

VAN DE GESCHIEDENIS LEERT MEN?

IMPACT, PUBLICATIEDRUK EN FAIR PLAY
IN DE (PSYCHOLOGISCHE) WETENSCHAP

“Als de onderzoeker, ten eerste, naar vermogen streeft naar objectiviteit en zuivere redeneringen en als hij zich, ten tweede, houdt aan de voorschriften van eerlijk onderzoek en open publicatie, die in dit boek zijn beschreven, dan treedt geen verwarring op. Anderen zullen dan desgewenst zijn werk kunnen overdoen en/of kunnen onderscheiden waar, zijns ondanks, invloeden van andere ideologieën dan die van de objectieve wetenschap in zijn definities zijn ingeslopen, of waar hij methodologische fouten heeft gemaakt.”¹

Dit citaat is al meer dan vijftig jaar oud en komt uit het boek *Methodologie* van Adriaan de Groot (1914-2006). De Groot is als hoogleraar aan de universiteiten van Amsterdam (lang) en Groningen (kort) verbonden geweest. Menigeen ziet hem nog altijd als meest invloedrijke Nederlandse psycholoog aller tijden. Zo geldt zijn proefschrift *Het denken van den schaker*² internationaal als intellectuele voorloper van de cognitieve psychologie.³ Een andere significante, meer nationale bijdrage is geweest dat hij sociaal wetenschappers in het algemeen en psychologen in het bijzonder heeft aangespoord hun beweringen empirisch te staven. In *Methodologie* heeft hij – ruim vijftig jaar geleden dus al – glashelder de spelregels uiteengezet hoe dat zo fair mogelijk te doen.

Generaties studenten sociale wetenschappen zijn opgevoed met zijn empirische cyclus en bijbehorende onderzoeksadagium: “Als ik iets weet, kan ik iets voorspellen; kan ik niets voorspellen, dan weet ik niets.”⁴ Die empirische cyclus begint met *observatie*: het systematisch verzamelen van empirisch feitenmateriaal en het vormen van hypothesen. De tweede fase is die van *inductie*: de hypothesen worden preciezer geformuleerd. Dan volgt de fase van *deductie*: het op basis van die hypothesen afleiden van

VITTORIO BUSATO

werd geboren in 1965. Is psycholoog en publicist. Is onder meer hoofdredacteur van *De Psycholoog*, het blad van het Nederlands Instituut van Psychologen (NIP). Schreef diverse boeken, onder andere de biografie *Weg met Piet Vroon* (Amsterdam University Press / Salomé, 2004) en het e-boek *Psychologie al dente* (Fosfor, 2014). Adres: www.vittoriobusato.nl.

toetsbare voorspellingen. Die voorspellingen worden vervolgens *getoetst* aan nieuw empirisch materiaal. De uitkomsten worden daarna, als vijfde fase, *geëvalueerd* op algemene theoretische geldigheid – waarna de cyclus weer van voor af aan begint.

In dit proces is het volgens De Groot cruciaal om expliciet onderscheid te maken én te handhaven tussen exploratie en toetsingsonderzoek. Zijn gebod: hypothesen stel je *vooraf* op en toets je aan *nieuw* materiaal. “Wie een exploratie in de verslaggeving voorstelt als een toetsingsonderzoek door te doen alsof de hypothese al scherp gesteld was voordat het onderzoek begonnen was – wat helaas vrij gemakkelijk gedaan kan worden – maakt zich schuldig aan een ernstig vergrijp tegen de sociale ethiek van de wetenschapsbeoefening. In de ‘open’ communicatie tussen wetenschapsbeoefenaars wordt erop gerekend, dat dergelijke misleidingen niet voorkomen.”⁵

Doelstelling van exploratief onderzoek, zo benadrukt De Groot, is hypothesevorming. Hypothesen worden er niet getoetst; die zijn zoals bij toetsingsonderzoek vooraf niet scherp geformuleerd, maar louter geëxploreerd om in lijn met vooraf bestaande theoretische bevindingen tot scherpere hypothesen te komen. “Het komt te vaak voor, dat ‘exploratie’ een eufemisme is voor onnodige contaminatie in een onderzoek, dat veel beter systematisch objectief-descriptief had kunnen worden opgezet. (...) Het gevolg is, dat een aanvankelijk als toetsingsonderzoek bedoeld, maar als zodanig slecht uitgevoerd project wordt gepresenteerd als ‘exploratief onderzoek’ – ter redding uit methodologische nood.”⁶

