

HOOFDSTUK 3. METEN

3.1 WAT IS "METEN"?

Systematische waarneming van objecten in de ervaringswereld is een belangrijk onderdeel van (empirische) wetenschapsbeoefening. In veel gevallen is het handig met waarnemingsresultaten allerlei berekeningen te kunnen uitvoeren. Bijvoorbeeld de afstand tussen twee punten op aarde en het gewicht van een voorwerp in een berekening om na te gaan hoeveel energie het verplaatsen van dat voorwerp zou kosten. Daarom streeft men er dikwijls naar om waarnemingsresultaten te verkrijgen in de vorm van getallen. Onder meten verstaat men dan ook van oudsher: de grootte van iets zo scherp mogelijk bepalen en uitdrukken in een maatgetal, waarvoor de gewone rekenregels gelden. (Met gewone rekenregels wordt in dit geval bedoeld, die welke gelden voor de reële getallen.)

Voor sommige eigenschappen van objecten, zoals lengte en massa, bestaat al een zeer lange traditie van meten. De manier waarop dat gebeurt lijkt dan ook heel "natuurlijk". T.a.v. andere eigenschappen, b.v. "nut", "intelligentie" of "macht" worden methoden van "meting" gebruikt die betrekkelijk nieuw zijn, en minder vanzelfsprekend lijken. Onder meer door de ontwikkelingen in de sociale wetenschappen, waar men naar het voorbeeld van de natuurwetenschappen eveneens "exact" te werk wilde gaan, verruimde zich het begrip meten tot "op objectieve wijze in schaal brengen". Daarbij wordt met "schaal" welliswaar nog steeds "getalschaal" bedoeld, maar deze behoeft niet langer metrisch te zijn, d.w.z. de "gewone" rekenregels behoeven niet langer te gelden (De Groot, 1970, blz. 226). Welke rekenregels voor die getallen gelden, is afhankelijk van het soort schaal, waarin zij gelezen moeten worden. Onder "objectief" wordt in dit verband verstaan: "niet in belangrijke mate afhankelijk van degene die de meting uitvoert".

Een verdere consequentie van deze ontwikkeling is, in de definitie van het begrip "meten" niet langer te spreken van getallen, maar van "code" of "meetwaarde" o.i.d.. Met name voor metingen in de z.g. "nominale schaal" (zie hieronder) kan men immers de aan een object toegevoegde code nauwelijks nog een getal noemen. Voorbeeld: een waarneming waarbij aan mannen een code 0 en aan vrouwen een code 1 wordt toegekend t.b.v. verwerking door middel van ponskaarten. Het kunnen uitvoeren van een meting houdt in, dat men beschikt over een bepaalde voorstelling, een model van de ervaringswereld, waaruit men een objectieve wijze van toekenning van meetwaarden aan bepaalde objecten heeft afgeleid.

”Meten” is het doen van uitspraken over objecten in de ervaringswereld volgens een bepaalde, objectieve, procedure.

De rekenregels, die voor de meetwaarden gelden, hangen af van de gehanteerde schaal en daarmee van het model, waarop de meetmethode berust.

Meetresultaten zijn ”data” als gedefinieerd in hoofdstuk 1, maar niet alle data zijn het resultaat van een korrekte meting.

3.2 VERWIJSMIDDELEN, VARIABELEN, CONTINU EN DISCREET

Om uitspraken te kunnen doen over objecten in de ervaringswereld is het nodig om op enigerlei wijze naar die objecten te kunnen verwijzen. Een ”verwijsmiddel” is: naar het bedoelde object wijzen, terwijl men erover spreekt. Het geven van een naam aan een objekt, b.v. een eigennaam voor een persoon, vormt een tweede mogelijkheid. Ook kan men gebruik maken van een ”coördinatenstelsel” in enigerlei vorm, een verwijsmiddel dat gebaseerd is op afspraken die het mogelijk maken een objekt in het tijd/ruimte-universum te localiseren. B.v. de aanduiding van een plaats op aarde in lengte- en breedtegraden. (Vgl. Nuchelmans, 1978, blz. 94 e.v.)

Aan objecten kunnen gewoonlijk verschillende waarneembare grootheden worden onderkend zoals grootte, gewicht, ouderdom e.d. Meetbare grootheden worden ”variabelen” genoemd.

