

# Het experiment in de sociale wetenschappen

Overal zijn experiment, observatie en theorie zeer nauw met elkaar verbonden. In de sociale wetenschappen is het vrijwel onmogelijk die drie begrippen scherp te scheiden. Voortdurend is er een wederzijdse beïnvloeding: het experiment geeft bepaalde mogelijkheden tot observatie, maar de observatie-mogelijkheden bepalen ook welke soort experimenten mogelijk zijn; bijna steeds is de observatie, gezien het veel omvattende geheel van de sociale wetenschappen, slechts gericht op een gedeelte en het gekozen gedeelte hangt dan in belangrijke mate af van de theorie waaruit men vertrekt; anderzijds is de opgestelde theorie meer dan eens sterk afhankelijk van de gekozen mogelijkheid van observatiemethode. De titel, « het experiment in de sociale wetenschappen », hebben wij derhalve opgevat in zeer brede zin en uitgebreid tot de drie met elkaar verbonden begrippen.

Allereerst moeten over die drie termen een paar woorden gezegd worden.

Als *theorie* beschouw ik een beschrijving van het gebeuren in het verleden, heden of toekomst, met enkele kenmerken van het gebeuren zelf en met enkele parameters, die de omstandigheden aangeven en die al of niet regeerbaar zijn (die ofwel van buitenaf gegeven, ofwel in meerdere of mindere mate door ons beïnvloedbaar zijn). De in de theorie opgenomen relaties en afhankelijkheden kunnen verklaringen geven van het gebeuren en kunnen soms tot prognoses leiden.

Een theorie voert bijna steeds tot mathematisering. « The great architect of the universe now begins to appear as a pure mathematician » (Jeans); dit geldt niet alleen voor de fysica, maar ook meer en meer voor de sociale wetenschappen. Dit wiskundige staat echter niet volledig apart, maar moet geïntegreerd worden in de wetenschap, waarop het toegepast wordt.

De moeilijkheid spruit vooral voort uit de splitsing in onze soorten opleiding: literair-filosofisch en wiskundig-wetenschappelijk. (De in Nederland gebruikte term « natuurfilosofische » is jammer genoeg ook alleen maar een woord-synthese). Degenen, die het meest met de sociale wetenschappen in aanraking komen zijn over het algemeen literair-filosofisch

gevormd en staan eerder vreemd tegenover de wiskundige benadering. De niet-wiskundige zou moeten beginnen met de eigenaardige angst voor de wiskunde af te leggen (angst, die ons, m.i. kunstmatig, opgedrongen wordt, doordat van in de lagere school traditioneel verklaard wordt, dat rekenen en later wiskunde « moeilijk » is!). Men moet inzien, dat de wiskunde slechts een bepaalde taal is om het denkbare gemakkelijk te kunnen denken; Poincaré zegde: « la science sert à l'économie de la pensée » en dit geldt vooral voor de wiskunde.

In het recente werkje van Huxley over literatuur en wetenschap geeft hij het gebruik van de taal aan als een van de essentiële verschillen tussen literatuur en wetenschap: de literaire taal gaat meer naar het globaliseren, het samenvatten van algemene beelden, zodat atmosferen worden opgeroepen, terwijl in de wetenschappelijke taal de woorden afgepeld worden tot op het skelet, waar geen enkele dubbelzinnigheid meer mogelijk is, maar waar ieder woord een heel beperkte betekenis heeft. In beide gevallen gaat het echter om het gebruik van de taal. Een theorie, in wiskundige termen opgesteld, is een taalgebruik met zeer scherp omschreven symbolen (woorden), waarmee we hele, ondubbelzinnig bepaalde, gedachtensprongen kunnen maken (het getal drie is bijv. reeds het resultaat van een sterke abstrahering van een hele gedachtengang; hetzelfde kan een wiskundige denken van een integraal, die ook een bondige samenvatting is van een lange redenering).