Exploratief onderzoek is volgens De Groot *vooronderzoek*. Ontbreekt het vervolg, namelijk exacte theorie- en/of hypothesevorming en -toetsing, dan heeft het nauwe-

lijks waarde. “Aan ‘gemengde’, theoretisch gerichte onderzoeken moet de reeds eerder genoemde eis worden gesteld, dat de onderzoeker *de verschillende vormen en werkwijzen in hun betekenis goed uit elkaar houdt*. De betekenis van het significantieniveau bijvoorbeeld hangt sterk af van de vraag of we met een toetsings- of een exploratief onderzoek te doen hebben. Exploreert men, bijvoorbeeld, net zo lang totdat men iets ‘significants’ heeft gevonden, dan is dit, door de voorafgaande selectie, niet meer significant in de statistische zin.”⁷

DE FRAUDULEUZE DUIM

Van de geschiedenis leert men, zo luidt het cliché. Wie evenwel een aantal recente gevallen van wetenschapsfraude beschouwt, zal toch zo zijn of haar twijfels hebben over de geldigheid ervan. Als “onderzoekers” zoals Diederik Stapel (Tilburg University), Dirk Smeesters (Erasmus Universiteit Rotterdam), Don Poldermans (eveneens Erasmus), Marten Bax (Vrije Universiteit Amsterdam), Peter Paul Rijkema (Universiteit van Amsterdam), Patrick van Calster (Rijksuniversiteit Groningen) en een anoniem gebleven (waarom eigenlijk?) Belgische hoogleraar reumatologie aan de Rijksuniversiteit Leiden *Methodologie* al bestudeerd hebben, dan is duidelijk dat ze het niet zo nauw hebben genomen met dat empirisch staven en met voorschriften van eerlijk onderzoek en open publicatie – om over het expliciete onderscheid tussen exploratie en toetsend onderzoek maar te zwijgen.

Wat hebben zij ook alweer op hun kerfstok? Sociaal psycholoog Stapel bleek een groot deel van zijn oeuvre uit de spreekwoordelijke duim gezogen te hebben; hij kwam met stip de internationale top tien van grootste wetenschappelijke bedriegers binnen. Zijn Belgische collega en hoogleraar consumentengedrag Smeesters, eveneens sociaal psycholoog, werd betrappt op gegoochel met data in een drietal inmiddels teruggetrokken artikelen.

Internist en hoogleraar Poldermans verzoonde gegevens van zo’n tweehonderd patiënten. Hoogleraar politieke antropologie Bax werd lang na zijn emeritaat alsnog beticht van wetenschappelijk wangedrag (overigens zonder gevolgen, alsof zijn frauduleuze duimzuigen verjaard zou zijn); hij zou onder meer niet bestaande publicaties, niet geleverde prestaties en nooit toegekende onderscheidingen in officiële documenten hebben opgevoerd en zich voor zijn publicaties op oncontroleerbare bronvermeldingen hebben gebaseerd.

Rijkema, hoogleraar algemene rechtsleer, zou flinke stukken van een onder zijn naam gepubliceerd handboek hebben overgenomen uit het boek van zijn voorganger zonder hem daar de credits voor te geven; als schending van de wetenschappelijke integriteit gold Rijkema’s handelwijze in tegenstelling tot die van Stapel, Smeesters, Poldermans en Bax uiteindelijk niet, wél als zeer onzorgvuldig handelen.

Het overschrijfwerk van de hoogleraar criminologie Van Calster kostte hem daarentegen weer wel zijn baan; de Vrije Universiteit Brussel trok zijn doctorstitel uit 2005 terug omdat grote delen van zijn dissertatie zonder bronvermelding waren overgenomen uit een veelgebruikt managementhandboek – de druppel voor zijn Groningse werkgever.

En A.S., die anonieme reumatoloog, werd ontslagen nadat collega's fraude in haar laboratoriumonderzoek hadden ontdekt. A.S., zo bleek, sloop geregeld 's nachts het lab binnen om bloedstalen van patiënten te manipuleren, om zo te verbloemen dat een eerder door haar ontworpen test niet deugde – gelukkig zijn haar patiënten, evenals die van Poldermans, nooit in gevaar geweest.