Bij een z.g. ”continue” variabele vormen alle denkbare meetwaarden een continuum: voor elke twee willekeurige waarden is het mogelijk een derde waarde te vinden, die ertussenin ligt. Voorbeelden zijn tijd, gewicht, lengte, temperatuur en vele andere variabelen die in een reeel getal uitgedrukt kunnen worden.

Variabelen, waarbij tussenliggende waarden niet altijd denkbaar zijn, worden ”discreet” genoemd. Voorbeelden zijn aantallen (b.v. het aantal mensen in een zaal), de sexe van personen en de plaats die iets of iemand inneemt in een rangorde.

Bij het meten van continue variabelen, hoe nauwkeurig ook, moet altijd afgerond worden. Variabelen worden daarom altijd discreet gemeten.

3.3 NIVEAUX VAN METEN

Metten kan gebeuren in verschillende schaaltypen. Ieder schaaltype correspondeert met een bepaald ”meetniveau”.

Tussen deze meetniveaux bestaat een bepaalde relatie: ieder hoger meetniveau sluit alle lagere niveaux in, in die zin dat aan alle voorwaarden die gelden voor de lagere niveaux voldaan moet zijn, plus nog tenminste een andere voorwaarde. Bij een laag meetniveau is het niet goed mogelijk de mate waarin een objekt een bepaalde eigenschap heeft aan te geven. Op een hoog meetniveau kan dat wel.

De meetniveaux die gewoonlijk worden onderscheiden zijn de volgende (Van den Ende, 1971, blz. 16 e.v.):

1. De nominale schaal (het nominale niveau).
2. De ordinale schaal (het ordinale of rangordeniveau).
3. De intervalschaal (het intervalniveau).
4. De ratioschaal (het rationiveau).

3.3.1 De Nominale Schaal

Bij meting op dit niveau is er nog geen "meer of minder" of "hoger of lager". De aan een objekt toegevoegde code bestaat dikwijls uit letters. Als getallen worden gebruikt zijn dit slechts kengetallen. Voorbeelden:

- Het classificeren van een groep mensen in personen die "man" en personen die "vrouw" zijn. I.p.v. man en vrouw kan men ook een code geven, b.v. man = code 0 en vrouw = code 1.
- Het determineren van planten.

Metten op nominaal niveau houdt dus in, dat men beschikt over criteria, volgens welke men codes toekent aan objecten. Daarbij gelden de volgende eisen:

1. Er zijn tenminste twee verschillende codes.
2. Objecten, die voldoen aan hetzelfde criterium krijgen dezelfde code, objecten die voldoen aan verschillende criteria krijgen verschillende codes.
3. Ieder objekt krijgt niet meer dan een code. (De klassen sluiten elkaar onderling uit.)
4. Voor ieder objekt is er een code. (De classificatie is uitputtend.)
5. Toekenning van de code aan het ene objekt is onafhankelijk van de toekenning van de code aan een ander objekt. (De klassen zijn onafhankelijk.)

Met de verkregen getallen of codes kan men niet zinvol berekeningen uitvoeren, in die zin dat aan de resultaten van de berekeningen conclusies verbonden mogen worden m.b.t. de geclassificeerde objecten. In bepaald sociaal wetenschappelijk onderzoek kan men wel aantreffen, dat de proefpersonen de gemiddelde score 0,58 hadden. Enige betekenis heeft het resultaat van deze berekening echter niet, aangezien men het genoemde getal pas weer kan begrijpen door terug te berekenen, dat van iedere 100 personen er 58 code 1 gehad moeten hebben. (Dat dit terugrekenen mogelijk is komt overigens niet door de nominale schaal, maar doordat de classificatie uit slechts twee groepen bestaat.)

Bij meten op nominaal niveau kan men, zonder de betekenis van het meetresultaat te veranderen, de toegevoegde codes vervangen door elke andere, mits dit op eenduidige wijze gebeurt. D.w.z. volgens een z.g. 1-1 transformatie, waarbij voor iedere code precies een andere code in de plaats wordt genomen.

3.3.2 De Ordinale Schaal

Het ordinale meetniveau is het niveau dat juist hoger ligt dan het nominale. Ging het bij meten op het nominale niveau nog slechts om het classificeren, op ordinaal niveau is er ook een eendimensionale ordening. Toegekende codes (die niet per se getallen behoeven te zijn) hebben de betekenis van rangnummers. Voorbeelden:

- Het toekennen (volgens een objectieve methode) van de letters "A" ("excellent"), "B" ("good") t/m "F" ("fail") aan bepaalde leerprestaties, examenresultaten o.i.d..
- Het rangschikken van het nut van bepaalde goederencombinaties in de economische theorie van het consumentengedrag (volgens Pareto, vgl. Heertje, 1970, blz.86). Hierbij laat men toe, dat verschillende objecten (i.c. goederencombinaties, b.v. 2 appels en 3 peren resp. 3 peren en 2 appels) eenzelfde ranggetal ("nut") toebedeeld krijgen.
- Het rangschikken van taken in de industrie naar hun zwaarte.

