

Essay

Twijfelachtige onderzoekspraktijken zijn een zorg voor de hele samenleving

Belangrijke beslissingen hangen soms sterk af van wetenschappelijke kennis. Wanneer die kennis ondeugdelijk is, kunnen de maatschappelijke gevolgen daarvan groot zijn, zeker in crisissituaties. De impact van twijfelachtige onderzoekspraktijken reikt tot ver buiten het wetenschappelijk bedrijf. In dit essay pleiten we er daarom voor om het gesprek over dit probleem, dat tot nu toe voornamelijk binnen de wetenschappelijke gemeenschap wordt gevoerd, breder te trekken naar beleidsmakers en maatschappij. Ook kijken we naar mogelijke oplossingen.

Slordigheid met grote gevolgen

Dat slordige wetenschap grote maatschappelijke gevolgen kan hebben, werd pijnlijk duidelijk aan het begin van de coronapandemie. Het Chinese *Journal of Epidemiology* publiceerde in maart 2020 een onderzoek dat stelde dat PCR-testen voor 50% onnauwkeurig waren bij het vaststellen van een coronabesmetting. De versneld gepubliceerde studie stond slechts enkele dagen online, maar dat was voldoende voor Amerikaanse deskundigen om de inkoop van testen af te schalen. Het tijdschrift besloot het artikel al snel weer in te trekken vanwege problemen met de methodologie, maar toen was de beslissing al genomen met alle gevolgen voor de testcapaciteit van dien¹.

Slordig onderzoek is een voorbeeld van een breder probleem, dat bekend staat als twijfelachtige onderzoekspraktijken. Twijfelachtige onderzoekspraktijken resulteren als onderzoekers methodologische keuzes maken die niet correct zijn. Meestal betreft dat geen opzet van de onderzoeker, maar komt het door gebrek aan methodologische kennis en/of een onderschatting van het belang van methodologische strengheid. We moeten hierbij niet denken aan frauduleuze zaken zoals het fabriceren van data of het plagiaat, maar aan acties die de onderzoeksuitkomsten net iets gunstiger doen lijken en daarmee als gebrekkig gezien kunnen worden.

De druk waaronder onderzoek gerelateerd aan COVID-19 is geproduceerd tijdens de coronapandemie heeft het probleem van twijfelachtige onderzoekspraktijken weliswaar verergerd, maar niet veroorzaakt. Twijfelachtige onderzoekspraktijken komen al decennia in allerlei wetenschappelijke disciplines voor, met name in de sociale en de medische wetenschappen, bij kwantitatief onderzoek dat hypothesen toetst. Wetenschappers hebben hier veel onderzoek over gedaan. Zo publiceerde de Tilburgse

1 Marcus, A., & Oransky, I. (2020). The Science of This Pandemic Is Moving at Dangerous Speeds. *Wired*

methodologiehoogleraar Jelte Wicherts samen met anderen² een checklist met maar liefst 34 twijfelachtige keuzes die een onderzoeker beter kan vermijden bij psychologisch kwantitatief onderzoek. Binnen andere disciplines, zoals de geesteswetenschappen, het kwalitatieve sociaalwetenschappelijk onderzoek en het meer toegepaste onderzoek, is er veel minder bekend over de aard en omvang van het probleem.

Twijfelachtige onderzoekspraktijken kunnen leiden tot onbetrouwbare onderzoeksuitkomsten. Daarom vormen ze in de disciplines waarin ze onderzocht zijn een groot probleem, en mogelijk in andere disciplines ook. Veel kwantitatieve studies liggen aan de basis van beleidskeuzes die daardoor soms slecht gefundeerd zijn. Daarnaast kunnen twijfelachtige onderzoekspraktijken leiden tot dalend vertrouwen in de wetenschap en een inefficiënt wetenschapssysteem.