PUBLICATIEDRUK

Wat beweegt wetenschappers, die weten dat ze via het (althans bedoeld als) zelfreïnigende vermogen van de *peer review* grote kans lopen ontmaskerd te worden, ertoe toch te frauderen? Is de oorzaak de hoge academische publicatiedruk zoals Stapel zelf als een verklaring voor zijn fraude aangaf?⁸ Moet de oorzaak voor die fraude wellicht worden gezocht in te veel ijdelheid, egotripperij of narcisme om desnoods dan maar via fraude in bladen met de hoogste *impact*-factor te komen, hét toverwoordje waar het in de wetenschap tegenwoordig om draait?⁹

Over die publicatiedruk, bijbehorende ratrace en nadruk op kwantiteit wordt wel vaker geklaagd – bijvoorbeeld in de krant *De Morgen* door een grote groep Belgische wetenschappers¹⁰; grotendeels op basis van aantallen publicaties en aantallen studenten en doctoraten verdeelt de Vlaamse overheid immers het geld over de academische instellingen (de beruchte outputfinanciering). De verleiding om de empirische werkelijkheid wat mooier voor te spiegelen om zo de kans op publicatie te vergroten, zou daarom aantrekkelijker zijn – temeer ook daar tijdschriftredacties veel meer geneigd zijn “significante” resultaten te publiceren.

Ook kunnen er van die druk perverse prikkels uitgaan, of beter gezegd van het dus eveneens tot de Alma Mater doorgedrongen marktdenken. Als universiteiten onder meer worden afgerekend op aantallen afgestudeerden, ligt het gevaar van academische inflatie op de loer; een norm voor een tentamen is zo bijgesteld. Als universiteiten meer geld krijgen voor aantallen publicaties, dan moet er niet vreemd van worden opgekeken dat het management daarop zal aansturen. Zo vertelde een hoogleraar mij die was gevraagd om in nota bene Tilburg op een hoogleraarspositie te solliciteren, dat hij/zij verleid werd “met bonussen van 5K” bovenop het salaris voor elke publicatie in een toptijdschrift – een praktijk die zij/hij overigens uniek voor die universiteit noemde.

Het is psychologie van de koude grond, maar narcisme zal vermoedelijk evenmin zijn uit te sluiten als mogelijke (mede)veroorzaker van wetenschapsfraude. Zei weten-

schapsfilosoof *sir* Karl Popper (1902-1994), overigens een goede kennis van De Groot, ooit dat het er in de wetenschap niet zozeer gaat om *wie* er iets zegt maar *wat* er wordt gezegd, dat lijkt in de huidige wetenschappelijke afrekencultuur steeds minder op te gaan. Zoek op het internet maar eens naar cv's van in de wetenschap werkzame psychologen; menigeen vermeldt daarin de impactcijfers van de tijdschriften waarin ze publiceren, alsmede aantallen citaties van hun artikelen – cijfers dus die over de tijd veranderen. Is dergelijke borstklopperij de manier om uit te drukken dat je “meetelt”? Zou je zo je werkelijke impact kunnen uitdrukken in relatie tot de opdracht van de psychologie: mensen begrijpen en helpen?

Hoe dan ook, die publicatiedruk of een teveel aan ijdelheid kunnen nooit dé verklaring voor wetenschappelijke fraude zijn. Er zijn immers heel veel meer al dan niet ijdele wetenschappers die last hebben van die vermeende druk die zich níét tot frauduleus handelen laten verleiden. Feit is dat mensen in het algemeen, en wetenschappers dus ook, sjoemelen, sjacheren, fouten maken, liegen, en dat sommigen echt over de scheef gaan.

GRIJS GEBIED

Binnen de wetenschap worden diverse integriteitsschendingen onderscheiden. Het verzinnen van gegevens staat met stip bovenaan, gevolgd door plagiaat en het verregaand aanpassen van data. Dan volgt er evenwel een beduidend grijzer gebied.