Metten op ordinaal niveau houdt dus in, dat men behalve te classificeren (zoals bij het meten op nominaal niveau) ook nog rangschikt, en dienovereenkomstig codes toekent aan de objecten. Behalve de eisen, reeds genoemd voor meten op nominaal niveau, geldt de eis:

6. Objecten, die van een bepaalde eigenschap "meer" hebben, krijgen ook een code, die een hogere plaats in de rangorde van de codes aan geeft.

Als voor de toe te kennen code getallen worden gebruikt, dan zijn dat ranggetallen, waarmee men niet zinvol kan optellen, vermenigvuldigen

e.d.. De rangorde binnen de codes kan van beneden naar boven verlopen (b.v. de 1e, 2e en 3e plaats in een sportklassement), of juist andersom (b.v. schoolcijfers: een 10 is de hoogste score, een 1 het laagst).

Bij meten op ordinaal niveau kan men, zonder de betekenis van het meetresultaat te veranderen, de gehanteerde reeks (getal-)codes wijzigen volgens een vast patroon, nl. door een willekeurige andere consequent ("monotoon") stijgende of dalende reeks.

3.3.3 De Intervalschaal

Het intervalniveau van meten is het niveau, dat juist boven het ordinale ligt. Ging het bij de ordinale schaal nog slechts om een ééndimensionale ordening, bij de intervalschaal kan er ook zinvol gesproken worden over hoeveel hoger of lager een bepaalde meetwaarde is vergeleken met een andere. Afstanden tussen meetpunten ("intervallen") kan men dan zinvol vergelijken. Voorbeelden:

- Het toekennen van een jaartal aan gebeurtenissen komt gewoonlijk neer op een meting op intervalniveau. Jaartallen drukken het tijdsverloop in jaren uit tussen een door conventie bepaalde gebeurtenis (b.v. de geboorte van Christus) en een andere.
- Temperatuurmeting volgens Celsius. Hierbij stelt iedere graad verhoging van de temperatuur een hoeveelheid toe te voeren warmte voor (nl. 1/100 van de warmte, nodig om een zekere hoeveelheid nog juist niet bevroren water bij een druk van 1 atmosfeer te verwarmen tot nog juist niet kokend).

Naast de eisen die gelden voor meten op ordinaal niveau geldt de eis:

7. Er wordt gebruik gemaakt van een vaste meeteenheid.

Hierdoor kan elk gevonden verschil tussen twee meetwaarden uitgedrukt worden in een aantal meeteenheden. Het is dan ook gebruikelijk om bij meting op intervalniveau codes in de vorm van getallen toe te kennen. Met de verschillen tussen deze getallen (de intervallen) kan men zinvol berekeningen uitvoeren, zoals verhoudingen bepalen en gemiddelden nemen.

Transformatie van de ene reeks codes naar een andere kan, zonder de betekenis van het meetresultaat te veranderen, volgens ieder rechtlijnig verband.

3.3.4 De Ratioschaal

Bij meting op dit ("hoogste") niveau heeft het niet alleen zin over de verschillen tussen de meetwaarden (de intervallen) te spreken, maar komt ook betekenis toe aan de verhoudingen (ratio's) tussen de meetwaarden. Voorbeelden:

- Temperatuurmeting volgens Kelvin. Deze vond, dat er een absoluut laagste temperatuur bestaat, waar beneden men geen warmte meer aan een objekt kan onttrekken. Dit noemde hij het absolute nulpunt (ca. -278 graden Celsius).
- Meting van afstand, snelheid, gewicht en vele andere natuurkundige grootheden. Zo ook het bepalen van de hoeveelheid informatie die een boodschap bevat zoals weergegeven in hoofdstuk 1.4.
- meting van tijdsduur vanaf een niet willekeurig tijdstip. B.v. de leeftijd van personen.