In het afgelopen decennium kwam aan het licht dat twijfelachtige praktijken vaak voorkomen. De resultaten van de *National Survey Research Integrity* lieten in juli 2021 zien dat ongeveer de helft van de Nederlandse wetenschappers, van allerlei disciplines, wel eens twijfelachtige onderzoekspraktijken gebruikt.³ Uit een groot internationaal replicatieproject in de psychologie dat in 2015 verscheen, bleek dat het aantal statistisch significante resultaten in gerepliceerde studies ('herhaalstudies') maar 39% is, waar dat in de originele set studies 97% was⁴. In een artikel uit *The Lancet* in 2009, wordt geschat dat maar liefst 85% van het gestarte medisch onderzoek *research waste* betreft, onderzoek dat door onvoldoende kwaliteit geen bijdrage levert aan wetenschap en/of maatschappij⁵.

Hoe zien twijfelachtige onderzoekspraktijken eruit?

Twijfelachtige onderzoekspraktijken kunnen plaatsvinden in meerdere fases van het onderzoeksproces: de opzet, de dataverzameling, de analyse en de rapportage. Zo kunnen onderzoekers bij het opzetten van een experimentele studie naar de werkzaamheid van medicatie een experiment plannen zonder een goed vergelijkbare controlegroep. In een observatie-studie naar autisme kunnen ze een casus selecteren, alleen maar omdat die in het persoonlijke netwerk van de onderzoeker zit. Bij een survey naar politieke voorkeur kunnen ze een suggestieve vraag stellen die de resultaten vertekent. Bij een experimentele studie naar een vaccin kunnen ze een steekproef gebruiken die te klein is om de resultaten te kunnen generaliseren.

-
- 2 Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., Van Aert, R., & Van Assen, M. A. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in psychology*, 7, 1832.
 - 3 Gopalakrishna, G., Wicherts, J. M., Vink, G., Stoop, I., Van den Akker, O., Riet, G. t., & Bouter, L. (2021, July 6). Prevalence of responsible research practices and their potential explanatory factors: a survey among academic researchers in The Netherlands (preprint). <https://doi.org/10.31222/osf.io/xsn94>
 - 4 Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251).
 - 5 Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *The Lancet*, 374(9683), 86-89.

In kwantitatief toetsend onderzoek, de discipline waarin veel geschreven is over twijfelachtige praktijken, speelt statistiek een grote rol. Een instrument dat onderzoekers van oudsher veel gebruiken om effecten te toetsen, is de nulhypothese-significantietoets. De nulhypothese, die de verwachting inhoudt dat de data geen effect zullen laten zien, wordt daarbij getoetst tegen de alternatieve hypothese, die een specifiek te verwachten effect formuleert. De zogenaamde p -waarde die uit de toets komt, bepaalt of de nulhypothese verworpen wordt en speelt daarmee een cruciale rol bij het formuleren van onderzoeksresultaten. De statistische conventie luidt dat die waarde onder de 0,05 moet liggen om de nulhypothese te verwerpen. Bij het toepassen van de toets moeten onderzoekers echter een aantal keuzes maken die de uitkomst ervan beïnvloeden. Daarbij weten die onderzoekers ook dat de kans dat een tijdschrift hun uitkomsten wil publiceren, groter wordt als ze een lage p -waarde rapporteren, waarmee ze het effect dat ze in hun hypothese geformuleerd hebben daadwerkelijk vinden. Het komt voor dat onderzoekers meerdere toetsen op hun data uitvoeren totdat ze een publicatie-waardig resultaat vinden. De data van een studie waarvan de resultaten geen sterk verband of effect laten zien, worden dan ‘uitgewrongen’ om toch nog ergens een interessante publicatiewaardige statistiek te vinden. Dit staat ook bekend als *data dredging* of *p-hacking*. Wanneer de verzamelde data de vooraf geformuleerde verwachting niet bevestigen, kiest de onderzoeker er soms voor om dit niet eenvoudigweg te rapporteren, maar op zoek te gaan naar een ander effect dat de data wel laten zien en daarover achteraf een verwachting te formuleren. Dit formuleren van een verwachting nadat de resultaten bekend zijn, staat bekend als *HARKing* (*hypothesizing after the results are known*). Het kan leiden tot inaccurate onderzoeksconclusies.