Aan de andere kant zal degene, die uit de meer wiskundig-wetenschappelijke richting komt, soms de « feeling » missen voor de sociale wetenschappen. De binding tussen de werkelijkheid en het gemathematiseerde model gaat dan verloren. Die binding moet vooral op twee punten bestaan: enerzijds bij het opstellen van het model, waarbij de werkelijkheid zeer grondig gekend moet zijn en waarbij men vooral moet nagaan hoe ze geschematiseerd kan worden teneinde een zo eenvoudig mogelijk en toch zo veel omvattend mogelijk wiskundig model te kunnen opstellen, waarbij veel zal afhangen van de reële observatiemogelijkheden; anderzijds mag men bij de confrontatie van de wiskundige resultaten met de werkelijkheid en meer algemeen bij het interpreteren van de conclusies van het schema met betrekking tot de werkelijkheid niet vergeten hoe men geschematiseerd heeft en wat men verondersteld heeft. Veelal vergt deze tweevoudige confrontatie van het wiskundig model met de werkelijkheid een samenwerking van personen met verschillende opleiding. Deze samenwerking zal vooral gunstige resultaten kunnen opleveren, indien de samenwerkenden een inzicht verworven hebben van een algemene synthese, gestoeld op een grondige specialisatie. Een minimum van synthetisch inzicht is noodzakelijk om elkaar te kunnen begrijpen.

*Observatie* van de kenmerken van de beschrijving en van de parameters zal de theorie moeten ondersteunen. Om in een mathematisering te passen moeten deze observaties door getallen kunnen uitgedrukt worden; dit wil echter niet zeggen, dat er steeds zo iets als een maatschaal zou moeten bestaan, het is dikwijls ook voldoende dat er een rangorde kan opgesteld worden of soms zelfs dat men slechts een grootte-relatie tussen ieder tweetal waarnemingen kan bepalen. De essentiële vraag is echter wat geobserveerd moet worden en hoe.

Een voorbeeld hiervoor is het economisch begrip « prijspeil ». Wat moet hiervoor geobserveerd worden? Dikwijls worden het theoretische begrip « prijspeil » en het meetinstrument « indexcijfer » (bijv. van de kleinhandelsprijzen) wederzijds gedefinieerd, hetgeen ons alleen maar in een kring laat ronddraaien zolang noch het prijspeil, noch het indexcijfer een nauwkeurige autonome omschrijving hebben. Het enige middel is, dat men een keuze doet van een nauwkeurig vastgelegde berekeningsmethode voor het indexcijfer, d.w.z. dat er een bepaald meetinstrument gekozen wordt; en hiermee wordt dan het prijspeil vastgelegd. In plaats van dus een observatie-methode te kiezen in functie van een theoretisch gedefinieerd begrip, wordt het theoretisch begrip nauwkeurig vastgelegd door de keuze van een observatie-methode. Hierbij zal men wel het meetinstrument indexcijfer voortdurend moeten aanpassen aan de gewijzigde omstandigheden van de werkelijkheid, opdat het er mee verbonden theoretische begrip prijspeil zinvol zou kunnen gebruikt worden in een aangepast model van die werkelijkheid.

Een ander voorbeeld vinden we in het meten van de intelligentie. Ook hier dringt de vraag zich op: wat is het eigenlijk en hoe kunnen we het meten? En opnieuw zal de nauwkeurige vastlegging van het theoretisch begrip gebeuren door de keuze van een bepaald meetinstrument, nl. van een bepaalde intelligentie-testbatterij).

Een *experiment* in de klassieke zin, is het observeren van een geïsoleerde situatie, waarvan alle externe omstandigheden (de parameters) gecontroleerd en beheerst gevarieerd kunnen worden. Als wij die beperkte omschrijving van een experiment moeten behouden, kan daarvan in de sociale wetenschappen slechts zelden sprake zijn. Meer algemeen kan men het experiment echter beschouwen als de observatie van een reactie op een nauwkeurig omschreven ingrijpen in de situatie en dat kan dan ook een sociale situatie zijn. Hier moet dan de vraag gesteld worden naar het verband tussen experiment en observatie; zijn dat verschillende begrippen of lopen zij in elkaar over. Is bijv. het toepassen van een psychometrische test een observatie — met een bepaalde test als « meetinstrument » — of een experiment — nl. het ingrijpen in de situatie door het voorleggen

van een bepaalde vragenlijst. De grens zal moeilijk te trekken zijn, maar geldt dat ook niet voor « observatie » en « experiment » in de atoomfysica?

Zijn deze hierboven omschreven wiskundig geformuleerde theorie, getalmatige observatie en experiment in algemene zin ook mogelijk in de sociale wetenschappen, zoals ze het zeer zeker zijn in de natuurkunde en aanverwante wetenschappen?