Meting op rationiveau houdt in, dat naast de eisen, reeds genoemd voor het intervalniveau, geldt:

8. Er wordt gebruik gemaakt van een niet-arbitrair ("natuurlijk") nulpunt.

De toe te kennen codes zijn reële getallen, waarvoor alle uit de wiskunde bekende rekenregels gelden.

Transformatie van de ene reeks codes naar de andere kan, zonder het meetresultaat te veranderen, alleen volgens een vaste vermenigvuldigingsfactor.

3.4 MEETNIVEAU EN ONTWIKKELINGSSTAND VAN EEN WETENSCHAP

Hierboven was sprake van de ontwikkeling van het begrip "meten" in die zin, dat tegenwoordig, althans in de sociale wetenschappen, niet alleen van "meten" wordt gesproken indien aan objecten getallen worden toegevoegd waarvoor de gewone rekenregels gelden (meten op rationiveau). Het weergeven van deze ontwikkeling houdt niet in dat naar mijn mening meten op lager dan rationiveau nu de voorkeur zou verdienen boven datgene wat vroeger onder "echt meten" werd verstaan. Integendeel: om te kunnen meten op rationiveau zijn hoogontwikkelde modellen, met een stevige empirische onderbouwing nodig. Uit de gegeven voorbeelden mo-ge dat blijken.

Grofweg kan men zeggen, dat op hoe hoger niveau in een bepaalde tak van empirische wetenschap gemeten kan worden, des te meer ontwikkeld die discipline in empirisch opzicht is (des te meer kennis in die tak van wetenschap is verworven). Gekonstateerd kan worden dat aan de minst ontwikkelde kant de sociologie staat, op een wat hoger niveau de psychologie en de economie, en de natuurwetenschappen zijn in dit opzicht het meest ontwikkeld.

3.5 METEN EN BEGRIPSVORMING

De definitie van "meten" is, overeenkomstig het gebruik in de sociale wetenschappen, zo gekozen dat het op alle vier de schaaltypen van toepassing is. Dit is niet noodzakelijk.

Hempel (1969, blz. 50 e.v.) spreekt van "begripsvorming" als algemene term. Hij onderscheidt vervolgens de procedures van:

1. Classificeren
2. Niet-metrisch ordenen
3. Meten

De Groot (1970, blz 226) vermeldt het standpunt van Torgerson (1969, hoofdstuk 2;3), die alleen van "meten" spreekt waar men kan zeggen met een "schaal" te doen te hebben die van "laag" naar "hoog" loopt. Meten begint bij hem dus bij de ordinale schaal, en de nominale schaal is geen "meetschaal".

Naar mijn mening zijn deze verschillen van terminologische aard en hanteren genoemde schrijvers in beginsel hetzelfde model met betrekking tot het doen van waarnemingsuitspraken volgens een bepaalde objectieve procedure.

3.6 VALIDITEIT EN BETROUWBAARHEID

Voor het beoordelen van de doelmatigheid van metingen zijn de criteria "validiteit" (of "geldigheid") en "betrouwbaarheid" van belang (vgl. De Groot, 1970, blz. 138)

De betrouwbaarheid houdt verband met de vraag: "in hoeverre wordt iets gemeten?", terwijl validiteit de vraag betreft in hoeverre iemand meet, wat hij meten wil. De vraag of een meting geldig is, is pas zinvol, als de betrouwbaarheid van de meting voldoende is. De betrouwbaarheid heeft onder meer betrekking op de vraag naar de mate, waarin toevallige fouten worden uitgesloten en op de nauwkeurigheid van de meting.

Veel toevallige fouten zijn menselijke fouten, waaronder waarnemings- (b.v. aflees-) en rekenfouten. (Berekeningen kunnen immers onderdeel uitmaken van een meetprocedure.) Maakt men afleesfouten, dan werken deze in latere berekeningen vanzelfsprekend door. De statistiek helpt de fouten die daarbij optreden te schatten. Toevallige fouten kunnen aan het licht komen door dezelfde meting op een ander tijdstip te herhalen, of door de meetmethode te variëren. Zijn er bijvoorbeeld voor een bepaalde

meting menselijke beoordelingen nodig, dan ligt het voor de hand verschillende beoordelaars te gebruiken. (De Groot (1970, blz. 244 e.v.) geeft een lijst van voorzorgen en controles die bij het toepassen van menselijke beoordelaars genomen c.q. uitgevoerd kunnen worden.)