Oorzaken in het wetenschapssysteem

Welke oorzaken liggen er aan de basis van twijfelachtige onderzoekspraktijken? We bespreken er vier.

1. Publicatiedruk

Hiervoor noemden we al dat wetenschappers in kwantitatief toetsend onderzoek graag een effect in hun studie vinden omdat dat publicatie in een tijdschrift waarschijnlijker maakt. Hoewel er nog niet veel onderzoek is gedaan naar oorzaken voor twijfelachtige onderzoekspraktijken op het niveau van het wetenschapssysteem, weten we wel welke prikkels excellentiebeleid in de wetenschap veroorzaakt. Publicatiedruk in het wetenschapssysteem leidt tot hoge werkdruk voor onderzoekers, waardoor ze soms niet de ruimte voelen om de best beschikbare onderzoeksmethoden toe te passen.^{6,7}

6 Tijdink, J., C.A. Maclaine Pont & J.F.M. de Jonge, Publicatiedruk bij medisch-wetenschappelijk onderzoek. Den Haag, Rathenau Instituut 2015

7 Scholten, W., L. van Drooge en P. Diederik (2018). Excellent is niet gewoon – Dertig jaar focus op excellentie in het Nederlandse wetenschapsbeleid. Den Haag: Rathenau Instituut

2. Beperkte methodologische kennis

Daarnaast zorgt hoge werkdruk ervoor dat er minder tijd overblijft voor onderzoekers om zichzelf te blijven scholen, bijvoorbeeld op het gebied van methodologie⁸. Voor onderzoekers is dat vaak ondergeschikt aan scholing op de inhoud. Kennis over de kwaliteit van keuzes in de methodologie blijft dan impliciet of onvoldoende aanwezig, waardoor onderzoekers dezelfde twijfelachtige praktijken blijven gebruiken.

3. Conventies

Net als veel andere instituties is het wetenschapssysteem in de kern conservatief: wetenschappers groeien op binnen het systeem en onderwijzen nieuwe wetenschappers ook weer volgens dezelfde principes. Competitieve praktijken zoals wetenschappelijke tijdschriften, academische loopbaanpaden, onderzoeksevaluaties en financieringsstromen hebben de neiging om bestaande opvattingen en normen te versterken. Veel gewoonten en gebruiken zitten in het systeem ingebakken, zo ook de methodologische taal die wetenschappers onderling spreken. Veel onderzoekers gebruiken in hun statistische analyses nog steeds de nulhypothese-significantietoets, die sinds enkele jaren als twijfelachtig wordt bestempeld en die vaak in verband wordt gebracht met twijfelachtige onderzoekspraktijken⁹. Dat het wetenschapssysteem in de kern conservatief is, is niet per se een slechte eigenschap. Wetenschappers benaderen nieuwe kennis kritisch en gaan daar niet onmiddellijk in mee, ook waar het gaat om inzichten die de methodologie betreffen. Recent onderzoek⁸ laat echter ook zien dat veel onderzoekers de nulhypothese-significantietoets blijven gebruiken omdat ze nou eenmaal gewend zijn om in de statistische analyse bepaalde keuzes te maken. Ze zijn opgevoed in die methodologische leer en hun werkomgeving is daarop ingericht: hun onderwijs, interne kwaliteitsprocedures, verwachtingen van collega's en tijdschriftredacties, subsidiecriteria en de softwarepakketten die standaard gebruikt worden voor statistische data-analyse. Wanneer we methodologisch gedrag van onderzoekers zouden willen veranderen, moeten we dus op systeemniveau substantieel prikkels veranderen.¹⁰

4. Gebrek aan leiderschap

Het eerder genoemde onderzoek⁸ toont aan dat er, specifiek in het onderzoek waarin nulhypothese-toetsen gebruikt worden, weinig verandert doordat wetenschappers een sterke afhankelijkheid van anderen voelen. Enerzijds in de verhoudingen met supervisors, zoals bij promotietrajecten waarin promovendi niet altijd de vrijheid krijgen om eigen methodologische keuzes te maken, anderzijds in de soms expliciete verwachtingen van lezers, tijdschriften en subsidieverstrekkers. Onderzoekers geven aan dat hun methodologische keuzes soms ook bepaald worden door wat hun

8 Stunt, J., van Grootel, L., Bouter, L., Trafimow, D., Hoekstra, T., & de Boer, M. (2021). Why we habitually engage in null-hypothesis significance testing: A qualitative study. *PloS one*, 16(10), e0258330.

