

Alle boeken de wetenschap uit; om te beginnen die in het Nederlands

De Academische Boekengids 23, oktober 2000, pp. 10-11.

Het Centrum voor Wetenschaps- en Technologische Studies verwerkt zijn beslissingen onmerkbaar in de geleverde scores. Dat is gebleken bij de VSNU-visitatie van de Nederlandse psychologie in 1999.

Op congressen zoals die van de American Psychological Association en de History of Science Society komen niet alleen wetenschappelijk onderzoekers. Ook alle grote uitgeverijen zijn er vertegenwoordigd. De University Presses van Cambridge, Harvard, en Chicago zijn aanwezig, Routledge is er, Sage, Gallimard, Elsevier, en noem maar op. Ze presenteren zich met prachtig ingerichte kramen, reclamemateriaal en vooral met heel veel boeken. Met trots worden de nieuwste uitgaven getoond. De manuscripten zijn grondig becommentarieerd door collega's, en eindredacteurs hebben de tekst nog eens met uiterste precisie geschoond. Eventuele illustraties zijn zo helder mogelijk weergegeven. In de pauzes tussen de sessies zwermen lezers en auteurs uit naar de kramen, waar ze met de vertegenwoordigers van de uitgeverijen praten over fondsen en individuele boeken.

Tijdens sommige congressen, bijvoorbeeld dat van de International Society for the History of the Social and Behavioral Sciences, worden zelfs volledige sessies gewijd aan nieuw verschenen boeken. Een aantal toonaangevende onderzoekers geeft een gesproken recensie van een enkel boek. De auteur reageert op de opmerkingen en ook de zaal bemoeit zich ermee.

De verborgen missie van het CWTS

Toen ik deze zomer zo'n congres meemaakte, verbeeldde ik me de reactie van medewerkers van het Leidse CWTS (Centrum voor Wetenschaps- en Technologie Studies). Ze zouden hun hoofden schudden over zoveel verspilling van tijd en geld. Naar eigen zeggen berekent het CWTS wetenschappelijke productie en de 'impact' daarvan vanuit het perspectief van verantwoording van onderzoeksgelden (Nederhof, Van Leeuwen en Visser, 2000, p. 1). Voor dat doel hanteert zij echter uiterst nauwe definities. Boeken, boekhoofdstukken, boekbesprekingen, en verwijzingen daarnaar, tellen niet mee bij het bepalen van wetenschappelijke productie en de impact daarvan. Van de Nederlandstalige publicaties doen ook de artikelen dat nauwelijks.

Niet dat het CWTS dit erbij vermeldt. Het verwerkt dergelijke beslissingen onzichtbaar in de geleverde scores. Het is als met het zogeheten 'swill' - gemalen etensresten uit ziekenhuizen en restaurants - in diervoedsel. Op het etiket van je duur betaalde blikjes gourmet-kattenvoer zul je dat ingrediënt niet vermeld zien (Trommelen, 1999). Wie de samenstelling echt wil achterhalen, moet de interne bedrijfsrapportage zien te verkrijgen en dan nog eigen analyses doen.

Zo ging het althans bij de VSNU-visitatie van de Nederlandse psychologie die vorig jaar werd afgerond. De visitatiecommissie liet het berekenen van de productie en impact van de 61 landelijke psychologieprogramma's over aan de Leidse sciëntometrische experts. Zonder dat dit aan de onderzochte wetenschapsbeoefenaren bekend werd gemaakt, gebruikte het CWTS bij de samenstelling van zijn prestatiescores ingrediënten die voor sommige programma's ronduit schadelijk zijn.

In de meeste psychologische onderzoeksgebieden wordt vooral via tijdschriftartikelen gecommuniceerd, in sommige echter veel via boeken. Zoals ik zal laten zien, kon aan het laatste type programma met de berekeningen van het CWTS geen recht worden gedaan. Waar dit soort 'verantwoording van onderzoeksgeld' wordt vereist, zullen goede boeken - om te beginnen die in het Nederlands - dan ook uit de wetenschap verdwijnen.