Er zijn tussen de twee groepen wetenschappen in ieder geval verschillen in de omstandigheden en in het materiaal, die gemaakt hebben, dat men gemeend heeft, dat dergelijke benaderingen in de sociale wetenschappen volstrekt onmogelijk waren. Men werd nog in die mening versterkt door enige pogingen, die — toch minstens gedeeltelijk — als mislukt moesten beschouwd worden. Zo wilde o.m. Quetelet in zijn « Physique sociale » rechtstreeks de methodes van de klassieke natuurkunde overbrengen naar de sociologie, hetgeen zeker een té eng keurslijf was om de volle problematiek van een menselijke wetenschap te vatten. Echter is ondertussen de natuurkunde evenmin in het starre kader van vroeger gebleven: op verschillende plaatsen zijn ontdekkingen gedaan van verschijnselen, die omverwierpen wat men lang als vaststaand had beschouwd. Bepaalde onzekerheden en probabilismen hebben veroorzaakt, dat de methodes van de klassieke natuurkunde moesten verruimd worden; in zekere zin is daardoor toch een brug geslagen naar de sociale wetenschappen. Mijn stelling is, dat we door de ontwikkeling in de moderne natuurkunde gezien hebben, dat de verschillen tussen sociale wetenschappen en natuurkundige wetenschappen niet zo zeer essentieel, dan wel gradueel zijn. De toepassing van de wiskunde is in de sociale wetenschappen wel moeilijker en maakt nieuwe wiskundige technieken noodzakelijk, maar het gaat steeds om graduele verschillen.

Ik zou een algemene classificatie (qua indeling heel zeker betwisbare) willen proberen van de verschillende punten van het graduele onderscheid tussen de sociale wetenschappen en de natuurwetenschappen (1).

1. Het niet isoleerbaar zijn van het werkelijke verschijnsel in beheersbare omstandigheden.

Het isolement zelf geeft waarschijnlijk reeds een verandering, zodat het niet meer het eigenlijk bedoelde verschijnsel is, dat men waarneemt. Men doet bijv. experimenten in de groepsdynamica, waarbij men personen zekere rollen laat spelen om te zien hoe de reacties en de interacties zijn.

(1) Dr. Denucé heeft reeds een drietal oorzaken van de verschillen aangegeven: 1) het unieke van de individualiteit; wanneer dit geldt in de biologie, dan geldt dit uiteraard nog meer in de menselijke wetenschappen; 2) het niet isoleerbare van hetgeen men wenst te onderzoeken; 3) het experiment wijzigt de situatie.

Men moet daarbij echter rekening houden met de waarschijnlijkheid, dat er door het kunstmatig samenbrengen anders gereageerd wordt dan in een normaal samen zijn. Slechts zelden is een onderzoek mogelijk in verschillende, maar steeds normale omstandigheden; een ideaal voorbeeld van zo'n uitzondering is het onderzoek van ééneiige tweelingen. De beschrijving van de omgeving is in zulke experimenten, waar de omstandigheden niet beheerst kunnen worden, maar als gegeven moeten aanvaard worden, van het hoogste belang.

In de natuurkunde op een bepaald niveau moet men echter ook afzien van de volledige beheersing van de omstandigheden van het experiment: in het micro-onderzoek kan men slechts een omgeving scheppen met een zekere toevalsverdeling. Op grond van een groot aantal proeven van dezelfde soort zal men dan proberen de invloed van de niet-beheerste omstandigheden te detecteren en eventueel te elimineren.

2. Een observatie is een niet-isoleerbaar element van de omstandigheden zelf. In het bijzonder is de mens tegelijkertijd de waargenomene en de waarnemer.

Het feit zelf, dat er een observatie gebeurt is een essentieel kenmerk van de situatie en mag niet verwaarloosd worden uit de beschrijving van het resultaat. Zo kan het stellen van een vraag in het kader van een opinie-onderzoek aanleiding geven tot het activeren van een gevoel of een gedachte, die slechts latent was. De publicatie van het resultaat van een onderzoek laat ook zijn sporen na; de bekendmaking van de resultaten van achtereenvolgende politieke opiniepeilingen over een aanstaande verkiezing hebben invloed op de verdere peilingen en op het definitieve verkiezingsresultaat. De intensiteit van dergelijke beïnvloedingen is echter onmeetbaar. Men kan hierbij misschien denken aan een analogie met de onzekerheidsrelatie van Heisenberg.

Een bijzonder probleem wordt in de sociale wetenschappen gesteld door de noodzaak de subjectiviteit van de waarnemer uit te schakelen. De waarnemer is enerzijds deel van het observatie-instrument en maakt anderzijds min of meer deel uit van het waarnemingsobject. Maar ook de mate van een subjectiviteitsfout kan niet rechtstreeks gemeten worden. Speciale technieken met vergelijking van verschillende observaties zullen nodig zijn om te komen tot een schatting van en uitschakeling van dergelijke fouten.