9 Van Dongen, N. & van Grootel, L. (preprint). Overview on the Null Hypothesis Significance Test: A Systematic Review on Essay Literature on its Problems and Solutions. *PsyArXiv Preprints | Overview on the Null Hypothesis Significance Test*

10 Scholten, W., L. van Drooge en P. Diederik (2018). *Excellent is niet gewoon – Dertig jaar focus op excellentie in het Nederlandse wetenschapsbeleid*. Den Haag: Rathenau Instituut

meerdere hen opdraagt of impliciet van hen verwacht. Er heerst een cultuur van reactiviteit in de wetenschappelijke wereld waarin verschillende stakeholders naar elkaar wijzen om het initiatief te nemen om onderzoekspraktijken te veranderen. Op individueel niveau zijn er bovendien grote verschillen in hoeverre onderzoekers zich verantwoordelijk voelen voor het up-to-date blijven met ontwikkelingen in de onderzoeksmethodologie en daarin mee te gaan of zelfs te pionieren.¹¹ Leiderschap en initiatief om een structurele verandering teweeg te brengen, ontbreken.

Maatschappelijke implicaties

Twijfelachtige praktijken zijn niet alleen een probleem voor het wetenschapsbedrijf, ook voor de samenleving kunnen ze grote gevolgen hebben.

Ten eerste is wetenschappelijk onderzoek een belangrijke informatiebron voor beleidsontwikkeling. Wanneer onderwijskundig onderzoek vaststelt dat een nieuwe leesmethode beter werkt voor kinderen met een taalachterstand, is het waarschijnlijk dat beleidsmakers in het onderwijsbeleid dit zullen opvoeren als reden om deze methode te implementeren. Wanneer dat wetenschappelijk bewijs gebaseerd is op twijfelachtig uitgevoerd onderzoek, kan het in de publicatie gesuggereerde resultaat in de praktijk afwezig zijn of afwijken.

Niet alleen kunnen gepubliceerde resultaten van een studie een directe invloed hebben. Ook draagt publicatie van de studie als onderdeel van de literatuur over een specifiek onderwerp bij aan een vertekend algemeen beeld van dat onderwerp. De vertekening die ontstaat wanneer positieve resultaten wel, maar negatieve of onduidelijke resultaten niet de wetenschappelijke tijdschriften halen, noemt men publicatiebias.

Ook de efficiëntie van de uitgaven aan wetenschap is in het geding. De Nederlandse overheid geeft jaarlijks zo'n zes miljard euro uit aan research en development.¹² De conclusie uit het artikel uit *The Lancet* van 2009 dat 85% van al het medisch onderzoek tot *research waste* leidt, en de observatie dat de helft van de Nederlandse wetenschappers wel eens twijfelachtige praktijken toepast die ook aan die *research waste* bijdragen, suggereren dat een substantieel deel van die zes miljard niet de waardevolle kennis voor de maatschappij oplevert die was beoogd.

Daarnaast is het vertrouwen van de samenleving in de wetenschap in het geding. Uit onderzoek van het Rathenau Instituut blijkt dat het publiek vertrouwen in de wetenschap als institutie al jaren hoog is.¹³ We weten echter ook dat burgers veel belang hechten aan de robuustheid van wetenschappelijke kennis en in het bijzonder

11 Stunt, J., van Grootel, L., Bouter, L., Trafimow, D., Hoekstra, T., & de Boer, M. (2021). Why we habitually engage in null-hypothesis significance testing: A qualitative study. *PloS one*, 16(10), e0258330.

12 Waarvan 4,9 miljard euro als subsidie en 1,1 miljard euro als fiscale steun aan bedrijven (Rathenau Instituut (2021). *Totale Investerings in Wetenschap en Innovatie 2019-2025*. Den Haag (auteurs: Broek-Honingh, N.G. van den, A. Vennekens).