Een nul voor impact

Wie neemt er nu de tijd om CWTS-cijfers te analyseren? Mijn speurtocht naar de geheimen van de sciëntometrie begon met het concept-visitatierapport Psychologie van 1999. Daarin werd mijn eigen programma, Theorie en Geschiedenis van de Psychologie, aan de Rijksuniversiteit Groningen ('RUG 7' in termen van het rapport) gesierd met een asterisk. Deze asterisk - ook in het definitieve rapport staat hij er - deelt RUG7 met acht andere programma's in de Nederlandse psychologie. Het visitatierapport licht toe dat het sterretje bepaald geen positieve kwalificatie symboliseert. Het gaat om onderzoeksprogramma's die minder dan twintig 'papers' hadden en die minder dan twintig keer zijn geciteerd door anderen in de periode 1993-1998 (VSNU, 1999, p.111-112).

Nu is Theorie en Geschiedenis een heel klein programma (2,7 fte vaste staf in de meetperiode), maar we wisten zeker dat we aardig wat hadden gepubliceerd. En wat de verwijzingen naar dat werk betreft: een getal als 1 of 2 citaties hadden we misschien nog geloofd, maar een 0 was domweg te onwaarschijnlijk. Het CWTS maakt gebruik van de diensten van het Amerikaanse ISI (Institute for Scientific Information). Ook naar eigen zeggen kan het ISI slechts in beperkte mate de impact van wetenschappelijk werk meten. Het is bekend dat het niet alle tijdschriften doorzoekt en dat verwijzingen in boeken helemaal niet

worden getraceerd. Maar zelfs in de beperkte set van voornamelijk Amerikaanse tijdschriften zou er toch tenminste één citatie naar RUG7 moeten zijn. Een collega deed een hertelling en vond 26 ISI-verwijzingen in de periode 1993-1998, zelf-citaties en binnengroep-citaties niet meegeteld.

Andere indicatoren

Het conceptrapport was naar de onderzoeksinstituten gestuurd met een uitnodiging om fouten te signaleren. De zelf gevonden citaties werden daarom gekopieerd en door het bestuur van het Heymans Instituut (het Instituut voor Fundamenteel Psychologisch Onderzoek aan de RUG) naar de VSNU gezonden. Omdat citaties in boeken niet konden worden geteld en Theorie en Geschiedenis volgens het conceptrapport *onzichtbaar* en zelfs *geïsoleerd* zou zijn, werd tevens gewezen op andere indicatoren van internationale productie en zichtbaarheid van het programma. Het Heymans Instituut noemde bijvoorbeeld een 'gereviewed' boek (in zijn geheel geschreven door één auteur) dat was gepubliceerd bij de Cambridge University Press, en dat was besproken in gerenommeerde tijdschriften zoals *The Bulletin of the History of Medicine* en het boekbesprekingstijdschrift *Contemporary Psychology van de American Psychological Association* (APA). Vermeld werd dat ook een ander boek, eveneens van één enkele auteur en eveneens na grondige 'peerreview', was geaccepteerd door de Cambridge University Press, alsmede door de Italiaanse uitgeverij Einaudi. Er werd gewag gemaakt van het bijzondere van een artikel in *Isis*, het tijdschrift van de Amerikaanse History of Science Society, dat doorgaans alleen artikelen over natuurwetenschappen publiceert.

Nog meer voorbeelden werden genoemd. Er was een eigen besprekingsartikel in *Contemporary Psychology* waarin men alleen op uitnodiging kan schrijven. Dat stuk was bovendien door de redactie 'gefeatured' (dat wil zeggen: als zeer goed artikel voorin geplaatst en voorzien van portretten van recensent en auteur). Er was ook een boekbespreking in *Isis* die eveneens alleen op uitnodiging kon worden geschreven. Voorts was er een redactielidmaatschap van het APA-tijdschrift *History of Psychology*, een (aangenomen) uitnodiging van de APA om voor haar congres van 1996 een sessie te organiseren over Contextual Historiography of Psychology, een (aangenomen) uitnodiging door de British Psychological Society, sectie History, om op te treden als gastspreker, een flinke onderzoekssubsidie van de Britse Wellcome Trust die men zonder goede buitenlandse reputatie niet kan verkrijgen, een jarenlang bestuurslidmaatschap van de European Society for the History of the Human Sciences, en verschillende optredens als gastdocent bij de Onderzoekersopleiding Wetenschapsstudies te Noorwegen. Gevoegd bij de 26 citaties in isi-artikelen leek dit voldoende argumentatie voor de conclusie dat Theorie en Geschiedenis internationaal zeker niet inactief en onopgemerkt was gebleven.

