

Enkele statistische aspecten van sociaal onderzoek

door

Mr W. H. Somermeijer

Wie sociale research bedrijft, komt in verschillende fasen van zijn onderzoek met statistische problemen in aanraking. Hoe eerder hij deze opmerkt en oplost: des te beter voor hem en zijn onderzoek. Helaas merkt niet iedere onderzoeker deze op, of wel behandelt hij ze en bagatelle: des te erger. Statistische problemen zijn o.m. verbonden aan:

- I het ontwerp van de vragenlijst;
- II het ontwerp van de steekproef;
- III de verwerking en interpretatie van de uitkomsten der enquête.

I. DE VRAGENLIJST

I. Vraag niet te veel

De in de vragenlijst op te nemen vragen moeten noodzakelijk en voldoende zijn ter bereiking van het gestelde doel. Dit betekent, dat het doel zo scherp mogelijk geformuleerd moet worden. Daartoe zal het soms in een aantal deeldoeleinden gesplitst dienen te worden.

Deze doeleinden kunnen vooreerst het gemakkelijkst worden vastgelegd in een schema van bewerkingen, waaraan men de uit de enquête te putten gegevens wil onderwerpen, zomede van tabellen, waarin men deze of daaruit af te leiden gegevens wil vastleggen. Dit schema van tabellen en bewerkingen dient aan de definitieve vaststelling van de vragenlijst vooraf te gaan¹). Men moet geen vragen stellen, „omdat ze wel aardig zijn“, of zo maar „om te zien wat er uitkomt“. Speciaal hier geldt: overdaad schaadt, d.i. de overdaad, welke blijkt uit het niet verwerken van gevraagde gegevens, welke geen controlerende of introductoire functie bezitten.

Ook al kost een overbodige vraag slechts weinig tijd en papier, zij bezwaart het onderzoek en kan de betrouwbaarheid der antwoorden op andere vragen ongunstig beïnvloeden, wanneer de geënquêteerden — niet geheel ten onrechte — de indruk zouden krijgen, dat toch maar „raak“ (dus mis) gevraagd wordt. Een dergelijke vragerij lokt lukraak antwoorden uit. Men onthoude: „In der Beschränkung zeigt sich erst der Meister“.

Een dergelijk voorwerk is wel noodzakelijk, doch niet voldoende om vast te

stellen of aan bovengenoemde eis voldaan is. Immers, tevens moet worden nagegaan of:

- a. de in de vragenlijst gebezigde begrippen ondubbelzinnig gedefinieerd, d.i. bij passende vraagstelling (door de enquêteur) slechts voor één uitleg vatbaar zijn;
- b. de vragen zodanig zijn, dat gezien de aard van de bevolkingsgroep waaronder de enquête gehouden zal worden, voldoende betrouwbaarheid en volledigheid bij de beantwoording gewaarborgd geacht mag worden²⁾;
- c. vragen stelt welke teveel in detail afdalen. vragen voldoende nauwkeurige antwoorden te verkrijgen, zulks in verband met de spreiding der laatste³⁾.

Met deze punten wordt helaas nog te weinig rekening gehouden, omdat men:

- a'. geen of onvoldoende besef heeft van de definitie-variabiliteit der gebezigde begrippen;
- b'. vragen stelt welke, wegens haar „onbescheiden" karakter of geladenheid met sentiment een foutief antwoord als het ware uitlokken;
- c'. vragen stelt welke teveel in detail afdalen.

2. Vraag duidelijk, definieer scherp

Statistisch onervarenen zullen versteld staan over de veelheid van definities voor begrippen welke men in het dagelijks leven argeloos pleegt te bezigen⁴⁾. Tussen die definities bestaat geen onderscheid van „juist" of „fout", doch hoogstens een van doelmatigheid.

Welke definitie in een bepaald geval doelmatig is, hangt o.m. af van het doel van het onderzoek. Het is echter minder belangrijk welke definitie men kiest, dan wel dat men een eenmaal gekozen definitie consequent toepast en tevens dat de definities van verschillende begrippen in eenzelfde vragenlijst met elkaar in overeenstemming zijn. Anders zondigt men, zoals zo vaak in verhandelingen van „geesteswetenschappelijke" stempel, tegen het denkprincipe der identiteit (dat men psychologisch kan interpreteren als het principe van de denkdiscipline).

Al blijft de definitie-keuze in zekere zin willekeurig, niettemin kan men hiervoor een houvast vinden aan wat op dit gebied „mos" geworden of zelfs „gecodificeerd" is.

