

AANBIEDING EINDRAPPORT “FALENDE WETENSCHAP”

door prof. Pim Levelt, voorzitter coördinerende onderzoekscommissie

Trippenhuis, 28 november 2012

De wetenschappelijke fraude van Diederik Stapel, die begin september 2011 in Tilburg aan het licht kwam, dank zij knap werk van drie klokkenluiders, heeft in binnen- en buitenland een schok doen gaan door de academische wereld. Immers, de basis voor alle wetenschappelijke samenwerking is onderling vertrouwen. Wanneer dat vertrouwen ernstig wordt geschonden, zoals in het geval van de heer Stapel, knaagt dat aan de fundamenten van het wetenschappelijk bedrijf.

De Rector Magnificus van Tilburg University besloot direct volstrekte openheid van zaken te geven. Hij stelde een onderzoekscommissie in, de Commissie Levelt, die aard en omvang van de fraude moest vaststellen en de onderzoekscultuur in kaart moest brengen die deze langdurige fraude had mogelijk gemaakt. Dat beeld zou niet volledig worden zonder ook de eerdere werkplekken van de heer Stapel in het onderzoek te betrekken, de Universiteit van Amsterdam, waar hij van 1993 tot 1999 verbleef en in 1997 promoveerde, en de Rijksuniversiteit Groningen, waar hij van 2000 tot 2006 hoogleraar was. In goed overleg met de rectores van die universiteiten werden ook daar onderzoekscommissies ingesteld, de Commissie Drenth in Amsterdam en de Commissie Noort in Groningen. In dit eindrapport doen de drie commissies gezamenlijk verslag van hun bevindingen.

In het interim-rapport van 31 oktober 2011 berichtten de Commissies over aard en omvang van de fraude: datafabricatie en –vervalsing in tenminste een dertigtal artikelen. Maar vooral besteedde dat interim-rapport aandacht aan persoonlijke slachtoffers van deze fraude, met name de promoti van de heer Stapel, die met raffinement en zo nodig intimidatie om de tuin waren geleid. Velen zagen hun publicaties in rook opgaan, met ernstige schade voor hun verdere wetenschappelijke carrière. De commissies concludeerden dat er in geen geval sprake was geweest van medeplichtigheid.

Maar het grote werk moest nog beginnen, althans naar het oordeel van de Commissies. Hier was sprake van een omvangrijk, maar besmet wetenschappelijk oeuvre. In een wetenschappelijke wereld die die naam verdient, moet hier het kaf van het koren te worden gescheiden. Tijdschriftpublicaties trekken soms lange sporen binnen en tussen disciplines. Het zelfreinigend karakter van de wetenschap vereist dat frauduleuze publicaties worden teruggetrokken en niet verder, als een agressief virus, in de literatuur voortwoekeren.

Het siert de drie betrokken universiteiten dat zij de Commissies de vrije hand hebben gegeven in het analyseren van alle publicaties van de heer Stapel. Daarbij werd medewerking van de coauteurs verzocht, die dat zonder uitzondering hebben verleend door onderzoeksmaterialen en data aan te leveren en in veel gevallen door ons in gesprekken te informeren over de gang van zaken bij het onderzoek. Tevens werden aan de drie universiteiten statistici ter beschikking gesteld van de Commissies om alle publicaties van de heer Stapel aan een

grondige analyse te onderwerpen. Het spreekt vanzelf dat het hier om een arbeidsintensieve en kostbare operatie ging. Het is bij ons weten nooit eerder gebeurd dat het hele wetenschappelijk oeuvre van een fraudeur op de korrel is genomen. Dat geldt ook voor recente fraudegevallen in de wetenschap, die nog vers in het geheugen liggen. Ik ben van mening dat daar te gemakkelijk van is afgezien. Het getuigt van wetenschappelijke moed dat Tilburg University, de Rijksuniversiteit Groningen en de Universiteit van Amsterdam de analyse van het gehele oeuvre van een fraudeur tot nieuwe standaard hebben verheven.

De Commissies hebben hun bevindingen over de frauduleuze publicaties van de heer Stapel cumulatief op hun website gepubliceerd, zodat co-auteurs, tijdschriften en de *peer community* zonder uitstel werden geïnformeerd. De slotsom is dat in 55 publicaties van de heer Stapel met zekerheid fraude is vastgesteld; de eerste daarvan dateert uit 2004. Tevens zijn er in 10 artikelen van vóór die datum, in de Amsterdamse en Groningse periode, sterke aanwijzingen voor fraude gevonden. Voorts is vastgesteld dat in 10 dissertaties die onder supervisie van de heer Stapel zijn verschenen één of meerdere hoofdstukken gebaseerd zijn op frauduleuze data, in enkele gevallen zelfs het gehele proefschrift. Het is nu zaak dat de tijdschriften met voortvarendheid alle frauduleuze artikelen terugtrekken, zodat de wetenschappelijke lucht gezuiverd wordt.

