappij).
- Graaf, P.M. de (1987). *De invloed van financiële en culturele hulpbronnen in schoolloopbanen*. ITS (dissertatie RUU), Nijmegen.
- Graaf, Paul M. & Harry B.G. Ganzeboom (1990). 'Intergenerational educational mobility in the Netherlands of birth cohorts 1891 through 1960.' *Netherlands Journal of Sociology*, forthcoming.

Boekbesprekingen

C.J.C. Rutenfrans, Criminaliteit en sexe. Een verklaring voor de verschillen in het criminele gedrag van vrouwen en mannen, Gouda Quint bv, Arnhem, 1989, 350 pagina's, prijs f 62,50.

In september 1989 promoveerde C.J.C. Rutenfrans op het proefschrift *Criminaliteit en sexe*. Dit goed geschreven boek deed nogal wat stof opwaaien door de beweerde bio-fysiologische verklaringen voor de verschillen in omvang en aard van criminaliseerbare en gecriminaliseerde gedragingen tussen mannen en vrouwen. Rutenfrans heeft zijn uitgebreide speurtocht niet beperkt tot biologische en fysiologische literatuur, hij nam ook een omvangrijke hoeveelheid psychologische, sociologische en (cultureel-) antropologische publikaties door. Hij komt tot een gecompliceerdere verklaring.

Rutenfrans brengt twee clusters van 'universele' verschijnselen met elkaar in verband. Het eerste cluster heeft betrekking op de verschillen tussen groepen mannen en groepen vrouwen qua regelovertreedende gedragingen. De volgende conclusies worden getrokken: 'a. Minder vrouwen dan mannen plegen delicten. b. Vrouwen begaan met een geringere frequentie delicten dan mannen. c. Het kwantitatieve verschil in het criminele gedrag van vrouwen en mannen is het grootst bij de geweldscriminaliteit en het kleinst bij vermogenscriminaliteit. d. Het overgrote deel van de door vrouwen gepleegde criminaliteit bestaat uit niet geweldadige vermogensdelicten, waarvan winkel-

diefstal kwantitatief het belangrijkste is...' (p. 21-22).

Het tweede cluster betreft de universele verschillen in 'manlijke en vrouwelijke patronen'. Daarover wordt, na behandeling van biologische en psychologische literatuur – van Cesare Lombroso aan het eind van de 19e eeuw tot Gisela Konopka in 1966 – verondersteld dat 'stabiliteit' kenmerkend is voor vrouwen (als groep) en 'variabiliteit' voor mannen (als groep). Een verbijzondering van die stabiliteit van vrouwen zou hun sterkere relatiegerichtheid zijn en hun streven naar bestendige, gelijkwaardige relaties. Dominantiegerichtheid van mannen wordt gezien als een verbijzondering van de mannelijke variabiliteit. Deze dominantiegerichtheid wordt omschreven als het streven naar een zo hoog mogelijke status binnen de groep waarvan men deel uitmaakt. Vrouwen zouden streven naar horizontale en mannen naar verticale relaties. Deze veronderstellingen over de 'fundamentele verschillen tussen de geslachten' worden tot 'feiten' na behandeling van cultureel-antropologische en 'cross-culturele psychologische' gegevens (p. 127). De onderscheiden mannelijke en vrouwelijke patronen blijken dan universeel te zijn. Op p. 168 is echter weer sprake van 'de door mij veronderstelde fundamentele verschillen tussen de geslachten'. De lezer verkeert voortdurend in onzekerheid over de 'hardheid' van dit essentiële onderdeel van het betoog.

Maar of het nu om een veronderstelling gaat of om een vaststelling, Rutenfrans ziet die fun-

Kritische bijdragen

'Achievement' en 'ascription': variaties op een thema

Frans Kerstholt

Een indruk die ik al langer had, is de afgelopen jaren steeds sterker geworden. Wij lijken aperte en vaak elementaire fouten in onderzoek eigenlijk alleen nog maar te ontdekken als wij dat onderzoek aan tweedejaars studenten moeten uitleggen. Dit roept vragen op over de kwaliteitscontrole door bijvoorbeeld redacties van wetenschappelijke tijdschriften, redacties en wetenschappelijke reeksen en promotiecommissies, die ik graag ter overweging wil voorleggen.