Praktijken als het weglaten van onwelgevallige gegevens, het aanpassen van statistische analyses omdat die beter uitpakken, het recyclen of opsplitsen van onderzoeksresultaten om maar tot meer publicaties te komen, het achteraf aanpassen en “opleuken” van hypothesen om wel significante resultaten te krijgen – was reeds opgemerkt dat De Groot dergelijke praktijken dik vijftig jaar geleden al aanstipte? –, heten tegenwoordig wel *questionable research practices*.¹¹ Daarvan is lang niet altijd duidelijk hoe moedwillig ze zijn toegepast. Wie er bijvoorbeeld rond voor uitkomt een zogeheten uitbijter (denk aan een proefpersoon die bewust een onderzoek vergalt) uit de dataset te hebben verwijderd, heeft in tegenstelling tot een onderzoeker die dat “stiekem” doet geen enkel probleem.

Hoe vaak zulke twijfelachtige onderzoekspraktijken voorkomen? Volgens een recente studie van de Harvard Business School zou één op de tien psychologen wel eens op onfaire wijze data vervalsen.¹² Volgens een nog recenter proefschrift zou de helft van de wetenschappelijke publicaties in de experimentele psychologie statistische slordigheden bevatten, zoals het ten eigen faveure afronden en oppoetsen van data.¹³ Eerder onderzoek onder leiding van Jelte Wicherts, thans universitair hoofddocent aan het Departement Methoden en Technieken van Tilburg University en initiatiefnemer van het pas opgerichte *Journal of Open Psychology Data*, illustreerde dat psychologen mogelijk ook wel wat te verbergen hebben als ze hun data niet openbaar maken.

Zo deed hij met collega's enkele jaren terug in *American Psychologist* verslag van een opmerkelijke enquête.¹⁴ Van de aangeschreven auteurs die in de laatste twee issues van 2004 in de zogeheten *high impact*-tijdschriften *Journal of Personality and Social Psychology*, *Developmental Psychology*, *Journal of Consulting and Clinical Psychology* en *Journal of Experimental Psychology: Learning, Memory, and Cognition* hadden gepubliceerd, stelde slechts een kwart (!) zijn data beschikbaar voor heranalyses. Saillant, want volgens de ethische richtlijnen van de *American Psychological Association* (APA) verplichten onderzoekers zich ertoe om hun gegevens tenminste vijf jaar beschikbaar te houden voor collega's. Sterker, in een andere studie vonden Wicherts en collega's dat onderzoekers die hun data niet beschikbaar stellen, ook opvallend meer statistische fouten maken.¹⁵

Onduidelijk blijft hoe structureel psychologen rommelen. En zouden die psychologen hun *unfair play* ook toegeven als ze niet anoniem mochten blijven? En hoeveel pakken hun onfaire praktijken of zelfs fraude slimmer en subtieler aan en glippen zo door de mazen van het net? Het is lastig te bepalen hoe vaak fraude in de psychologie exact voorkomt – al zal het grote publiek wellicht denken dat sinds de affaire-Stapel de fraudegevallen zich opstapelen en dat strenge controle absolute noodzaak is. Zo speelt er op het moment dat ik de laatste hand aan dit artikel leg (mei 2014) mogelijk alweer een nieuw geval van datamanipulatie door een sociaal psycholoog aan de Universiteit van Amsterdam.¹⁶ Maar bij een beroepstak die zich bij uitstek kenmerkt door onderling vertrouwen past geen honderdprocentcontrole als waren het potentiële bolletjes-slikkers zoals bij een vlucht vanuit Curaçao naar Schiphol.

TERUG BIJ AF

Neemt niet weg dat er in de psychologische vakpers steeds luider geluiden klinken om transparantie en *fair play* in het onderzoek terug te brengen door onder meer het stimuleren van replicaties, het delen en openbaar maken van data, en het preregistreren van experimenten (waarbij je vooraf vastlegt wat je gaat onderzoeken, hoe je dat gaat doen en welke conclusies je dus wel en niet zult kunnen trekken). Vraag is echter wel hoe nieuw die initiatieven zijn en of veel meer aandacht voor het werk van een oude meester in het vak als De Groot niet gewoon een veel simpelere remedie is.