Het Heymans Instituut sprak ook het vermoeden uit dat boeken bij het vaststellen van de productie niet goed waren geteld, ervan uitgaande dat een volledig boek door één auteur gelijk was gesteld aan een enkele publicatie door die auteur. Het instituut wees daarnaast nog op het opmerkelijke feit dat de landelijk gezien extreem hoge score voor *Nederlandstalige* wetenschappelijke publicaties die RUG7 desondanks had (we hebben Douwe Draaisma in ons team) dit programma eenzelfde eindcijfer voor productiviteit had opgeleverd als andere programma's die nationaal veel lager scoorden en internationaal ongeveer gelijk.

Over de brede ontvangst van ons Nederlandstalige werk is verder maar gezwegen. Impact schijnt immers alleen iets internationaals (lees Angelsaksisch) te zijn. Er kwam geen reactie van de VSNU of de commissie. In september verscheen het definitieve rapport. De enige verandering was een verhoging van het kwaliteitscijfer van 3 naar 3.5 (op een schaal van 1-5). Het productiviteitscijfer bleef een 3.5, de nul aan citaties naar RUG7 was blijven staan, alsmede de asterisk en de opmerkingen over de zorgwekkende onzichtbaarheid van programma (VSNU, 1999, p.20 en p.45).

Geldige en ongeldige citaties

Een 3.5 voor kwaliteit en productiviteit is nu ook zo laag weer niet. En voor de beide andere categorieën 'relevance' en 'viability' kreeg het programma een 4. Maar waren boeken echt slechts als een enkele publicatie geteld? En vooral: waarom stond die nul aan impact nog steeds in het rapport? Wat was er toch met die zelf gevonden isi-citaties naar RUG7 gebeurd? De laatste vraag heb ik toen maar per e-mail recht-streeks aan het CWTS gesteld. De antwoorden in de daaropvolgende correspondentie waren opmerkelijk. Er viel uit op te maken dat het CWTS zelf nog eens beperkingen aanbrengt in de toch al gelimiteerde reeks van het ISI.

Niet alleen worden verwijzingen in boeken door het isi ongeteld gelaten, niet alleen is het aantal ISI tijdschriften beperkt, het CWTS accepteert zelfs niet alle ISI-citaties! Met name boekgerelateerde verwijzingen worden genegeerd. Ten eerste telt het CWTS verwijzingen naar boeken niet mee, ook al stonden deze in tijdschriften opgenomen door het ISI. Ten tweede telt het CWTS boekrecensies niet mee, zelfs niet als het om artikelrecensies gaat, en zelfs niet als ze in ISI-tijdschriften staan. Ten derde telt het CWTS citaties uit de onderzochte periode (in dit geval 1993-1998) niet mee als deze verwijzen naar werk gepubliceerd voor deze periode. Dit laatste werkt nadelig uit voor boekgeoriënteerde programma's vanwege het erkende sciëntometrische gegeven dat sommig, bijvoorbeeld experimenteel, onderzoek relatief snel wordt geciteerd, en ander, bijvoorbeeld theoretisch/historisch onderzoek, relatief laat.

Uit de publicatielijsten van psychologisch Nederland is dus een deelverzameling gehaald van die publicaties die *volgens het CWTS* citeerwaardig zijn. Het visitatierapport vermeldt dit niet en geeft er dus ook geen verdediging van. Er staat alleen dat de impact van de artikelen van een programma is

afgemeten aan de citaties ernaar (VSNU, 1999, p.112).

Behalve de manier waarop de citatiescore tot stand was gekomen, was het ook nog onduidelijk hoe de productie was berekend. Hoe kwamen we aan die asterisk die niet alleen staat voor minder dan twintig citaties, maar ook voor minder dan twintig papers? Het begon me te dagen dat de CWTS-deelverzameling van citeerwaardige publicaties eveneens was gebruikt voor het vaststellen van het aantal gepubliceerde papers. De kleine lettertjes in het visitatierapport leerden bovendien dat eenzelfde besluit was genomen bij de *algemene* berekening van wetenschappelijke productie: boekhoofdstukken en hele boeken waren daarbij niet slechts als één artikel geteld, zoals het Heymans Instituut in zijn reactie nog veronderstelde. Ze waren *helemaal niet* meegerekend. Het ging bij productie alleen maar om aantallen publicaties in internationale respectievelijk nationale 'gereviewde' tijdschriften (VSNU, 1999, p.112).

