

DE REPRODUCEERBAARHEID VAN HET ONDERZOEK IN VLAANDEREN: FEITENONDERZOEK EN AANBEVELINGEN

KVAB Denkersprogramma 2022

Sabina Leonelli
Stephan Lewandowsky



KVAB STANDPUNTEN

81

Koninklijke Vlaamse Academie van België
voor Wetenschappen en Kunsten - 2022

**DE REPRODUCEERBAARHEID
VAN HET ONDERZOEK
IN VLAANDEREN:
FEITENONDERZOEK EN
AANBEVELINGEN**

KVAB DENKERSRAPPORT 2022



KVAB Press

KVAB STANDPUNTEN

81

Cover concept: Francis Strauven
Ontwerp cover: Charlotte Dua
Afbeelding: Shutterstock

De tekening van het Paleis der Academiën is een reproductie van het originele perspectief van Charles vander Straeten in 1823. Jozef Cantré ontwierp het logo van de KVAB in 1947.

De KVAB Standpunten worden gepubliceerd door de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Hertogsstraat 1, 1000 Brussel.
Tel. 00 32 2 550 23 23 – info@kvab.be – www.kvab.be

DE REPRODUCEERBAARHEID VAN HET ONDERZOEK IN VLAANDEREN: FEITENONDERZOEK EN AANBEVELINGEN

KVAB DENKERSRAPPORT 2022

Sabina Leonelli
Stephan Lewandowsky



Gedeeltelijke reproductie is toegelaten mits uitdrukkelijke bronvermelding.
Partial reproduction is permitted provided the source is mentioned.
Aanbevolen citeerwijze: Sabina Leonelli, Stephan Lewandowsky, *De
reproduceerbaarheid van het onderzoek in Vlaanderen: Feitenonderzoek en
aanbevelingen*, KVAB Denkersrapport 2022, KVAB Standpunt 81, 2022.

© Copyright 2022 KVAB
D/2022/0455/08
ISBN 978 90 656 921 77

Drukkerij Universa

DE REPRODUCEERBAARHEID VAN HET ONDERZOEK IN VLAANDEREN: FEITENONDERZOEK EN AANBEVELINGEN KVAB DENKERSPROGRAMMA 2022

INHOUDSOPGAVE

Inleiding en overzicht	16
Methoden	18
DEEL 1. De vele gezichten van reproduceerbaarheid	20
<i>KADER 1: Reproduceerbaarheid vereist niet noodzakelijk replicatie</i>	21
<i>Zes vormen van reproduceerbaarheid</i>	23
<i>Enge en ruime interpretaties van reproduceerbaarheid</i>	26
<i>KADER 2: Pre-registratie</i>	30
DEEL 2. Implementatie van reproduceerbaarheid binnen academische instellingen	33
<i>Implementatiepraktijken</i>	33
<i>Spanningen tussen belanghebbenden</i>	36
DEEL 3. De spanning tussen veerkracht en reproduceerbaarheid	38
<i>KADER 3. Wetenschapscommunicatie</i>	38
<i>KADER 4: The Belfast Project en gevolgen</i>	43
<i>KADER 5: Reproduceerbaarheid en transparantie: belangrijkste uitdagingen</i>	43
DEEL 4. Het spanningsveld tussen openbaar en particulier belang	47
<i>KADER 6: Onderzoekinstellingen aan de translationele spits</i>	48
Conclusies: Hoe helpt het streven naar reproduceerbaarheid ons de wetenschappelijke crisis aan te pakken?	54
Referenties	56

Samenvatting

Dit rapport is opgesteld in het kader van een Denkersprogramma van de KVAB. Het onderzoekt drie aspecten van reproduceerbaarheid met als doel om aan belanghebbenden aanbevelingen te formuleren over de manier waarop de reproduceerbaarheid van wetenschappelijk onderzoek kan worden gewaarborgd. Replicatie op zich is slechts één aspect van reproduceerbaarheid, en om reproduceerbaarheid grondig te begrijpen moet een bredere definitie worden gehanteerd die componenten omvat zoals transparantie, onderzoekspraktijken en de rol van theorie in de wetenschap. Ten tweede kan het verband tussen deze componenten en reproduceerbaarheid complex en contra-intuïtief zijn. Ten derde, als wetenschappelijk onderzoek robuuste, reproduceerbare en geloofwaardige kennis wil opleveren, dan moeten we de onderzoeksinstellingen, de evaluatie van wetenschappelijk onderzoek en de onderzoekspraktijken zo veranderen dat er uiteindelijk niet meer over reproduceerbaarheid gepraat hoeft te worden. We zullen reproduceerbaarheid hebben bereikt wanneer we er niet meer over discussiëren. Reproduceerbaarheid is belangrijk, maar ze neemt verschillende vormen aan die passen bij verschillende onderzoekssituaties, afhankelijk van de specifieke doelstellingen, methoden, materialen en omstandigheden van het onderzoek. Gezien deze noodzakelijke variatie moet reproduceerbaarheid worden geconceptualiseerd en toegepast op een manier die niet te eng is, zodat het onderzoek niet onbedoeld wordt beperkt op een manier die niet nuttig is, noch te ruim om zinvol te zijn.