Zoals aangestipt publiceren tijdschriften heel veel meer significante dan niet-significante resultaten. Replicatiestudies maken vrijwel geen kans op publicatie. Dat is vreemd, een gerepliceerd effect maakt een gevonden effect immers sterker en een niet-gerepliceerd effect zet dat effect meer in empirisch perspectief. Het Open Science Framework initieert thans tal van replicatie-initiatieven die onderdeel moeten gaan vormen van de wetenschappelijke cyclus, met als doel het opleveren van robuustere kennis; één onderzoek is immers geen onderzoek.¹⁷ Dat is beslist lovenswaardig, maar ruim vijftig jaar geleden verbaasde De Groot zich er ook reeds over dat replicaties zo

sporadisch voorkomen: “En als zij worden uitgevoerd, worden de resultaten, geheel ten onrechte, vaak niet gepubliceerd, vooral niet als deze negatief zijn.”¹⁸

Ook over die transparantie en het delen van data, waarvoor onder meer Wicherts en collega's zich lovenswaardig en luidruchtig hard maken, liet De Groot zich lang geleden al uit – herlees het openingscitaat van dit artikel nog maar eens. En eerlijk is eerlijk: ook onderzoekers¹⁹ die pleiten voor preregistratie zullen zich toch ook (weer) tot De Groot moeten wenden: “Een zo volledig mogelijke uitwerking op papier van de toetsings- (c.q. experimentele) opzet vooraf, is in ieder geval sterk aan te bevelen.”²⁰

Fair play in onderzoek heeft ook te maken met diegenen credits geven die dat historisch verdienen – dé manier om daadwerkelijk iets van de geschiedenis te leren. Het is in de wetenschap en zeker in de psychologie daarom tijd voor een pas op de plaats. Het soort ambachtelijke vakmanschap dat De Groot in *Methodologie* ruim vijftig jaar geleden dus al beschreef, moet weer verplichte literatuur worden in de opleiding. Ook moeten er voor promovendi, postdocs, professoren (en sowieso voor alle sociaal psychologen) verplichte *Methodologie*-bijscholingscursussen worden georganiseerd.

Staat dat alles eenmaal stevig op poten, dan zal de te grote variatie in kennis over methodologie en statistiek onder psychologen die er nu is significant afnemen. Dan durf ik te voorspellen dat niet alleen twijfelachtige onderzoekspraktijken tot het verleden zullen behoren, maar dat ook duidelijk zal worden dat psychologen nog altijd ontzettend weinig weten. Veel kennis in de psychologie berust, mede door die *unfair play*, op theoretisch drijfzand. Het verplicht stellen van *Methodologie* zal ertoe bijdragen dat er eerst fatsoenlijk geheid wordt alvorens er verder gebouwd wordt.

En ik durf ook wel te voorspellen dat een bijeffect zal zijn dat de nadruk op kwantiteit in de (psychologische) wetenschap vanzelf afneemt. Want als ik iets weet dankzij De Groots meesterlijke boek: meten is weten, ja, maar simpelweg tellen is kwellen.

P.S.

Nog heel even over impact. Een vriendin van mij is weduwe, haar zoon en dochter verloren op respectievelijk zes- en achtjarige leeftijd hun vader. Haar dochter is nu veertien. Onlangs viel me op dat zij zo veranderd was, alsof haar schichtigheid en onzekerheid plots hadden plaatsgemaakt voor een welhaast volwassen zelfbewustheid. Wat bleek? De dochter van mijn vriendin was naar een rouwverwerkingsweekend geweest voor kinderen die een ouder, broer of zus hadden verloren. Dat weekend stond onder leiding van een psycholoog van wie ze naar eigen zeggen ontzettend veel had geleerd over hoe ze haar emoties over het verlies van haar vader beter een plek kon geven en die met anderen kon delen. Dát is impact. Zou die therapeut dat ook pontificaal in zijn of haar cv vermelden?