Na het verschijnen van het definitieve rapport stuurde het Heymans Instituut opnieuw een reactie naar de VSNU. Ook hierop volgde geen nadere uitleg over de wijzen van tellen. Wel relativeerde (toenmalig) VSNU-voorzitter Meijerink in een schriftelijk antwoord het belang van de CWTS-scores. Hij stelde dat ze slechts een hulpmiddel waren geweest bij de beoordelingen, en hetzelfde deed mondeling de voorzitter van de commissie (respectievelijk R. Meijerink, brief van 11-11-1999 aan P.L.C van Geert, en A.F. Kalverboer, mondelinge mededeling aan T. Dehue).

Als de impact van de CWTS-berekeningen inderdaad verwaarloosbaar klein is, vraag je je af hoe het onderzoeksgeld kan worden verantwoord - bijna fl.67000,- in dit geval - dat het CWTS in rekening brengt. De cijfers werden echter al in een vroeg stadium aan de commissie bekendgemaakt, en het heeft er alle schijn van dat ze hebben meegewogen.

In alle inleidingen in de psychologie worden de experimenten uit de jaren zeventig besproken van de Amerikaan Rosenthal. Deze liet proefpersonen de prestaties van ratten beoordelen. Sommigen van hen maakte hij wijs dat hun ratten bij eerdere tests heel goed hadden gepresteerd, en sommigen juist dat hun ratten het bij andere tests slecht hadden gedaan. De proefpersonen moesten vervolgens nauwkeurig registreren hoe snel hun ratten de weg door een doolhof vonden. Het bleek dat hun waarneming sterk afhing van hetgeen hun tevoren over de dieren was verteld. Ratten waarover was gezegd dat ze op andere tests lage scores haalden, kregen lagere doolhofcijfers dan ratten waarover de beoordelaars hadden gehoord dat ze eerder hoge scores haalden.

Iedere psycholoog leerde aan de hand van dit voorbeeld over het zogeheten 'experimenter expectancy effect'. Toch lijkt dat effect bij de commissie een rol te hebben gespeeld. Tijdens het 'gesprek op locatie' vroeg de visitatiecommissie wat ik dacht te gaan doen aan de zichtbaarheid van ons programma. Als enige onwetend van de CWTS-nul begreep ik de portee van de vraag niet, en beantwoordde hem merkbaar niet naar tevredenheid.

Voorwaar een doolhoftest. Maar waarschijnlijk was het hoe dan ook niet meer mogelijk geweest het eenmaal opgeroepen beeld te corrigeren. De later door het Heymans Instituut toegestuurde reeks ISI-verwijzingen naar Theorie en Geschiedenis en vele andere blijken van internationale erkenning konden immers het gecreëerde imago ook niet meer bijstellen. De eindcijfers voor productiviteit en kwaliteit van diverse programma's wijzen daar eveneens op. De commissie vermeldt dat deze cijfers tot stand zijn gekomen op basis van haar algemene impressies, waarbij zij meer factoren in beschouwing heeft willen nemen dan alleen de gegevens van het CWTS (VSNU, 2000, p.5-6). Toch ziet het ernaar uit dat het 'experimenter expectancy effect' zich hierbij heeft doen gelden. Uit de tabellen achterin het rapport blijkt bijvoorbeeld dat het buitengewoon hoge aantal Nederlandstalige wetenschappelijke publicaties van RUG7 helemaal niets heeft bijgedragen aan het uiteindelijke productiviteitscijfer van dit programma: het productiviteitscijfer is gelijk aan dat van andere programma's met ongeveer eenzelfde hoeveelheid internationale wetenschappelijke artikelen. Voor kwaliteit kreeg Theorie en Geschiedenis zelfs een lager cijfer dan andere programma's met een ongeveer gelijke score voor internationale wetenschappelijke artikelen (VSNU, 1999, p.109-112). Dat laatste wordt in het verbale rapport rechtstreeks toegeschreven aan gebrek aan 'impact on the international scene' (VSNU, 2000, p.45).

Wat heet tellen?