Zo verdient het aanbeveling om zoveel mogelijk gebruik te maken van de in „codes" door de United Nations of haar Specialized Agencies neergelegde richtlijnen t.a.v. definities en classificaties⁵⁾. Hierin is zoveel denk- en praktische ervaring van de gehele wereld geïncorporeerd als door geen enkele individuele onderzoeker afzonderlijk opgebracht zou kunnen worden. Hierdoor vertegenwoordigen haar resultaten het summum van doordacht-

heid. Toepassing dier „codes" maakt de resultaten van individuele onderzoekingen met die van andere vergelijkbaar, dus verhoogt haar betekenis. Helaas spelen individualisme en eigenwijsheid op dit gebied, onder het motto van „aanpassing aan de bijzondere plaatselijke omstandigheden", nog hoogtij. Voor de vertegenwoordigers dezer categorie is het misschien goed te weten, dat bovenbedoelde internationale codes door haar decimale hiërarchie zo soepel zijn dat door consolidatie enerzijds en differentiatie anderzijds b i n n e n h e t k a d e r d i e r c o d e s de gewenste „aanpassing" bereikt kan worden.

3. Wilt U fouten vermijden: bereid U op haar voor!

Twee soorten fouten zijn te onderscheiden: toevallige en systematische. De eerste zijn minder ernstig dan de tweede: zij zullen elkaar immers ten naaste bij compenseren, al zullen zij correlatie- en regressie-coëfficiënten vitiëren⁶). Systematische fouten zijn echter, ook in de sociale sfeer, ten minste zo talrijk als de toevallige. De mens is een kuddedier, ook in zijn psychische aberraties. De onderzoeker houde verder steeds voor ogen, dat de mens zich ten overstaan van de statistiek aan alle zwakheden overgeeft, welke hij zich als homo „sapiens" eigen pleegt te maken. Dit blijkt bv. reeds uit een schijnbaar zo simpele vraag als die naar de leeftijd; de bij de antwoorden op deze vraag geloosde leugens en minder opzettelijke onwaarheden kunnen bv. voortkomen uit: ijdelheid (te lage leeftijdsopgave bij vrouwen), angst (bij een Volkstelling in het vooroorlogse Tsjecho-Slowakije bleken opmerkelijk veel jongemannen de 17-jarige, en opmerkelijk weinig de 18-jarige, d.i. dienstplichtige leeftijd te bezitten), bijgeloof (in sommige streken opmerkelijk veel, in andere, met afwijkende levensopvatting, opmerkelijk weinig geboorten op Zondag), getallenfetichisme (voorkeur voor bepaalde eindcijfers, i.h.b. 0 en 5, daarna 2 en 8: voor leeftijden, als in Brazilië, of voor geboortejaren, als in Polen), onwetendheid of vergeetachtigheid (t.a.v. het feit of t.a.v. de invulling), in de aantallen „leeftijd onbekend" tot uiting komende.

Een tweede voorbeeld: egoïsme blijkt uit de te lage opgaven voor de inkomstenbelasting, welke tevens aan de inkomensstatistiek ten grondslag ligt.

Een derde voorbeeld: schaamte blijkt uit het onwaarschijnlijk geringe aantal personen dat volgens de Volkstelling Nederland 1920 met lichaamsgebreken behept zou zijn. Inziende dat men op een dergelijke vraag toch geen betrouwbare antwoorden mag verwachten, heeft men haar — terecht — bij volgende Volkstellingen achterwege gelaten: „who doesn't ask questions, doesn't hear lies". Moraal: men stelle geen vragen waarop men geen betrouwbare antwoorden mag verwachten. M.a.w.: men geve zo min mogelijk reden tot liegen.

Uiteraard is een grote mate van ervaring vereist om te weten wat men redelijkerwijs wel en wat men niet kan vragen. Dit hangt ook van de te enquêteren bevolkingsgroep af. Een bepaalde vraag, welke in zeker milieu taboe is, kan in een ander milieu gerust gesteld worden.

Ook houde men rekening met het beperkte begrip der mensen, mede in verband met hun ontwikkelingspeil. Men schiete vooral niet te hoog. Men houde zich voor de geest de uitspraak van Dr Luning Prak, dat het intelligentie-peil van de gemiddelde Nederlander gelijk staat met dat van een kind van 14 jaar, en dat de gemiddelde Nederlander degene is die moeite heeft met 6 jaar lagere school.

Ook stelle men niet te strenge eisen aan het geheugen, d.w.z. men stelle niet vragen betreffende een te ver in het verleden liggende periode⁷⁾.