3. Veel karakteristieken kunnen niet gerationaliseerd worden of zijn niet rechtstreeks meetbaar.

Sommige kenmerken zijn alleen latent aanwezig en niet rechtstreeks waarneembaar. We kunnen alleen vermoeden dat ze oorzaak zijn van bepaalde uiterlijke wel te observeren reacties of er toch minstens invloed

op uitoefenen. Dit stelt echter grote moeilijkheden bij het opstellen van een theorie, tenzij men zich daarbij tot louter uiterlijkheden, die soms slechts secundaire kenmerken zijn, zou beperken. Daarbij geven deze niet-waarneembare kenmerken soms aanleiding tot irrationele reacties (althans tot reacties, waaraan we geen rationele fundering kunnen geven). Zo is de veronderstelling in de economie van een zuiver rationeel reagerende mens, die steeds zal streven naar een maximum « nuttigheid », slechts een schematisering van de werkelijkheid. Bij een verder uitbouwen van de theorie zal een irrationeel gedrag moeten geïncorporeerd worden (slechts één voorbeeldje van het irrationeel gedrag: verschillende verpakkingen van identiek hetzelfde waspoeder gaven duidelijk verschillende waardeeringen na gebruik.)

Voor vele karakteristieken is het onmogelijk een metrische schaal te vinden. Dit vraagt dan wiskundige methodes, waarbij alleen de rangorde gebruikt wordt; die rangschikking behoeft zelfs niet transitief te zijn. De theorieën worden er wel ingewikkelder door.

Een veel voorkomende moeilijkheid is het discontinu karakter van de karakteristieken. Dit is echter evenmin essentieel; wel maakt het meestal het te gebruiken wiskundige apparaat iets moeilijker (bijv. differenties in plaats van differentialen).

4. De waarneembare verschijnselen zijn multivariant. En dit in twee opzichten: enerzijds is er een groot aantal kenmerken nodig om een verschijnsel te karakteriseren en anderzijds moet men rekening houden met de onderlinge afhankelijkheid van die afzonderlijke kenmerken.

Bij het onderzoek naar de onderlinge afhankelijkheden mag de factor tijd niet verwaarloosd worden. Een gevolg kan onmiddellijk optreden na de oorzaak, maar er kan ook een zeker tijdsverloop (de *lag*) tussen blijven. Ook moet onderscheid gemaakt worden tussen enerzijds gevolgen, die uitsluitend bepaald worden door het eindresultaat van de verandering, die er de oorzaak van is en anderzijds gevolgen, die ook afhangen van de wijze waarop die verandering zich voorgedaan heeft (bijv. ook van de snelheid van de verandering in het oorzakelijke kenmerk). Beschouwen we als voorbeeld het verloop van een groep prijzen, dan zullen deze onderling afhankelijk zijn, ze zullen ook afhankelijk zijn van andere prijzen, van lonen e.d. Deze prijsreacties kunnen onmiddellijk optreden ofwel na een zeker tijdsverloop (bijv. wanneer de beïnvloeding gebeurt via de kostprijs van een produktie, die slechts na enige tijd afgewerkt in de handel komt). Ook de snelheid van de veranderingen kan een specifieke invloed hebben (bijv. via min of meer speculatieve aankopen van grondstoffen). Een model van die groep prijzen zal dan ook multivariant moeten zijn, de tijd moet als veranderlijke optreden en het zal differentiaal-vergelijkingen

(voor de snelheden) moeten bevatten of nog beter differentie-vergelijkingen om rekening te houden met het discontinue karakter van een aantal van de veranderlijken.