Omdat het andere voorbeeld in een bijdrage van onze student Van Kolverschoten aan de orde wordt gesteld, beperkt ik mij tot één van de zure vruchten van zo maar een onderwijssemester. Uit inspectie van door vooraanstaande Nederlandse stratificatie-onderzoekers gehanteerde procedures kwam naar voren, dat zij bijna een gewoonte hebben gemaakt van de constructie van ondeugdelijke data. Op basis van deze data werden vervolgens over het algemeen hoog gewaardeerde publicaties tot stand gebracht (boeken, waaronder minstens één proefschrift, artikelen in verschillende vaktijdschriften en waarschijnlijk een niet al te gering aantal internationale papers).

In 1978 had de Leidse statisticus De Leeuw in de *Sociologische Gids* (1978, 1) ondubbelzinnig de onjuistheid van een door Dronkers en De Jong gebruikte procedure aangetoond. In een artikel in *Mens en Maatschappij* (1987, 1) hebben Ganzeboom, Luijckx en 5(!) anderen dezelfde fout herhaald. Evenals hun collega-stratificatie-onderzoekers negen jaar eerder hebben de Utrechtse onderzoekers de gegevens uit een groot aantal datasets – het aantal was nu 10 – in één omvattend bestand samengevoegd. Zij gaven hiervoor de volgende rechtvaardiging: 'De hier te analyseren gegevens zijn afkomstig uit 10 verschillende surveys, die alle een toevalssteekproef uit de Nederlandse beroepsbevolking omvatten ...' (Ganzeboom, Luijckx en 5 anderen, 1987, p. 23) Wat hier in de eerste plaats opvalt is de uiterst vage omschrijving van de populatie waaruit de 10 steekproeven getrokken zijn. De beroepsbevolking verandert nogal snel. Daarom had in de omschrijving van de populatie toch minstens een historische specificatie van de populatie gegeven moeten worden. Dan had de lezer zich al iets geïnformeerder kunnen afvragen of de 10 surveys inderdaad allemaal op dezelfde manier aselekte steekproeven uit dezelfde populatie zijn. Nu ontbrak elke informatie. Er was evenmin informatie over de misschien wel verschillende procedures van steekproeftrekking, die in de 10 surveys

waren gehanteerd. Ik moest dus maar aannemen dat men niet alleen voor het in mijn werkomgeving uitgevoerde IVA-onderzoek, maar in alle gevallen een aselekte steekproef uit huishoudens en alleenstaanden getrokken had. Zoals gemakkelijk is in te zien, impliceert de aangeduide procedure verschillende kansen voor individuen om deel uit te maken van een steekproef. De grootte van huishoudens is immers variabel. In tegenstelling tot de suggestie van Heinen en Maas is de steekproef dus niet zelfwegend (Heinen & Maas, 1984, p. 2 en noot 2, p. 74). Ik neem verder aan, dat verkiezingsonderzoekers rechtstreeks in individuen geïnteresseerd zijn. Daarom kan ik mij niet voorstellen, dat de steekproeven voor de gebruikte kiezersonderzoeken volgens de door het IVA gebruikte methode getrokken zijn. Ik hecht dan ook nauwelijks geloof aan de hierboven geciteerde mededeling van het 'Utrecht Mobility Seminar'.