De macht van het ISI is nergens zo groot als in Nederland. Elders kunnen nog gedegen geschiedenissen van de statistiek worden geschreven. Dergelijke boeken laten zien dat getallen de sociale werkelijkheid niet zomaar afbeelden. Ze tonen hoe de statistische methoden de wereld *scheppen* die men meent er louter mee te beschrijven (o.a. Porter, 1995; Desrosières, 1998; Marks, 1997). Voor Nederland kan dat idee heel goed worden geïllustreerd aan de productie en impact-metingen van het CWTS. De sciëntometrische methoden creëren onmiskenbaar het soort wetenschap dat ermee kan worden gemeten. 'Het natuurwetenschappelijk publicatiemodel wordt dominant' stelt CWTS'er Ton Nederhof onder zijn opgewekt glimlachend portret (in Wouters, 1998, p.9), daarmee bedoelend dat ook de sociale en de geesteswetenschappen voornamelijk in Engelstalige artikelen zullen gaan communiceren. Dat haal je de koekoek. De voorspelling doet denken aan 'De nieuwe kleur wordt blauw' in een *Elle Mode* vol met blauwe kleding.

Er wordt een hoge prijs betaald voor deze verschraling van wat wetenschap mag heten. Als de VSNU haar plan doorvoert om voortaan alle disciplines door het CWTS te laten reguleren (Wouters, 1998),

zullen gedegen Nederlandstalige artikelen en vooral boeken spoedig niet meer geschreven kunnen worden. En hetzelfde geldt voor boeken in het algemeen. Engelstalige boeken en verwijzingen daarnaar worden immers eveneens ongeteld gelaten. Het gaan tellen van citaties in boeken, zoals het ISI van plan schijnt te zijn, zou trouwens niet veel helpen. Het probleem zit in het tellen zelf.

Als kwantiteit de basis van beoordeling is, zal er vooral kwantiteit moeten worden geleverd. En dat gaat het allerbest met onderzoek dat zich leent voor publicatie in gestandaardiseerde artikelvorm. In de psychologie en de sociale wetenschappen bestaat dat onderzoek uit het verzamelen van getallen en het daarop toepassen van statistische toetsen. Dat mag op zich moeilijk zijn, maar de verslaglegging ervan is tot op het niveau van het vocabulaire en de zinsbouw vastgelegd in *publication manuals*. Aan de combinatie van precisie en variatie in uitdrukkingwijzen hoeft dus niet te worden gewerkt. Ook ligt de opbouw van deze artikelen vast: eerst komt de hypothese, dan de werkwijze, de resultaten, en ten slotte de conclusie. Bij het maken van dergelijke teksten is weinig creativiteit en verbale competentie vereist. Derhalve kun je ze even gemakkelijk schrijven in een taal die de jouwe niet is. Je lezers zijn wat dit aangaat niet veeleisend. Ze zijn wel wat basaal Engels gewend, of hebben zelfs nooit geleerd het verschil te zien. Omdat deze stukken in veel mindere mate een persoonlijk stempel dragen van de auteur dan theoretisch of interpretatief werk, kunnen ze bovendien worden geschreven door hele groepen tegelijk. Dat verhoogt ieders productie en zichtbaarheid nog eens aanmerkelijk. Eigenlijk levert vooral die sociale wetenschap kans op hoge scores die lijkt op het sciëntometrisch onderzoek zélf. Zo scheidt het CWTS de wereld naar zijn evenbeeld.

Trudy Dehue is hoogleraar Theorie en Geschiedenis van de Psychologie aan de Rijksuniversiteit Groningen.

Besproken boeken:

Desrosières, A., *La politique des grands nombres. Histoire de la raison statistique*. Parijs: La Découverte, 1993.

Marks, H.M., *The Progress of Experiment. Science and Therapeutic Reform in the United States, 1900-1990*. New York: Cambridge University Press, 1997.

Nederhof, A.J., T.N. van Leeuwen en M.S. Visser, *Bibliometric Profiles of Academic Psychology Research in the Netherlands*. Leiden: CWTS Report 2000-05, 2000.

Porter, T.M., *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press, 1995.

Trommelen, J., 'Keukenafval in varkensvoer ondanks verbod'. *De Volkskrant*, 22-6-1999.

VSNU, *Psychology*. Utrecht: VSNU, 1999.

Wouters, P., 'Boeken tellen niet'. *Hypothese*, oktober 1998.