Uitgebreid overleg met belanghebbenden, meer bepaald met vertegenwoordigers van universiteiten, onderzoekscentra, koepelorganisaties, de farmaceutische industrie en financierende instellingen, leverde de Denkers inzichten op over huidige praktijken en mogelijke toekomst in het Vlaamse wetenschappelijk onderzoek.

Positieve acties en aandachtspunten in onderzoeksinstellingen

- Een erkenning van de rol van data stewards als integraal onderdeel van het onderzoeksecosysteem, waarbij veel universiteiten investeren in permanente posities voor data stewards om ondersteuning te bieden bij dataverzameling, -conservering en -beheer.
- Een groter bewustzijn van het belang van integriteit binnen het onderzoek, met als gevolg een toename van ethische commissies en integriteitsfunctionarissen aan verschillende universiteiten.
- Nieuwe opleidingsmiddelen voor onderzoekers en professioneel ondersteunend personeel om te leren hoe zij hun werk methodologisch meer verantwoord en robuuster kunnen maken, MAAR nog steeds een gebrek aan erkenning voor de expertise en inspanning die daarmee gepaard gaan.

Bevindingen voor de financierende instellingen

- Een bereidheid om rekening te houden met integriteit en Open Science-praktijken bij financierende instellingen zoals het FWO, waar negatieve resultaten momenteel worden erkend als een legitieme output voor onderzoeksbeoordelingen; etikettering van dergelijke overwegingen als "tweede as" voor onderzoek – MAAR kan worden geïnterpreteerd als een bevestiging dat kwesties van integriteit, kwaliteit en reproduceerbaarheid geen prioriteit hebben.
- Moeilijkheden bij de beoordeling van transdisciplinair onderzoek, dat veel instellingen in theorie van het grootste belang achten om mondiale uitdagingen aan te gaan, maar dat in de praktijk zeer moeilijk te beoordelen blijft.
- De erkenning van Open Science als een belangrijk fenomeen en pogingen om een cultuurverandering teweeg te brengen binnen academische instellingen. Dergelijke pogingen kunnen concrete vormen aannemen, zoals de wijziging van de criteria voor de aanwerving of promotie van academici zodat ze inspanningen omvatten op het gebied van dataconservering, reproductie van andere studies en kwaliteitsbeoordeling.

Spanningen tussen belanghebbenden

- Spanning tussen enerzijds door de overheid gefinancierde onderzoekers, voornamelijk aan universiteiten, die geacht worden volledig transparant te zijn, en anderzijds onderzoekers die in het bedrijfsleven werken en die door regelgevende instanties worden gecontroleerd, maar niet door academische vakgenoten.
- Spanningen tussen onderzoekers met verschillende niveaus van anciënniteit, voornamelijk omdat de extra kosten en inspanningen die nodig zijn voor het uitvoeren van onderzoek volgens een hoger criterium van reproduceerbaarheid worden gedragen door junioronderzoekers, die waarschijnlijk meer relevante vaardigheden hebben verworven.
- Belangrijk probleem bij het streven naar reproduceerbaarheid: de institutionele belonings- en stimuleringsmechanismen blijven achter en draaien nog steeds om conventionele maatstaven.
- Zelfs wanneer instellingen de prestatiecijfers aanvullen met een kwalitatieve evaluatie, waardoor inspanningen op het gebied van reproduceerbaarheid worden beloond, blijft de nadruk liggen op originaliteit en innovatie boven replicatie of kwaliteit.
- Reviewing is een arbeidsintensieve activiteit, maar is vrijwillig en bijna volledig onbeloond.