4. De sociaal onderzoeker zij geen detaillist, wel controleur

Weliswaar bezigt men vaak een steekproef ter verkrijging van detail-gegevens, voor welke verzameling een volledige telling onbruikbaar of inefficiënt is⁸⁾, maar men zij zich bewust van de toevallige fouten, welke aan een steekproef inhaerent zijn. Deze maken beperking ten aanzien van de detaillering nodig, opdat de uitkomsten wegens onnauwkeurigheid niet onbruikbaar worden. Die beperking zal des te geringer zijn naarmate de omvang van de steekproef groter, zomede de spreiding der antwoorden kleiner is; zij is a.h.w. de prijs welke men moet betalen om de tijds- en geldbesparing, welke een steekproef ten opzichte van een volledige telling oplevert, deelachtig te worden.

Om de uitkomsten van de steekproef zo betrouwbaar mogelijk te maken is het negatieve: bepaalde vragen niet te stellen, niet voldoende. Daartegenover zal men goed doen bepaalde contrôle-vragen in te lassen; deze kunnen het karakter van strik-vragen verkrijgen⁹⁾.

Mèt deze vragen kunnen onopzettelijk, en voor een deel ook opzettelijk verkeerde antwoorden op andere vragen achterhaald worden: overtuigend en consequent liegen eist een mate van intelligentie waarover niet veel mensen beschikken. Uit steekproef-technisch oogpunt bezien heeft gebrek aan intelligentie dus niet alleen iets tegen (zie hiervòòr), doch ook iets vòòr. Papieren contrôle alleen is echter niet voldoende: feitelijke contrôle, voor zover mogelijk, is beter.

5. Statistique: l'art de poser des questions

De betrouwbaarheid der antwoorden is voorts afhankelijk van de wijze waarop de vragen gesteld worden.

Het is nl. uit proeven met alternatief geformuleerde, doch materiëel gelijke vragen gebleken, dat de vraagformulering invloed heeft op de antwoorden^{9a)} Hoewel men a priori bepaalde formuleringen kan verwerpen, omdat zij suggestief gesteld zijn¹⁰⁾ of tot misverstand aanleiding kunnen geven, zal men voor de definitieve vaststelling der vragen goed doen met alternatieve formuleringen proeven te nemen¹¹⁾. Evenwel zullen deze niet altijd omtrent de optimale formulering uitsluitel kunnen geven. Vooreerst kan de betrouwbaarheid der antwoorden niet steeds beoordeeld worden. Ook is het mogelijk dat er geen (statistisch) significant verschil in betrouwbaarheid tus-

sen alternatieve formuleringen te constateren is; dan maakt het weinig verschil welke formulering men kiest. Bovendien is het mogelijk dat men slechts dan zo betrouwbaar mogelijke antwoorden verkrijgt, wanneer men de vraag-formulering aanpast aan sociaal en geografisch milieu, eventueel ook aan geslacht en leeftijd der geënquêteerden. Dit kan verband houden met verschil in woordgebruik door verschillende bevolkingsgroepen¹²⁾, zomede verschil in instelling der geënquêteerden tegenover de vragen.

Hiermee kan bij de opstelling van het vraag-formulier slechts in beperkte mate rekening worden gehouden¹³⁾. Het zal dan niet mogelijk blijken te zijn aan de enquêteur de precieze formulering der vragen voor te schrijven (zo deze zich hieraan zou kunnen houden); hierbij wordt van diens mogelijke eigenwijsheid of zucht tot afwisseling nog afgezien¹⁴⁾. Het hieraan verbonden nadeel van grotere subjectieve interpretatie-vrijheid der enquêteurs kan bij deugdelijke selectie, scholing, instructie der enquêteurs, en door proefenquête onder controle wel verminderd, doch niet opgeheven worden. Of dit nadeel groter of kleiner is dan hetwelk voortvloeit uit de alternatieve methode met vaste formulering, moet voor elk onderzoek afzonderlijk worden vastgesteld.

- 1) Bovendien is dit uit organisatorisch oogpunt wenselijk. Dit desideratum is door de United Nations in haar „Population Census Handbook” voor Volkstellingen uitgesproken; het is echter evenzeer voor andere sociale onderzoeken (van beperkter omvang, doch groter diepgang) van betekenis.
- 2) Wat in een gegeven geval als grens voor „voldoende” (uit te drukken in een frequentie-verdeling voor toelaatbare afwijkingen, resp. percentage voor „non-response”) aangemerkt dient te worden, hangt af van de variabiliteit van gegevens en de aard der te verwachten fouten, met het oog op de vereiste nauwkeurigheid der totaal-gegevens, af te leiden uit het doel van het onderzoek.
- 3) Acht men de vraag van voldoende belang (gesteld tegenover de daarvoor te maken kosten) dan dienen omvang en stratificatie der steekproef zodanig te zijn, dat de steekproeffouten voor de hiermee te verkrijgen uitkomsten binnen toelaatbare grenzen blijven.
- 4) Men leze bv. in het compilatie-werk der United Nations: „Population Census Methods” op hoe verschillende manieren „gezin”, „plattelandsbevolking”, „beroepsbeoefenaar” en „beroep”, ja zelfs „bevolking” (met de „de jure” en „de facto”-hoofdvarianten) gedefinieerd, en daarmee statistisch benaderd kunnen worden.
- 5) Behalve de „Population Census Methods” zie men bv. ook de beroepsclassificatie van het International Labour Office en de doodsoorzaken- en ziekteclassificatie van de World Health Organisation.
- 6) Men zie bv. J. Sittig-H. Freudenthal: „De juiste maat”, in het bijzonder Hoofdstuk II, § 3; in dit onderzoek konden de toevallige fouten nog voor een groot deel gecorrigeerd worden — een voorbeeld ter navolging!
- 7) Zo werd in het onderzoek van Kinsey c.s., „Sexual behavior in the human male”, veel van het geheugen der geënquêteerden gevergd; de daarbij