13 Rathenau Instituut (2020). *Balans van de wetenschap 2020* Den Haag (auteurs: L. Koens et al.)

aan repliceerbaarheid van onderzoek.¹⁴ Het naleven van wetenschappelijke regels vinden burgers belangrijk, vooral wanneer het onderzoek gebeurt in opdracht van de overheid. Dit betekent dat voor hen de kwaliteit van onderzoek een belangrijk aandachtspunt is.

Oplossingsrichtingen

Het probleem van twijfelachtige onderzoekspraktijken staat volop in de aandacht binnen de wetenschap, die zelf verschillende richtingen aangeeft voor oplossingen. Een belangrijke rol hierin speelt de nieuwe discipline van meta-onderzoek, die zich richt op vragen over de opzet van onderzoek, de keuze van analysemethoden en het correct rapporten van alle resultaten. Deze meta-onderzoeken hebben niet alleen problemen gesignaleerd, maar ook protocollen opgeleverd om twijfelachtige onderzoekspraktijken beter te monitoren en uiteindelijk te voorkomen. We noemen er drie.

1. Preregistratie

Een eerste oplossingsrichting is preregistratie, het vooraf vastleggen van een onderzoeksplan. Op daarvoor bedoelde open onlineplatforms geven onderzoekers precies aan welke onderzoeksvraag zij willen beantwoorden, hoe ze hun onderzoek gaan opzetten en welke methoden ze gebruiken voor het verzamelen en analyseren van de data. Deze platforms, zoals het *Open Science Framework*, zijn gratis beschikbaar maar worden nog niet veel gebruikt. Het *preregistreren* is onder andere bedacht om te voorkomen dat wetenschappers uit hun analyse een resultaat selecteren dat hun artikel beter publicabel maakt, terwijl ze het antwoord op andere onderzoeksvragen niet rapporteren. Ook moet het voorkomen dat ze hun vooraf opgestelde hypothese aanpassen zodra de resultaten bekend zijn. Preregistratie betekent niet dat onderzoekers niet meer van hun plan kunnen afwijken, zolang hun keuzes in de diverse onderzoeksfases maar transparant zijn. Het neemt de mogelijkheid weg om tijdens het onderzoeksproces toe te werken naar wenselijke resultaten die beter publicabel zijn, en verkleint daarmee potentiële prikkels voor twijfelachtige onderzoekspraktijken. Preregistratie kan in een aantal gevallen dus een heel geschikt instrument zijn, maar heeft ook nadelen. De administratieve lasten, de vermeende inflexibiliteit en beperkte geschiktheid voor exploratief onderzoek maken het geen uniform toepasbare oplossing, zie ook de Rathenau-publicatie over Preregistratie.¹⁵ Voor het verminderen van *p*-hacking en gerelateerde problemen, in kwantitatief toetsend onderzoek, is preregistratie echter veelbelovend en verdient het een plek in het gesprek over het tegengaan van twijfelachtige onderzoekspraktijken.

2. Publiceren van nulresultaten

Een tweede mogelijke oplossing is het eenvoudiger maken om zogenoemde nulresultaten te publiceren in wetenschappelijke tijdschriften. Van een nulresultaat is

14 Rathenau Instituut (2021). *Vertrouwde wetenschap* – Een kwalitatieve studie naar het publieke vertrouwen in wetenschap en opdrachtonderzoek. Den Haag (auteurs: Scholvinck, A., J. Elahi, N. van den Broek-Honingh en P. Faasse)