Dat de betrokkenen zich niet bewust zijn geweest van de problemen waar het hier om gaat, blijkt bijna 3 jaar later. In de boekaflevering van *Mens en Maatschappij* verschijnt in 1989 een artikel waarin verslag wordt gedaan van analyses op hetzelfde samengevoegde bestand (Ganzeboom & P.M. de Graaf, 1989). Hoewel, de auteurs geven nu een iets specifiekere rechtvaardiging voor de gevolgde procedure: 'Het gebruik van meerdere surveys zorgt er niet alleen voor dat het aantal te analyseren eenheden groot is, maar ook dat er toevallige vertekeningen die eigen zijn aan de steekproefprocedures en metingen worden gerandomiseerd. Het totale bestand vormt een steekproef uit de Nederlandse bevolking geboren tussen 1891 en 1960 en nog in leven zijnde tussen 1970 en 1985' (Ganzeboom & P.M. de Graaf, 1989, p. 64). De nieuwe verantwoording heeft twee voordelen. In de eerste plaats is zij specifiek en dus meer open voor kritiek. Een nog duidelijker omschrijving van de door Ganzeboom en De Graaf aangeduide populatie luidt mijns inziens: de Nederlandse bevolking geboren tussen 1891 en 1960 en niet overleden voor het kiezersonderzoek van 1970. In de tweede plaats blijkt de ondeugdelijkheid van de gevolgde procedure nog veel duidelijker. Het is immers duidelijk dat de leden van de aldus omschreven populatie nogal verschillende kansen hebben om deel uit te maken van de tussen 1970 en 1985 getrokken steekproeven. Ik heb mij laten vertellen, dat die ene in 1891 geboren respondent in 1970 vrijwel onmiddellijk na zijn interview overleden is.

Hierboven vermeldde ik al, dat De Leeuw in 1978 een zelfde fout signaleerd heeft. Het lijkt mij goed hem nog eens aan het woord te laten: 'Er is veel te doen geweest over de vraag of men gegevens uit verschillende onderzoekingen wel mag combineren om ze te gebruiken in een enkele analyse. (...) Alles mag, alleen moet men wel laten zien dat de conclusies die men trekt inderdaad uit de gedane bewerkingen volgen. In dit geval willen we uitspraken doen over de parameters van een multivariate verdeling. Op basis van aselekte steekproeven kunnen we deze parameters schatten. Dit betekent dus dat D&DJ van de impliciete veronderstelling uitgaan dat de (...) (gebruikte steekproeven, F.K.) (...) allemaal aselekte steekproeven uit dezelfde populatie zijn. Deze aanname lijkt onhoudbaar. (...) Uit het oogpunt van eenvoudige wetenschappelijke kwaliteitscontrole is het betreurenswaardig dat analyses gepubliceerd worden waarvan niet eens de mate van misleidendheid vastgesteld kan worden.' (De Leeuw, 1978, p. 36) Ik meen in het voorgaande aannemelijk gemaakt te hebben, dat het oordeel van De Leeuw opnieuw onverkort geldt. De kritiek van De Leeuw gold indertijd onderzoek van een

nog steeds actief Nederlands stratificatie-onderzoeker. Dat vooraanstaande Nederlandse stratificatie-onderzoekers van de les van De Leeuw weinig geleerd hebben, roept vragen op over de kwaliteitshandhaving in de Nederlandse sociologie.

Ik heb mij ook afgevraagd wat de mobiliteitsonderzoekers ter redding van hun projecten zouden kunnen ondernemen. Enigszins tot mijn verbazing leverde ook de paragraaf over 'Cumulating Cases and Combining Statistics from Samples' in Kish (1987, p. 183-192) weinig aanknopingspunten op.