Vrijwel ieder verschijnsel uit de menselijke wetenschappen zou alleen maar volledig kunnen beschreven worden door een groot aantal kenmerken. Een conjunctuurverloop kan onmogelijk door één maatstaf gemeten worden, een intelligentie evenmin. Het aantal verschillende karakteristieken is meestal zo groot, dat het onmogelijk is ze allemaal in te bouwen in een hanteerbaar model. We moeten een keuze maken en een schematisering opstellen, waarin alleen die karakteristieken verwerkt worden, die een zo goed mogelijke beschrijving geven van de essentiële kenmerken van het verschijnsel. De gekozen karakteristieken moeten leidinggevend en representatief zijn, zodanig, dat men kan verwachten, dat de andere, verwaarloosde, karakteristieken toch in grote lijnen dezelfde evolutie volgen als de in het model gebruikte kenmerken. Hier treedt een moeilijke interactie op tussen de keuze van de te gebruiken karakteristieken enerzijds en de theorie, die men over het werkelijke verschijnsel denkt op te stellen anderzijds. De beoordeling van het al of niet essentieel en representatief zijn van bepaalde karakteristieken kan de subjectiviteit van de onderzoeker introduceren. Welke karakteristieken, welke prijzen van welke produkten en van welke kwaliteiten van die produkten zijn bijvoorbeeld leidinggevend en representatief, zodat ze samen gebracht in een indexcijfer een voldoende nauwkeurige beschrijving toelaten van de evolutie van het prijspeil?

Naast het probleem van de keuze is er ook dat van de combinatie: welk aggregaat van de gekozen karakteristieken zal ons een maximum-informatie geven? Door een dergelijk aggregaat van waarneembare karakteristieken kan men soms ook pogen een beschrijving te geven van niet-rechtstreeks waarneembare, verborgen karakteristieken. Dit kan onderzocht worden met behulp van de correlatie-rekening en in het bijzonder met de factor-analyse. Men brengt hier alle verbindingen (correlatie-coëfficiënten) van het multivariant geheel bij elkaar en rangschikt ze zodanig, dat men kan vaststellen van welke karakteristieken de grootste correlatieve uitstraling uitgaat naar de andere kenmerken. Daarbij kan het dan voorkomen, dat men groepen aantreft van karakteristieken die in éénzelfde richting evolueren en waarvan men kan veronderstellen, dat ze als het ware gestuwd worden door één verborgen karakteristiek, één verborgen factor. De factor-analyse is ontstaan in het intelligentie-onderzoek en was bedoeld om na te gaan wat het deel « algemene intelligentie » was bij de resultaten van diverse intelligentie-tests; men heeft dit moeten uitbreiden tot het zoeken naar — overigens zo klein mogelijk aantal — verschillende

factoren, die verbonden worden met de niet-rechtstreeks waarneembare verschillende aspecten van de intelligentie.

5. De verschijnselen in de sociale wetenschappen zijn massaal.

Dit geldt ook, wanneer men een individu apart onderzoekt, omdat een sociaal wezen, de mens, niet geïsoleerd kan worden uit zijn milieu met een massa van invloeden.

Eenzijds moet men met dit massale rekening houden bij de beschrijving. De variabiliteit, eigenlijk de hele frequentieverdeling, is minstens zo belangrijk en soms zelfs belangrijker dan het gemiddelde of het meest voorkomende. Daarom was de té beperkte poging van Quetelet, die slechts sprak over de « gemiddelde mens » tot mislukking gedoemd. Sociaal-economisch is het gemiddeld inkomen veel minder belangrijk dan de inkomensspreiding; de bevolkingspolitiek zal meer rekening moeten houden met de hele leeftijdsstructuur van de bevolking dan met de gemiddelde leeftijd. Een analoog is hier misschien te vinden in het belang van de variabiliteit van de beweging van de moleculen in een gas.

Anderzijds brengt het massale mee, dat iedere observatie eigenlijk slechts een steekproef is. Het is ofwel een echte steekproef, omdat we slechts een deel kunnen onderzoeken uit de totale populatie; ofwel onderzoeken we weliswaar een hele bevolking, maar is dit in de tijd gezien en rekening houdende met alle niet gerealiseerde, maar evenwaardige mogelijkheden, toch te beschouwen als een steekproef uit een denkbeeldige populatie van mogelijkheden. De resultaten moeten dus gecontroleerd worden met de steekproeftheorie, zodat er aan de metingsfouten, die we eventueel maken, ook een steekproeffout moet toegevoegd worden. Terwijl een metingsfout zeer nauwkeurig gelimiteerd kan worden tussen wel bepaalde grenzen, kunnen de grenzen van een steekproeffout nooit absoluut, maar alleen probabilistisch bepaald worden. De zekerheid, dat het werkelijke kenmerk zal liggen binnen een bepaald interval van foutengrenzen, is niet meer absoluut; er blijft altijd een zekere kans, dat we ons vergissen. Die foutengrenzen zelf moeten overigens bepaald worden in functie van het vergissingspercentage, dat we nog als aanvaardbaar zullen geschouwen. Het meest gebruikte percentage is 5% als aanvaardbare kans op een vergissing. Indien men strenger wil zijn, bijv. omdat een vergissing ernstige gevolgen zou kunnen hebben, nemen we veelal 1% overschrijdingskans (dus kans dat de grenzen overschreden zijn, dat we ons vergist hebben). Het resultaat van een steekproef-onderzoek naar de gemiddelde lengte van een groep mensen zal de volgende vorm hebben:  $170 \pm 3$  cm, overschrijdingskans 5%. Dat wil dan zeggen, dat de werkelijke waarde tussen 167 cm en 173 cm zal liggen, maar dat er 5% kans is dat we ons vergissen en dat het aangegeven interval de werkelijke waarde niet bevat. Wensen we