gestelde vragen hadden betrekking op het gehele (sexuele) leven der geënquêteerden, dat zich voor oudere personen over 4 of 5 decennia uitstrekt. Misschien echter herinneren de ouderen zich de verder in het verleden gelegen feiten beter dan de meer recente (retrograde amnesie), hetgeen een uitzondering op de regel vormt.

- 8) Ook kan, bij goede opzet, de kwaliteit der uitkomsten van een steekproef, door de daarbij mogelijke strengere selectie der tellers en intensievere controle (ter plaatse!) beter zijn dan die van een volledige telling. Daarom wordt vaak de betrouwbaarheid van een volledige telling met behulp van een steekproef gecontroleerd.
- 9) Een oud trucje is het vragen naar leeftijd zowel als naar geboortjaar.
- 9a) Vgl. b.v. W. E. Deming, „On errors in surveys“ (American Sociological Review, 1944), i.h.b. „variability of response“; voorts E. S. Marks-W. P. Mauldin, „Response errors in census research“ (Journal of the American Statistical Association, 1950), waarin een deel van de ervaring van het Bureau of the Census met vraagonderzoek door middel van proefenquêtees is neergelegd.
- 10) Bij uitzondering kan een suggestieve vraag boven een neutrale te prefereren zijn, nl. als door de suggestie een neiging tot foutief antwoorden gecompenseerd wordt (contra-suggestie). Bv. wanneer aan een ongehuwde de vraag wordt gesteld: „**wanneer** hebt U voor het eerst (of laatst) geslchtsgemeenschap gehad“, waarmee de enquêteur de vanzelfsprekendheid hiervan suggereert en daarmee de geënquêteerde over eventuele schaamte heenhelpt, in plaats van de neutrale: „Hebt U reeds geslchtsgemeenschap gehad?“ Ook al zijn de als antwoord genoemde tijden als zodanig onbetrouwbaar, de door het „wanneer“ geconcretiseerde eerstgenoemde formulering draagt er toe bij, dat de kans dat geslchtsgemeenschap erkend wordt, *wanneer in feite geen heeft plaats gehad, althans in bepaalde milieus*, geringer is dan dat bij laatstgenoemde formulering geslchtsgemeenschap ontkend wordt, *wanneer deze in feite heeft plaats gehad*. Vgl. Kinsey a.o., „Sexual behavior in the human male“.
- 11) Dit kan op eenvoudige wijze geschieden met behulp van A- en B-formulieren, waarbij alle oneven genummerde geënquêteerden de vragen op A-wijze, en alle even genummerde geënquêteerden de vragen op B-wijze geformuleerd krijgen. Men kan dan onderzoeken of de A- en B- deel-steekproeven ten aanzien dier vragen significante verschillen vertonen.
- 12) Zo zal het „deftige“ woord „geslchtsgemeenschap“ in het voorbeeld van noot 10) bij enquëtering onder personen van geringere ontwikkeling door een meer familiale, eventueel platte term vervangen moeten worden; hierdoor wordt het „rapport“ tussen enquêteur en geënquêteerde groter, hetgeen de betrouwbaarheid der antwoorden ten goede komt. Vgl. wederom Kinsey, i.c.
- 13) De vraagformules kunnen het best op een afzonderlijk formulier, **alleen ten behoeve van de enquêteur**, worden gedrukt (of gestencild); het enquêteformulier zelf kan dan beknopt en overzichtelijk blijven, hetgeen aan de snelheid der bewerking ten goede komt.
- 14) Op de wijze waarop de enquêteur zijn moeilijke en verantwoordelijke taak moet verrichten, kan hier niet worden ingegaan. Men zie daarom bv. de speciaal voor enquêteurs van het Institut National de Statistique et de Etudes Economiques samengestelde handleiding „Manuel des enquêteurs par sondage“.