Een artikel van Adriaanse (1976) bracht mij daarentegen op een idee dat wellicht serieuze overweging verdient. Adriaanse laat ons eerst 'bedenken dat een statistische toets is opgebouwd uit drie elementen: een populatie, een steekproef en een selectieprincipe (de wijze waarop de steekproef uit de populatie verkregen is)'. Vervolgens stelt hij: 'Nu is statistische toetsing of voorspelling mogelijk indien van deze elementen er twee bekend zijn' (Adriaanse, 1976, p. 165). Het probleem wordt dan hoe de situatie van de Utrechtse mobiliteitsonderzoekers zodanig geherinterpreteerd kan worden dat ook voor hen geldt, dat er twee elementen van de drie bekend zijn. Mij lijkt, dat men de 10 gebruikte steekproeven zou kunnen beschouwen als steekproeven uit de populatie, die zij gezamenlijk vormen. Zij zouden dus steekproeven kunnen zijn uit de steekproef, die Ganzeboom, Luijckx en 5 anderen zien als een representatieve steekproef uit iets waarvan ik de kenmerken niet ken. Zo gezien zou getoetst kunnen worden of elke steekproef een aselekte steekproef vormt uit de door de onderzoekers samengevoegde bestanden. De verschillende multivariate verdelingen mogen dan niet significant van die van het samengevoegde bestand afwijken. Indien de verschillende bestanden die toets zouden overleven, zou er wellicht enige serieuze reden zijn om te bedenken uit welke populatie de samengevoegde steekproef op haar beurt een steekproef zou kunnen zijn.

Tenslotte zou men kunnen proberen met behulp van econometrische methoden, zoals die bijvoorbeeld in een recente publikatie van Sayrs (1989) voor sociologen toegankelijker gemaakt worden, aannemelijk te maken, dat de verschillende steekproeven op belangrijke punten dezelfde structurele kenmerken en relaties vertonen. Vooral de methode waarbij voor de steekproeven dummy-variabelen gemaakt worden, lijkt in dat verband een aantrekkelijke kandidaat of minstens een richtinggevend model. Het is mij echter niet duidelijk hoe men vanuit de orde, die men in zijn data gezien heeft, dwingend kan concluderen tot de kwaliteit, bijvoorbeeld de representativiteit, van diezelfde gegevens. Hanushek en Jackson (1977, p. 128-129) attenderen zelfs op een ernstig statistisch probleem. Zij stellen, dat bij de in ons verband relevante onderzoekshypothese, namelijk dat de geaggregeerde waarnemingen uit dezelfde populatie komen, een ontoetsbare nul-hypothese hoort. Deze nul-hypothese houdt in, dat de 'coefficients of subsample differences... are different from zero.' Zij is ontoetsbaar, omdat '...our statistical tests are all designed to test simple hypotheses - that the parameter equals a given value...' Zij suggereren voorzichtig gebruik van een principieel niet afdoende statistische toets. De voorzichtigheid ligt in een 'lower level of acceptance', bijvoorbeeld een significantie-niveau van 25% in plaats van 5%. Mij lijkt, dat het gebruik van een niet adequate statistische toets in combinatie met een altijd enigszins willekeurig aangepast significantie-niveau wat al te veel ruimte laat voor geloof (of juist het tegendeel daar-

van) in de kwaliteit van de data. De kans op 'wishful' analyseren lijkt mij onaanvaardbaar groot.

Mijns inziens is hierboven de vinger op een aantal premature publikaties gelegd. Men kan zich afvragen waar de gesignaleerde nogal zorgeloze houding op het punt van de kwaliteit van data haar oorsprong heeft. Ik heb het vermoeden dat een houding die volgens mijn krant door Bourdieu als 'archaisch positivisme' is aangeduid en waartegen ik in mijn proefschrift (Kerstholt, 1989, hst. 4) ook enig voorzichtig tegengas heb willen geven, een belangrijke rol speelt. Archaisch positivistische sociologen, die op zoek zijn naar de invariante wetmatigheden van het sociale leven, menen al te makkelijk geen representatieve data nodig te hebben. Algemene wetmatigheden gelden immers ook voor niet-representatief materiaal. Vooral bij zeer representativiteitsgevoelig onderzoek, en mobiliteitsonderzoek gericht op de vergelijking van cohorten is dat nu eenmaal, kan deze houding fatale gevolgen hebben.