de kans op vergissing kleiner te maken, dan zal echter de uitspraak, die het resultaat van het onderzoek weergeeft ook vager worden. Dat de gemiddelde lengte van een groep mensen tussen 50 cm en 300 cm gelegen is, kan nauwelijks betwijfeld worden; de overschrijdingskans is zo goed als 0, maar een dergelijke uitspraak brengt ons ook geen enkele nieuwe informatie.

6. De variabiliteit is o.m. een uiting van het probabilisme van de sociale verschijnselen.

De sociale wetenschappen onderscheiden zich hierdoor wel van de klassieke natuurkunde, waar alles streng gedetermineerd is. Maar naarmate men verder doordringt bij het onderzoek naar het wezen van de materie, naarmate men dus gaat naar de steeds kleinere, maar in massa voorkomende deeltjes, speelt het probabilisme een steeds belangrijker rol.

Men kan zowel in de moderne natuurkunde als in de sociale wetenschappen de vraag stellen of dit probabilisme een essentiële eigenschap is, of slechts een praktisch hulpmiddel, dat een gebrek aan voldoende kennis moet overbruggen. Het is wel aanlokkelijk om te zeggen, dat het probabilisme in de sociale wetenschappen essentieel is en dat het verbonden is met de vrije wil van de mens. Dit zou ons nochtans tot een filosofisch probleem brengen, waarvan de begrippen eerst nauwkeurig zouden moeten gedefinieerd worden.

Mogen we het probabilisme slechts beschouwen als een praktisch hulpmiddel, dan zouden we moeten veronderstellen, dat alles — zowel het resultaat van de worp van een dobbelsteen als iedere menselijke daad — zou kunnen voorspeld worden door een menselijk, maar enorm veel omvattend brein (de demon van Laplace); het zou slechts door het gebrek aan voldoende grote en voldoende snel werkende hersenen zijn, dat iets ons als « toevallig » en « vrij » zou schijnen.

Hoe het ook zij, wij zijn er ons wel van bewust, dat hetgeen we in praktische toepassingen als toevalsfactor beschouwen, zeker wel wat meer is dan datgene dat we misschien als essentieel probabilisme mogen beschouwen. Er is in elk geval wel een reeks beïnvloedende factoren, die eigenlijk wel beschreven zouden kunnen worden in ideale omstandigheden, maar die té complex zijn om in een model opgenomen te worden, wil dit enigszins overzichtelijk blijven.

Het gevolg is, dat we in de sociale wetenschappen vrijwel nooit kunnen spreken over deterministische wetten, maar hoogstens over wetmatigheden in de zin van waarschijnlijkheidsverdelingen. Zo kan men over het aantal kinderen per familie geen sociale wet opstellen, die zegt hoeveel kinderen een familie in bepaalde omstandigheden zal hebben. Wat men wel kan vaststellen, is dat een gezin in bepaalde sociale omstandigheden

een zekere waarschijnlijkheidsverdeling van familie-grootte zal hebben. Overigens zal men ook geen twee gezinnen vinden, waarvan de omstandigheden identiek zijn en wanneer we dan een model willen opstellen, dat probabilistisch moet zijn, dient men niet alleen rekening te houden met die waarschijnlijkheidsverdeling voor iedere familie op zich zelf, maar ook met de afwijkingen tussen de omstandigheden van de verschillende families. Het waargenomen verschijnsel moet zo beschouwd worden als het resultaat van een samengestelde waarschijnlijkheidsverdeling. (Voor « volgroeide » families wordt het een samengestelde Poisson-verdeling, d.i. het resultaat van Poisson-verdelingen, waarvan de parameter, het gemiddelde, zelf ook een waarschijnlijkheidsverdeling heeft).