15 Rathenau Instituut (2022) *Hoe preregistratie de kwaliteit van onderzoek kan bevorderen*. Den Haag (auteurs: Van Grootel, L. en P. Diederik)

sprake wanneer een verondersteld verband niet uit de *verzamelde* data naar voren komt. Als het eenvoudiger wordt om deze nulresultaten te publiceren, zal dat niet alleen druk wegnemen bij wetenschappers om met een aantrekkelijk onderzoeksresultaat te komen, maar ook helpen om publicatiebias te voorkomen. Verschillende wetenschappelijke tijdschriften, waaronder het hoog aangeschreven *Nature*, moedigen onderzoekers aan om hun nulresultaten in te dienen voor publicatie. Het Duitse *Berlin Institute of Health* beloofde vorig jaar duizend euro subsidie voor onderzoekers die hun nulresultaten of een replicatiestudie indienden, onder het motto *fight publication bias*. Sinds enkele jaren zijn er een paar wetenschappelijke tijdschriften die zich specifiek richten op nulresultaten (*Journal of Negative Results in...*) of een apart format hanteren om nulresultaten publicabel te maken (*Section Null Results in Brief*). Wijdverbreid is deze praktijk, die een prikkel voor de onderzoeker weghaalt om in zijn/haar resultaten 'op zoek' te gaan naar een significant resultaat, nog niet. Deze ontwikkeling die ingezet is door uitgevers van tijdschriften, vindt nog geen weerklank in nationaal wetenschapsbeleid.

3. Replicatie en reproductie

Een derde initiatief is het stimuleren van replicatie en reproductie. Replicatie staat in de wetenschap voor het herhalen van een bestaand onderzoek op basis van hetzelfde onderzoeksprotocol. Dit betekent dat er opnieuw data verzameld worden. De focus ligt daarbij vooral op het valideren van inzichten uit de eerdere studie. Wetenschappers hebben diverse standaardprotocollen ontwikkeld om replicatie voor individuele onderzoekers makkelijker te maken. Bij *reproductie* gaan onderzoekers geen nieuwe data verzamelen, maar nieuwe analyses uitvoeren op bestaande data. De focus ligt daarmee vooral op de controle van de initiële data-analyse. De trend dat steeds meer onderzoeksdata openbaar beschikbaar komen, moedigt reproductie aan. Toch lijken replicatie- en reproductieonderzoek minder aantrekkelijk voor zowel financiers als wetenschappelijke tijdschriften, en daarmee ook voor onderzoekers. Financiers steken hun middelen wellicht liever in het exploreren van onontgonnen terrein, dan in het repliceren of reproducen van wat al bekend heet te zijn. Er is binnen de wetenschap ook volop discussie over de wenselijkheid van replicatie en reproductie van meer exploratief onderzoek. In de metawetenschap wordt gesproken hoe exploratief onderzoek gereproduceerd en gerepliceerd zou kunnen worden, maar dat onderwerp is nog onderbelicht. Dat reproductie en replicatie niet zondermeer gestandaardiseerd ingezet kunnen worden, bemoeilijkt de vertaling naar beleidsmaatregelen.

Een factor die de integratie van alle drie oplossingen kan belemmeren, is hun relatieve onbekendheid. Vaak zijn ze nog geen onderdeel van de routines die onderzoekers meekrijgen in hun opleiding. Het instituut of departement waar een onderzoeker werkt, blijkt in de praktijk vaak bepalend voor de vraag of iemand bekend is met preregistratie, replicatie en reproductie en het gebruik daarvan.

Hoewel deze initiatieven vooral vanuit de wetenschap zelf komen, investeren ook Nederlandse onderzoeksfinanciers op verschillende manieren in kwaliteitsverbetering van het onderzoek. In 2016 trok ZonMw drie miljoen euro uit voor replicatieonderzoek.

Deze subsidie onderstreept het belang van replicatie en reproductie voor de wetenschap, maar zorgt er nog niet meteen voor dat dit structureel wordt ingezet voor diverse typen wetenschappelijk onderzoek.

Daarnaast stelde ZonMw voor de periode 2016-2020 een subsidie van een kleine vier miljoen euro beschikbaar voor onderzoek naar het bevorderen van verantwoorde onderzoekspraktijken, zowel op het niveau van het individu, de organisatie, als het gehele wetenschapssysteem.