Het belangrijkste nieuws echter dat dit eindrapport te bieden heeft betreft de onderzoekscultuur waarbinnen deze grootschalige fraude zolang onopgemerkt bleef. Als onverwachte “bijvangst” leverde de analyse van Stapel’s oeuvre een beeld op van onderzoekspraktijken waarin *sloppy science* aan de orde van de dag was. Ik zal dat verder aanduiden met “slodderwetenschap”. In veel publicaties, ook niet-frauduleuze, werden grondprincipes van ordentelijk wetenschappelijk onderzoek met voeten getreden. We geven daarvan in het eindrapport veel gedetailleerde voorbeelden. Centraal staat de zg. “verificatiebias” die in veel van deze publicaties werd aangetroffen. Gegevens die de onderzoekshypothesen niet bevestigden werden op allerlei manieren weggewerkt. Ik geef u daarvan drie voorbeelden.

- (1). Een experiment levert niet het verwachte resultaat. Dan wordt het, met kleine variaties, een paar keer herhaald en men rapporteert alleen het experiment dat tenslotte wel de verwachte waarden oplevert. Zo wordt een hypothese altijd bevestigd.
- (2). Er worden in een experiment drie condities getoetst. Elk daarvan moet een specifiek resultaat opleveren als de hypothese klopt. Twee condities doen dat, maar de derde doet het niet. In het artikel worden alleen de uitkomsten van de twee positieve condities vermeld, de derde is weggemoffeld.
- (3). Een uitgevoerd experiment bevestigt niet de hypothese. Maar als de resultaten van een of twee proefpersonen worden verwijderd, is die bevestiging er wel. Post hoc wordt dan een reden bedacht om de data van die proefpersonen te verwijderen. Bijvoorbeeld: ze hadden al eens eerder aan zo’n experiment meegedaan.

Zo krijg je uitkomsten die te mooi zijn om waar te wezen.

Ook constateerden de commissies een grote mate van statistische incompetentie in de publicaties waarvan de heer Stapel co-auteur is.

Het oeuvre van de heer Stapel is natuurlijk niet representatief voor zijn vak, de sociale psychologie. Nederland telt vele voortreffelijke sociaal-psychologen, die zich verre houden van zulke slodderwetenschap. Maar het gaat hier wel om publicaties met allerlei co-auteurs uit binnen- en buitenland; wij telden 70 verschillende co-auteurs op de publicaties van de heer Stapel. Het zijn bovendien publicaties waarvan er veel door promotie- en leescommissies zijn beoordeeld. En vooral, het zijn publicaties die de hele reviewprocedure van toptijdschriften op dit vakgebied hebben doorlopen. Het gebeurde zelfs dat reviewers de auteurs adviseerden om onwelgevallige resultaten weg te laten om de lijn van het verhaal niet te verstoren, of om condities waarin geen effecten werden gevonden niet te rapporteren. Hier heeft de kritische functie van de wetenschap van laag tot hoog gefaald. Geen wonder dat het zo lang moest duren voor de fraude werd ontdekt. Dat gebeurde ook niet door verantwoordelijke functionarissen in de academische wereld, maar uiteindelijk door drie van de heer Stapel afhankelijke junior-onderzoekers.

Er is hier op grote schaal sprake geweest van falende wetenschap. Helaas worden een aantal bevindingen van onze Commissies weerspiegeld in de recente open email die Nobelprijswinnaar Daniel Kahneman over de situatie in de sociale psychologie publiceerde: "your field is now the poster child for doubts about the integrity of psychological research. Your problem is not with a few people who have actively challenged the validity of some priming results. It is with the much larger population of colleagues who in the past accepted your surprising results as facts when they were published. These people have now attached a question mark to the field, and it is your responsibility to remove it." Kahneman ziet "a train wreck looming. I expect the first victims to be young people on the job market." Als dit rapport, en de daarin gesignaleerde misstanden, kan bijdragen aan die saneringsoperatie, dan heeft de Stapel-fraude naast alle aangerichte, vooral persoonlijke schade, wellicht nog een positief effect.

Het ziet er gelukkig naar uit dat die uitdaging snel wordt opgepakt door de wetenschappelijke gemeenschap. Ons eindrapport vermeldt, als inleiding van hoofdstuk 6 "Aanbevelingen", een aantal zelf-corrigerende initiatieven die de laatste maanden zijn genomen. Ik noem er één, het meest recente. Het internationale tijdschrift *Perspectives on Psychological Science* heeft nog deze maand november een omvangrijke *special issue* uitgegeven over het probleem van onderzoeksreplicatie in de psychologie, met 23 uitstekende artikelen, waaraan opvallend veel Nederlandse experts hebben bijgedragen.

Tot slot moge ik de finale conclusie van dit rapport citeren: "de bespreking van de factoren die de fraude van de heer Stapel hebben gefaciliteerd, zijn machtige positie van coryfee en de falende kritische functie van zijn wetenschappelijke wereld laten onverlet de conclusie van de Commissies, dat de fraude begaan door de heer Stapel dusdanig is dat hij in die factoren en omstandigheden geen enkel excuus voor zijn handelen kan vinden."