#### 7. De onderlinge beïnvloedingen zijn probabilistisch.

We hebben reeds gesproken over het multivariante karakter van de verschijnselen; de onderlinge relaties, die voorkomen zijn bijna steeds probabilistisch, correlatief. (Een correlatie kan zowel de uitdrukking zijn van een eenzijdig gerichte afhankelijkheid als van een wederzijdse beïnvloeding). Indien A plaats vindt is er meestal niet slechts één bepaalde mogelijkheid B, die ook noodzakelijk moet plaats vinden; het verschijnsel A bepaalt slechts een waarschijnlijkheidsverdeling voor B.

Is het verband tussen A en de waarschijnlijkheidsverdeling van B volledig metrisch te beschrijven, dan heeft men te maken met een maatcorrelatie. In de sociale wetenschappen is het echter dikwijls zo, dat de metrische schalen onbestaande of onzeker zijn en dan kan er slechts sprake zijn van een rangcorrelatie: de volgorde van een aantal resultaten betreffende het kenmerk A zal een invloed uitoefenen op de overeenkomstige resultaten van het met A gecorreleerde kenmerk B.

#### 8. Het begrip significantie.

Indien de probabilistische wetmatigheden van de sociale wetenschappen een grote variabiliteit — grote spreiding — vertonen, kunnen de toevalligheden die zich bij het waarnemen realiseren de wetmatigheid overwoekeren. Het significantieonderzoek is er op gericht om in dergelijke omstandigheden toch de wetmatigheid te detecteren. In de gewone natuurkunde is dit niet nodig, omdat we daar door het beheersen van de experimentele omstandigheden de toevalligheden van de waarnemingen grotendeels kunnen uitschakelen; er blijven hoogstens toevallige metingsfouten, die echter klein gemaakt kunnen worden vergeleken met de veranderingen in het onderzochte kenmerk zelf.

Een bewijs, dat gebaseerd is op het begrip significantie zou ik een « bewijs uit het onwaarschijnlijke » willen noemen, naar analogie van de « bewijzen uit het ongerijmde » (waarvan ook in de lessen meetkunde in het middelbaar onderwijs reeds enkele voorbeelden gegeven worden). Bij

een dergelijk bewijs uit het ongerijmde wordt een veronderstelling verworpen, omdat ze tot een logisch onmogelijke conclusie leidt; die verwerping is absoluut en laat geen enkele twijfel. Bij het bewijs uit het onwaarschijnlijke wordt een veronderstelling verworpen, omdat ze tot een conclusie leidt, die zeer onwaarschijnlijk is; zo'n verwerping is echter niet absoluut en kan een vergissing betekenen, omdat de verkregen conclusie weliswaar zeer onwaarschijnlijk, maar niet volstrekt onmogelijk is.

Het schema van een bewijs uit het ongerijmde is als volgt:

Te bewijzen stelling: S.

Hypothese: niet-S.

Redenering op grond van streng causale afleidingen.

Is de conclusie 3 onmogelijk, dan moet de hypothese niet-S verworpen worden en blijft alleen de stelling S over.

Bij een bewijs uit het onwaarschijnlijke is het schema ongeveer hetzelfde:

Te bewijzen stelling: S.

Hypothese: niet-S.

Redenering: het opstellen van de waarschijnlijkheidsverdeling van een bepaalde karakteristiek K van het verschijnsel, indien de hypothese juist is.

Daaruit leest men de « overschrijdingskans » P af, d.i. de waarschijnlijkheid van de waargenomen waarde  $K_0$  van de karakteristiek K en van alle waarden van K, die even sterk of nog sterker dan  $K_0$  afwijken van de hypothese.

Is die overschrijdingskans P zeer klein, is dus binnen de gestelde hypothese de gevonden waarde  $K_0$  zeer onwaarschijnlijk, dan moet die hypothese niet-S verworpen worden en blijft er weer alleen de stelling S over.

Nemen we als voorbeeld de emancipatie van de vrouw, zoals ze hier vandaag tot uiting komt. We zouden de stelling willen formuleren, dat de gelijkheid nog niet bereikt is. Er zijn hier op dit ogenblik drie dames en achttien heren. Op zichzelf is dit echter nog geen voldoende bewijs voor de mindere mogelijkheid of de mindere belangstelling bij de dames. Het zou immers kunnen zijn, dat er in werkelijkheid wel minstens een gelijkheid bestaat, maar dat die overwoekerd is door de toevalligheden (en vele zeer complexe invloeden, die we als toeval zullen beschouwen van het al of niet aanwezig zijn).