Eind 2019 startten VSNU, NWO, ZonMw, KNAW en NFU het programma *Erkennen en waarderen*. Dit programma doet op vijf thema's beleidsvoorstellen om de beloningsstructuur voor onderzoekers beter in balans te brengen en perverse prikkels te beperken. Twee van deze thema's hebben ook te maken met verantwoorde onderzoekspraktijken, namelijk focussen op kwaliteit en het stimuleren van *open science*. Daarnaast hebben de KNAW, NWO en ZonMw in 2019 hun handtekening gezet onder de San Francisco Declaration on Research Assessment (DORA). De DORA is een wereldwijd initiatief uit 2012 waarin aanbevelingen worden gedaan voor onderzoekers, financiers, wetenschappelijke tijdschriften en onderzoeksinstellingen om tot evaluatiecriteria voor onderzoek te komen die meer uitgaan van de inhoudelijke waarde en impact van onderzoek en minder van bibliometrische indicatoren, zoals aantallen publicaties en citaties. Deze stap betekent dat evaluaties van Nederlandse wetenschapsorganisaties de publicatiedruk niet verder aanjagen. Verder zullen ze meer waardering geven aan open onderzoekspraktijken, waar bovengenoemde drie oplossingsrichtingen ook sterk mee verwant zijn.

Het oprichten van het Landelijk Orgaan Wetenschappelijke Integriteit en het formuleren van een Gedragscode Wetenschappelijke Integriteit, zijn andere initiatieven die georganiseerde wetenschappers hebben ondernomen om de integriteit van de Nederlandse wetenschap beter te waarborgen.

Toch zullen deze initiatieven niet alle vormen van twijfelachtige onderzoekspraktijken uit de wereld kunnen helpen. Wetenschappelijke integriteit is in het geding wanneer onderzoekers bewust onzorgvuldig handelen. Maar daarbij gaat het lang niet altijd om keuzes die bewust en opzettelijk onzorgvuldig zijn. Vele zijn het gevolg van onwetendheid, gewenning en een te grote keuzevrijheid in onderzoeksprocedures.

Een brede discussie is nodig

De dialoog over twijfelachtige onderzoekspraktijken verdient een breed platform. Opvallend is dat de huidige initiatieven om verantwoorde onderzoekspraktijken te bevorderen allemaal afkomstig zijn uit de wetenschap zelf. Het zijn onderzoekers en meta-onderzoekers die de directe gevolgen van twijfelachtige onderzoekspraktijken hebben geïdentificeerd. Zij stellen ook oplossingsrichtingen voor, zoals preregistratie en replicatieonderzoek.

Uiteraard hebben wetenschappers zelf het beste zicht op knelpunten in de dagelijkse praktijk en op de aanknopingspunten om beter gedrag te stimuleren. De oplossingen die ze hebben aangedragen, verdienen bredere bekendheid, zodat ze op grotere schaal kunnen worden uitgetest. Maar gezien de maatschappelijke gevolgen van twijfelachtige onderzoekspraktijken, dragen ook de politiek en wetenschapsorganisaties een verantwoordelijkheid. De verschillende initiatieven om twijfelachtige onderzoekspraktijken te beperken, zijn nog niet doorgedrongen tot de standaardpraktijk van wetenschappelijk onderzoek. Oplossingsrichtingen die leunen op de motivatie van individuele onderzoekers zijn kwetsbaar. Mogelijk moet de verantwoordelijkheid voor het bestrijden van twijfelachtige onderzoekspraktijken meer op het systeemniveau worden belegd, bijvoorbeeld door toezicht of audits te organiseren op onderzoeksgroepen, tijdschriften of onderzoeksprogramma's. Gezien de belangen die er op het spel staan, mag de samenleving daarover meepraten. Willekeurige burgers hebben misschien weinig zicht op methodologische kwetsbaarheden, maar zijn wel afhankelijk van de kennis die publiek gefinancierde wetenschap produceert.