Noten

- 1 A.D. DE GROOT, *Methodologie. Grondslagen van onderzoek en denken in de gedragswetenschappen*, Mouton & Co, Den Haag, 1961. Vertaald in 1969 als *Methodology. Foundations of inference and research in the behavioral sciences*. Citaat op p. 372.
- 2 A.D. DE GROOT, *Het denken van den schaker. Een experimenteel-psychologische studie*, Noord-Hollandische Uitgeversmaatschappij, Amsterdam, 1946. Vertaald in 1965 als *Thought and choice in chess*.
- 3 Bijv. VITTORIO BUSATO, “Adriaan de Groot: Meester in de psychologie”, in: VITTORIO BUSATO, MINEKE VAN ESSEN & WILLEM KOOPS (RED.), *Vier grondleggers van de psychologie*, Bert Bakker, Amsterdam, 2014, 280 p.
- 4 Zie noot I, citaat p. 20.
- 5 Ibidem, citaat p. 55. Cursivering De Groot.
- 6 Ibidem, citaat p. 323-324.
- 7 Ibidem, citaat p. 351. Cursivering De Groot.
- 8 DIEDERIK STAPEL, *Ontsporing*, Prometheus, Amsterdam, 2012, 315 p. Zie ook: www.demorgen.be/dm/nl/2461/Opinie/article/detail/1700383/2013/09/08/Brief-van-een-ontspoorde-wetenschapper.dhtml
- 9 Een impactfactor wordt berekend aan de hand van het gemiddelde aantal citaties van alle verschenen artikelen in een tijdschrift binnen een periode van twee jaar. Hoe hoger de impactfactor, des te hoger het wetenschappelijke prestige van een tijdschrift. Betere tijdschriften worden door meer wetenschappers gelezen. Dat vergroot de kans dat artikelen uit dat blad worden geciteerd. Dat vergroot weer de impact van een tijdschrift – enzovoort, enzovoort.
- 10 www.demorgen.be/dm/nl/2461/Opinie/article/detail/1690107/2013/08/21/Zo-kan-het-niet-langer-aan-de-universiteit.dhtml
- 11 L.K. JOHN, G.F. LOEWENSTEIN & D. PRELEC, “Measuring the Prevalence of Questionable Research Practices with Incentives for Truth-Telling”, in: *Psychological Science*, 23, 2012, pp. 524-532. Zie ook: M. Bakker, *Good science, bad science. Questioning research practices in psychological research*, academisch proefschrift, Universiteit van Amsterdam, 2014.
- 12 Zie eerste referentie noot II.
- 13 Ibidem, tweede referentie.
- 14 J.M. WICHERTS, D. BORSBOOM, J. KATS & D. MOLENAAR, “The poor availability of psychological research data for reanalysis”, in: *American Psychologist*, 61, 2006, pp. 726-728.
- 15 J.M. WICHERTS, M. BAKKER & D. MOLENAAR, “Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results”, *PLoS ONE*, 2011. doi: 10.1371/journal.pone.0026828.
- 16 www.nrc.nl/nieuws/2014/04/29/uva-hoogleraar-manipuleerde-data-van-onderzoek/
- 17 Zie www.openscienceframework.org. Zie bijvoorbeeld ook: D. LAKENS, A. HAANS & S. L. KOOLE, “Eén onderzoek is géén onderzoek: Het belang van replicaties voor de psychologische wetenschap”, in: *De Psycholoog*, 9, 2012, pp. 10-18.
- 18 Zie noot I, citaat p. 36.

- 19 Bijv. M. BAKKER, A. DIJK VAN & J.M. WICHERTS, “The rules of the game called psychological science”, in: *Perspectives on Psychological Science*, 7, 2012, pp. 345-551; E.-J. WAGENMAKERS, R. WETZELS, D. BORSBOOM, H.L.J. VAN DER MAAS & R.A. KIEVIT, “An agenda for purely confirmatory research”, in: *Perspectives on Psychological Science*, 7, 2012, 632-638. Opvallend is dat deze groep onderzoekers, allen verbonden (geweest) aan de Programmagroep Psychologische Methodenleer van de UvA, A.D. de Groot (geestelijk vader van die programmagroep) lijkt te hebben (her)ontdekt. Nemen zij in de zojuist genoemde (*high impact*-)publicaties geen verwijzingen op naar diens werk, dat doen zij in een recente Nederlandse publicatie nadrukkelijk wél: M. BAKKER, E.-J. WAGENMAKERS, D. BORSBOOM, J.M. WICHERTS & H.L.J. VAN DER MAAS, “Spelregels in de psychologie”, in: *De Psycholoog*, 12, 2013, pp. 68-76.
- 20 Zie noot 1, citaat p. 142.