Als stelling kunnen we dan formuleren, dat de kans, dat een van de aanwezigen vrouwelijk is, minder dan de helft is. De te stellen hypothese wordt dan daarentegen, dat diezelfde kans een half of meer zou zijn. De waarschijnlijkheidsverdeling, die we dan moeten bepalen is die van het aantal dames K; het is een binomiale waarschijnlijkheidsverdeling met 21 trekkingen en met een elementaire kans  $p \geq \frac{1}{2}$ . Om de overschrijdingskans P te bepalen moeten we de waarschijnlijkheden samentellen van

$K=3$  (d.i. de waarneming  $K_0$ ),  $K=2$ , 1 en 0 (dat zijn de nog sterker afwijkende mogelijkheden).

Is  $p$  juist een half dan geeft berekening :

$$\begin{aligned} \text{waarschijnlijkheid } K=0 &: \quad \left(\frac{1}{2}\right)^{21} \\ \text{» } K=1 &: \quad 21 \left(\frac{1}{2}\right)^{21} \\ \text{» } K=2 &: \quad 210 \left(\frac{1}{2}\right)^{21} \\ \text{» } K=3 &: \quad 1330 \left(\frac{1}{2}\right)^{21} \end{aligned}$$

dus de som  $P = 1562 : 2^{21} = 0,08\%$ .

(De coëfficiënten in de berekening zijn de binomiale coëfficiënten uit de ontwikkeling van Newton voor de macht van een tweeterm of ook de getallen uit de driehoek van Pascal of nog de aantallen combinaties). In de hypothese  $p \geq \frac{1}{2}$  weten we dan zeker dat de overschrijdingskans  $P \leq 0,08\%$  is.

Is dit nu té onwaarschijnlijk? Er komt hier een element van arbitraire beslissing, nl. de keuze van de « significantie-drempel », d.i. de bovengrens van wat we nog als te onwaarschijnlijk zullen beschouwen. Nemen we die drempel hoog, dan zullen we veel onwaarschijnlijk noemen — en op grond daarvan dikwijls de hypothese verwerpen en de stelling als bewezen beschouwen, — maar dan lopen we ook grote risico's om ons daarbij te vergissen. Kiezen we daarentegen een lage drempel, dan zal het niet dikwijls voorkomen, dat de overschrijdingskans  $P$  nog kleiner is en zullen we dus slechts zelden een stelling kunnen bewijzen met een dergelijk bewijs uit het onwaarschijnlijke. Wel te verstaan moet de keuze van de significantie-drempel a-priori gebeuren en mag ze niet beïnvloed worden door het gevonden resultaat. De meest gebruikte drempel in de sociale wetenschappen is die van 5%. Nemen we die drempel in ons voorbeeld, dan is de overschrijdingskans veel kleiner dan de drempel — we zeggen dan, dat het waarnemingsresultaat  $K_0=3$  « significant » is, — zodat de hypothese verworpen moet worden en de geformuleerde stelling als bewezen kan worden beschouwd.

Deze besluitvorming op grond van het onwaarschijnlijke lijkt mij een typisch kenmerk van de moderne aanpak van de sociale wetenschappen en is een noodzakelijke voorwaarde voor de toepassing van het experiment in die wetenschappen. Eigenlijk is het slechts, wat in het dagelijks leven voortdurend gebeurt en ook moet gebeuren, wil men niet in een gevaarlijke besluiteloosheid verzanden. Is het bijvoorbeeld niet zo, dat  $U$  straks op straat een zekere kans hebt te verongelukken? Die waarschijnlijkheid is echter zo klein, dat  $U$  er niet werkelijk rekening mee zult houden.

Significantie-bewijzen hebben een vrijwel onbeperkt toepassingsgebied, waarbij overigens het berekenen van de te gebruiken waarschijnlijkheidsverdeling niet altijd zo eenvoudig is als bij het gegeven voorbeeld. Wel

zijn er heel wat van de in aanmerking komende waarschijnlijkheidsverdelingen opgenomen in statistische tabellen-verzamelingen. Slechts één recent voorbeeld nog op een heel ander terrein: het onderzoek naar het auteurschap bij de geschriften van Paulus. Rekening houdende met zekere stijleigenschappen en toevallige afwijkingen er van, vergelijkt men verschillende teksten en vraagt men zich af of die afwijkingen verklaard kunnen worden door toevalligheden bij eenzelfde auteur of dat dit té onwaarschijnlijk zou zijn en dat er meerdere auteurs moeten aangenomen worden.

Hein PICARD