Soms is het mogelijk schattingen te maken van de betrouwbaarheid. B.v. in meerkeuzetests voor leerprestaties kan men de formule van Kuder-Richardson (Kuder/Richardson, 1937, blz. 159) hanteren, die aangeeft in hoeverre de onderdelen van een test elkaar ondersteunen. Op deze wijze opgevat betekent betrouwbaarheid "interne consistentie". Een verwante manier bestaat daaruit dat men een test in twee helften splitst en nagaat, in hoeverre de resultaten op beide testen overeenkomen ("split-halves reliability" of halveerbetrouwbaarheidscoëfficiënt, De Groot, 1970, blz. 288).

Een meetprocedure is nauwkeuriger in algemene zin, naarmate een meetuitkomst gemiddeld meer informatie verstrekt (vgl. De Groot, 1970, blz. 279). Aldus geformuleerd is dit uitsluitend een kwestie van de gebruikte meetschaal en wel van de mate van differentiatie van meetuitkomsten, die de gebruikte schaal mogelijk maakt. In aansluiting op onze definitie van "informatie" in hoofdstuk 1.4 is de informatie die een meetuitkomst verstrekt gelijk aan:

$$(1) \quad -^2 \log \frac{1}{p(a)}$$

waarin $p(a)$ de waarschijnlijkheid van de meetuitkomst (a) voorstelt. Bij n mogelijke meetuitkomsten is de gemiddelde informatie:

$$(2) \quad \sum_{i=1}^n p(i) \cdot -^2 \log \frac{1}{p(i)}$$

d.i. de som van de informatie die iedere meetuitkomst geeft, steeds vermenigvuldigd met de kans op die meetuitkomst.

Voor het eenvoudigste geval, twee mogelijkheden, a en b ($n=2$), is de gemiddelde informatie dan:

$$(3) \quad p(a) \cdot -^2 \log \frac{1}{p(a)} + p(b) \cdot -^2 \log \frac{1}{p(b)}$$

Voor dit eenvoudige geval is gemakkelijk in te zien, dat deze gemiddelde informatie maximaal is, als $p(a)=p(b)$. Dit betekent, dat beide gebeurtenissen even vaak voorkomen. Aangezien $p(a)+p(b)=1$, is $p(a)=p(b)=0,5$. De gemiddelde informatie is dan $\log 2 = 1$. Neem b.v. eens aan, dat $p(a)=0,9$ en $p(b)=0,1$ is. Dan doet zich meestal (in 90% van de gevallen) de triviale mogelijkheid van $p=0,9$ voor, met een informatie van slechts $\log 10/9 = \text{ca. } 0,152$ en slechts een op de 10 gevallen levert $\log 10 = \text{ca. } 3,32$ op.

Men kan aantonen, dat ook in het algemene geval van n mogelijkheden de gemiddelde informatie maximaal is als voor iedere mogelijkheid (i), $p(i)$ gelijk is. Bij n mogelijkheden is dan $p(i) = 1/n$. Derhalve is in het optimale, algemene geval de gemiddelde informatie ($p(i)$ substitueren in (2)):

$$(4) \quad n \cdot \frac{1}{n} \cdot \log n = \log n$$

M.a.w. de theoretisch maximale nauwkeurigheid (of differentiatie) is gelijk aan $\log n$, en dus voor continue variabelen oneindig groot.

De hier uiteengezette differentiatie maat voor meetschalen heeft als voordeel, dat hij alleen afhangt van het aantal mogelijkheden (klassen) en de relatieve frequentie daarvan, en voor alle schaaltypen, van nominaal tot en met metrische, gebruikt kan worden.

Bij metrische (interval- en ratio-) schalen ligt het volgens De Groot (1970, blz. 280) voor de hand, tevens gebruik te maken van de "afstanden" tussen de schaalwaarden, waarover men bij deze schalen immers zinvol kan spreken. Vaak wordt daarom gewerkt met de "variantie" of met de wortel daaruit, de "standaardafwijking" als differentiatie maat. (Voor de relaties tussen deze maten en de hierboven gegeven maat wordt verwezen naar Shannon en Weaver, 1949, blz. 54 en Attneave, 1959, blz. 95.)

De vraag naar de validiteit van een meting komt aan de orde, nadat de vraag van de betrouwbaarheid bevredigend is beantwoord. Als iemand iets meet, meet hij dan ook, wat hij meten wil?