Het besproken voorbeeld maakt tenslotte duidelijk, dat wij in ons vak niet al te lichtvaardig op geloof in reputaties ('ascription'!) moeten varen. De kwaliteitscontrole dient zodanig verbeterd te worden, dat ernstige fouten niet pas in de onderwijssituatie ontdekt worden. Reputaties kunnen dan op een wellicht meer verantwoorde manier ('achievement'!) tot stand komen en tevens object van regelmatige evaluatie zijn. Dit is overigens alleen al noodzakelijk indien wij een echte toekomst voor het vak wensen.

Literatuur

- Adriaanse, R. (1976). De toepasbaarheid van statistische toetsen op populatiegegevens, *Sociologische Gids*, 3, p. 159-171.
- Dronkers, J. & U. de Jong (1978). Jencks en Fägerlind op zijn Hollands: Een aanzet tot de studie van de relaties tussen milieu, intelligentie, onderwijs, beroep en inkomen, *Sociologische Gids*, 1, p. 4-30.
- Dronkers, J. & U. de Jong (1978). *Sociologische Gids*, 1, p. 39-42.
- Gadourek, I. & J. Peschar (reds.) (1989). *De open samenleving?*, Boekaflevering Mens en Maatschappij, 64e jaargang.
- Ganzeboom, H. & P. de Graaf (1989). Intergenerationele opleidingsmobiliteit in Nederland van geboortecohorten 1891-1960, *Sociale Wetenschappen*.
- Ganzeboom, H., R. Luijckx en 5 anderen (1987). Intergenerationele klassenmobiliteit in Nederland tussen 1970 en 1985, *Mens en Maatschappij*, 1, p. 17-43.
- Hanushek, E. & J. Jackson (1977). *Statistical Methods for Social Scientists*, New York.
- Heinen, A. & A. Maas (1984). *Het NPAO-arbeidsmarktonderzoek. Resultaten van de eerste analyses: mobiliteitsgeneigdheid en segmentering van de arbeidsmarkt*, IVA, Instituut voor Sociaal-Wetenschappelijk Onderzoek van de Katholieke Hogeschool Tilburg, Tilburg.
- Kerstholt, F. (1989). *Tussen rationele keuze en durkheimiaanse rationaliteit*, Tilburg University Press.
- Kish, L. (1987). *Statistical Design for Research*, New York.
- De Leeuw, J. (1978). De politieke relevantie van correlaties, *Sociologische Gids*, 1, p. 31-39.
- Sayrs, L. (1989). *Pooled Time Series Analysis*, Sage University Paper on Quantitative Applications in the Social Sciences, nr. 70, Newbury Park.

Bestaat het wetenschappelijk forum eigenlijk wel?

Peter van Kolverschoten*

Hoe belangrijk zijn methoden en technieken van sociologisch onderzoek? Tijdens mijn propaedeutische sociologie had ik een docent die het belang hiervan onderstreepte met bedreigingen en verwensingen. Nog steeds ben ik ervan overtuigd dat hij, als ik te erg tegen de regels zondig, 's nachts aan mijn bed zal verschijnen. Het is er dus goed ingeprent en naar mijn mening terecht. U weet waar het om gaat: betrouwbaarheid en geldigheid. Gelukkig – zo werd ons verteld – is er het wetenschappelijk forum, dat genadeloos corrigeert. Bestaat dat forum eigenlijk wel?

In het kader van mijn studie vrijetijdswetenschappen aan de KUB heb ik het boek van H. Ganzeboom, 'Leefstijlen in Nederland', bestudeerd. Mijn oordeel was heel wat minder positief dan dat van de recensent van Mens en Maatschappij (november 1989). Nu ben ik natuurlijk nog maar een eenvoudig student. Daarom wil ik het ook niet wagen in dit kader in te gaan op alle aspecten die twijfels bij mij oproepen omtrent de kwaliteit van het boek.