De samenleving heeft bovendien ook invloed op het klimaat waarin wetenschappers actief zijn. In tijden van crisis verwachten politiek en burgers dat wetenschappers snel met resultaten komen die direct toepasbaar zijn op de problemen van het moment. De samenleving daagt wetenschappers uit tot flexibiliteit en responsiviteit, maar verwacht tegelijk een stevig fundament van wetenschappelijke waarden, procedures, methoden en theorieën. Hoe kijken burgers aan tegen het waarborgen van wetenschappelijke kwaliteit en het vernieuwen van onderzoeksmethoden door investeringen in onderzoek dat niet onmiddellijk antwoord geeft op actuele vragen? En wat zijn de verwachtingen van de samenleving over de mate van zekerheid die wetenschappelijke kennis kan bieden? Een wetenschappelijke publicatie draagt een klein stukje bij aan een veel grotere puzzel. Verschillende wetenschappers kunnen tot verschillende conclusies komen. Goede wetenschap kan niet bestaan zonder debat. Wanneer we met burgers in gesprek gaan over verwachtingen van de wetenschap en het tegengaan van twijfelachtige onderzoekspraktijken, kan dit onderwerp niet los worden gezien van een bredere visie op de werking van de wetenschap.

Welke rol speelt het probleem van twijfelachtige onderzoekspraktijken in de verwachtingen die de samenleving heeft en welke keuzes liggen er? Willen we meer gaan investeren in de methodologische kwaliteit van de wetenschap en in hoeverre mag dat ten koste gaan van snelheid en andere investeringen? Hebben we meer inzicht nodig in de mechanismen en oorzaken van twijfelachtige onderzoekspraktijken? Of moeten we onmiddellijk de kwaliteitszorg aanscherpen? Is het methodologisch onderwijs voor toekomstige onderzoekers op orde? Hoe kunnen we, door de prikkels op systematisch niveau te veranderen, onderzoekers ondersteunen in de kwaliteitsbewaking van hun studies? Dit is een aantal actuele vragen die nadere aandacht verdienen. Twijfelachtige onderzoekspraktijken, de gevolgen daarvan voor wetenschap en onderzoekers, en mogelijke oplossingen daarvoor, zijn het afgelopen decennium in de meta-wetenschap bestudeerd. De maatschappelijke gevolgen en de

maatschappelijke bijdrage aan oplossingen zijn dat nog onvoldoende. De verantwoordelijkheid daarvoor kan niet alleen bij de wetenschap liggen. Naast wetenschappers moeten beleidsmakers, subsidieverstrekkers en politici het gesprek voeren over de wenselijkheid van de geboden oplossingen en de mogelijkheden voor institutionele inbedding ervan.

Het Rathenau Instituut zet zich graag in om dit gesprek mogelijk te maken.

Auteurs

Leonie van Grootel, Laurens Hessels en Paul Diederer

Bij voorkeur citeren als:

Rathenau Instituut (2022). *Twijfelachtige onderzoekspraktijken zijn een zorg voor de hele samenleving*. Essay. Den Haag. Auteurs: Van Grootel, L., L. Hessels en P. Diederer

© Rathenau Instituut 2022

Verveelvoudigen en/of openbaarmaking van (delen van) dit werk voor creatieve, persoonlijke of educatieve doeleinden is toegestaan, mits kopieën niet gemaakt of gebruikt worden voor commerciële doeleinden en onder voorwaarde dat de kopieën de volledige bovenstaande referentie bevatten. In alle andere gevallen mag niets uit deze uitgave worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie of op welke wijze dan ook, zonder voorafgaande schriftelijke toestemming.

Open Access

Het Rathenau Instituut heeft een beleid voor open access. Rapporten, achtergrondstudies, wetenschappelijke artikelen, software worden vrij beschikbaar gepubliceerd. Onderzoeksgegevens komen beschikbaar met inachtneming van wettelijke bepalingen en ethische normen voor onderzoek over rechten van derden, privacy, en auteursrecht.

Contactgegevens

Anna van Saksenlaan 51
Postbus 95366
2509 CJ Den Haag
070-342 15 42
info@rathenau.nl
www.rathenau.nl

Twijfelachtige onderzoekspraktijken zijn een zorg voor de hele samenleving

Het Rathenau Instituut stimuleert de publieke en politieke meningsvorming over de maatschappelijke aspecten van wetenschap en technologie. We doen onderzoek en organiseren het debat over wetenschap, innovatie en nieuwe technologieën.