Ook de validiteit kan weer onder verschillende gezichtspunten worden bezien. Een eerste variant is de prediktieve validiteit. Dit is de mate, waarin de meting (of meetmethode) de variabele, die men eigenlijk wil bepalen, de "criteriumvariabele", voorspelt. Als voorbeeld kan dienen de ouderdomsbepaling van voorwerpen d.m.v. de koolstof-isotoop (of C^{14}) methode (zie Grote Winkler Prins, 1969 (Amsterdam), dl. 8, blz. 214). Nadat de mogelijkheid om op deze wijze de ouderdom van b.v. historische voorwerpen te bepalen in principe was ontstaan, rees natuurlijk de vraag naar de validiteit van dit soort metingen. Bij het onderzoek daarnaar ging men als volgt te werk: men bepaalde de ouderdom van voorwerpen, waarover historici een uitgesproken idee hadden. M.a.w. de criteriumvariabele was hier: "ouderdom volgens sommige historici". Toen uit de ervaringen bleek, dat de koolstofmethode een goede prediktievevaliditeit bezat en qua betrouwbaarheid superieur was, konden de rollen worden omgedraaid: thans valideert men het oordeel van historici aan de koolstofmethode. Deze ontwikkeling heeft iets paradoxaals, en doet denken aan het verhaal van de man die zich aan z'n eigen laarzen uit het moeras trok. Vandaar dat De Groot (1979, blz. 268) spreekt van "bootstrap"- of "Muenchhausen"-effect.

Prediktieve validiteit is slechts een substituuut, omdat de criteriumvariabe-

le ook weer een variabele is, die gemeten moet worden, etc.. De situatie is natuurlijk lang niet altijd zo gunstig als bij de koolstofmethode.

Denk b.v. aan examens, die uit moeten maken of leerlingen geschikt zijn voor een bepaalde opleiding. Een (empirische) criteriumvariabele voor gebleken geschiktheid is moeilijk te vinden, onder meer, omdat het enige jaren kan duren voordat vaststaat, in hoeverre voor het examen geslaagde studenten de opleiding al dan niet met succes hebben doorlopen. Behalve dat deze informatie wat laat komt, is het bovendien moeilijk uit te maken in hoeverre het niet slagen gedurende de opleiding nu een gevolg is van 'ten onrechte geschikt verklaard zijn', dan wel van andere factoren, zoals de kwaliteit van de opleiding of wellicht de geringe doelmatigheid van het eindexamen daarvan. Nog moeilijker is het een criteriumvariabele te vinden voor gebleken ongeschiktheid. De leerlingen die het toelatingsexamen niet gehaald hebben, worden immers afgevoerd, zodat hun 'ongeschiktheid' niet meer gefalsificeerd kan worden! In zijn "Selektie voor en in het hoger onderwijs" noemt De Groot dit de "valkuil voor psychologen" (De Groot, 1972, blz. 129). Dit probleem, meer algemeen geformuleerd: het probleem van de selekte (d.w.z. niet representatieve) groepen, speelt met name ook criminologen parten. Een beroemd voorbeeld vormt Lombroso's "theorie van de geboren misdadiger", die slechts gefundeerd was op metingen van een selekte groep gevangenisbewoners. (Zie Mannheim, 1965, blz 130. Een oplossing voor deze problematiek levert soms het mede-onderzoeken van een z.g. "contrôlegroep".)

Waar het bij de validiteitsvraag om gaat, is de relatie tussen het meetresultaat en de gemeten variabele. Dit aspect noemt De Groot: "begripsvaliditeit". Om tot bepaling van deze begripsvaliditeit te komen zijn weer een aantal nieuwe substituten ontwikkeld, waaronder "soortgenootvaliditeit" (samenhang met vergelijkbare meetprocedures) en "inhoudsvaliditeit" (samenhang met een verzameling van allerlei of zelfs alle beschikbare meetprocedures).

De problematiek van de validiteit overziend, kom ik tot de volgende conclusie. De validiteitsvraag is de vraag van de samenhang tussen de gevonden meetwaarden en de eigenschappen van het objekt, die iemand te weten komen wil. Wanneer deze wens van precieze aard is, d.w.z. dat men over het objekt al het nodige weet, anders gezegd, dat men over het objekt een realistisch model, een verantwoorde theorie heeft, is deze samenhang gewoonlijk aanwezig. Bevindt de kennis over het objekt zich in de schemer, zoals in de sociale wetenschappen dikwijls het geval is, dan is de validiteitsvraag actueel. (Voor een overzicht van sommige problemen zie Campbell, 1957, blz. 297 e.v.)