Er is echter één fout die het noodzakelijk maakt te reageren. Een fout die eerst gemaakt is door P.M. de Graaf in zijn promotieonderzoek (P.M. de Graaf: 'De invloed van financiële en culturele hulpbronnen in schoolloopbanen', Nijmegen, 1987). Ganzeboom gebruikt deze gegevens ter empirische ondersteuning van zijn studie. Deze fout vermindert de empirische waarde van zijn studie zodanig, dat ik mij afvraag of publicatie wel verantwoord is.

Ter zake. Uit de correlatie-matrix, die Ganzeboom presenteert (pag. 82-83) blijkt dat zowel de culturele als de economische beroepsstatus van de respondenten sterker samenhangt met de opleiding van de partner dan met de eigen opleiding. Dat leek mij een nogal opmerkelijk onderzoeksresultaat, waar het gaat om inwoners van de stad Utrecht tussen 25 en 64 jaar. Aangezien Ganzeboom hier echter niet op ingaat, ben ik in het boek van P.M. de Graaf (1987) gaan kijken hoe hij aan zijn gegevens is gekomen.

Wat is er gebeurd? Utrechtse studenten zijn er op uit gestuurd om gegevens te verzamelen op bepaalde adressen. Degene die de deur opendeed moest ondervraagd worden. Toen de gegevens binnenkwamen viel het de onderzoekers op dat vrouwen oververte-

* Student Vrijetijdswetenschappen K.U. Brabant

genwoordigd waren. Begrijpelijk, de man (kostwinner) is naar het werk, zo dachten zij. Er werd dus gecorrigeerd naar geslacht. De gegevens van de mannen kregen een gewicht van 1,5. Met het hierboven gememoreerde vreemde resultaat.

Inderdaad, vrouwen waren oververtegenwoordigd. Waarom? Omdat de man in het algemeen nog steeds kostwinner is. Echter, het is mogelijk of misschien zelfs aannemelijk dat een groot deel van de ondervraagde mannen 'huisman' is met een werkende vrouw.

Er is hier blijkbaar onderzoek gedaan naar *niet-kostwinners*. Voor deze categorie is het inderdaad *niet* verwonderlijk dat de culturele- en economische beroepsstatus sterker samenhangt met de opleiding van de partner dan met die van de respondent zelf.

Ganzeboom lijkt deze fout niet ontdekt te hebben en maakt daardoor onverantwoorde generalisaties. Dit maakt mijns inziens de resultaten van zijn analyses dubieus of zelfs waardeloos.

Ik kom terug op mijn vraag: Bestaat het wetenschappelijk forum eigenlijk wel? De Graaf is op zijn proefschrift gepromoveerd. Een hele promotiecommissie vol geleerde heren heeft hier als wetenschappelijk forum gefaald. Dan volgt Ganzeboom. Hij had toch zeker de empirie, die hij gebruikt als illustratie van zijn overigens theoretisch verhaal, kritisch moeten bekijken. Ook Ganzeboom faalt in dit opzicht.

Zo volgt dan blijkbaar de ene autoriteit de andere. Het Sociaal Cultureel Planbureau, uw recensent Bart Bakker en nogmaals Harry Ganzeboom, die in zijn volgende boek: 'Cultuurdeelname in Nederland' (1989) weer met dezelfde empirische gegevens aankomt. Hij vermeldt deze keer wel de oververtegenwoordiging van vrouwen in het bestand en de correctie, door mannen 1,5 maal zo zwaar te wegen, maar hij heeft blijkbaar nog steeds niet ontdekt wat er werkelijk fout is.

De gevolgen van het ontbreken van een wetenschappelijk forum kunnen snel om zich heen grijpen. Zo wordt het boek van Ganzeboom (1988) gebruikt in een artikel van de journalist Hans Rube in 'Het Nieuwsblad' (dagblad voor Midden-Brabant). Het wetenschappelijk forum lijkt niet te werken. Of zijn het uiteindelijk studenten die 's nachts aan het bed van de hooggeleerden moeten verschijnen?