Wetenschapscommunicatie

- Misvatting van het zogenaamde "tekortmodel", d.w.z. dat het publiek kennis mist die wetenschapsvoorlichters kunnen verschaffen, zodat het informatietekort bij het publiek wordt verholpen.

- Om het publiek in staat te stellen de waarde van wetenschappelijke beweringen te beoordelen, is het noodzakelijk de wetenschap te begrijpen als een zichzelf corrigerend onderzoeksproces, eerder dan als een reeks onfeilbare waarheden.
- De opvatting van wetenschappelijke communicatie als "informatieverstrekking" biedt geen mogelijkheid voor niet-onderzoekers om een dialoog aan te gaan en hun eigen expertise in te brengen (burgerwetenschap / "citizen science").
- Meerderheid van publieke betwistingen van gegevens door gereguleerde sectoren, lobbyisten en handelsorganisaties.
- Het algemene gebruik van "open gegevens" bij onderzoek dat gevoelige informatie omvat, is problematisch en in strijd met de doelstellingen van diversiteit en inclusiviteit.
- De afwegingen tussen transparantie en privacy en tussen reproduceerbaarheid, veerkracht, vertrouwen en eerbiediging van de privacy moeten eerlijk en grondig worden onderzocht.

Opmerkingen voor onderzoek in het openbaar en particulier belang

- De inzet is hoger voor commercieel onderzoek, en met name onderzoek dat gericht is op het omzetten van wetenschappelijke inzichten in handelswaar voor de wereldmarkt, gekenmerkt door een enorme druk om resultaten te reproduceren en zeer nauwkeurige en gestandaardiseerde kwaliteitscontroles toe te passen.
- Zowel publiek als privaat gefinancierde onderzoekers worden blootgesteld aan stimulansen en eisen om zo snel mogelijk resultaten te produceren, waarbij weinig tijd wordt besteed aan het onderzoeken van implicaties en het documenteren en naar behoren onderzoeken van onderzoeksprocessen.
- Voor particulier gefinancierde onderzoekers zijn er ernstige belemmeringen voor het vrij delen – en breed onderzoeken – van gegevens.
- Bij door de overheid gefinancierd onderzoek bestaat de neiging geheimhouding te verkiezen boven open samenwerking, vooral in drukbezette domeinen zoals de biogeneeskunde, en te veel te vertrouwen op geautomatiseerde systemen.
- Voor onderzoeksinstituten die geavanceerd translationeel onderzoek verrichten, is er geen institutioneel orgaan dat erop toeziet dat hun transnationale onderzoek maatschappelijk verantwoord is.

Executive summary

The report is produced in the context of a Thinkers' program of KVAB. It explores three aspects of reproducibility with the goal of providing recommendations to stakeholders about how to ensure reproducibility. Replication *per se* is only one aspect of reproducibility, and to thoroughly understand reproducibility requires casting a wider net that includes components such as transparency, research practices, and the role of theory in science. Second, the relationship between these components and reproducibility can be complex and counterintuitive. Third for science to yield robust, reproducible and credible knowledge, we must transform research institutions, evaluation and practices so that, ultimately, it is no longer necessary to talk about it. We will have achieved reproducibility when we no longer debate it. Reproducibility matters, but it takes different forms suited to different research situations, depending on the specific goals, methods, materials, and conditions of research. Given this necessary variation, reproducibility needs to be conceptualised and implemented in ways that are not too narrow, thereby inadvertently constraining research in ways that are unhelpful, nor too broad to be meaningful.

Extensive consultations with stakeholders provided the Thinkers with insights about current practices and possible futures in Flemish research, more specifically, Universities, Research Centers, Umbrella organizations, Pharmaceutical industry and Funders.

Positive actions and attention points in research institutes

- A recognition of the role of data stewards as integral parts of the research ecosystem, with many universities investing in permanent data steward positions to offer support in data collection, curation and management.
- A heightened awareness of the significance of integrity within research, resulting in a boost of ethics boards and integrity officers at various universities.
- Novel training resources for researchers and professional support staff to learn how to make their work more responsible and robust methodologically, BUT still a lack of recognition for the expertise and effort involved.

Findings for the funders

- A willingness to take account of integrity and Open Science practices by funders such as the FWO, where negative results are currently recognised as a legitimate output for research assessments, label of such considerations as "second axis" for research: BUT can be interpreted as a confirmation that questions of integrity, quality and reproducibility are not a priority.
- Difficulties to assess transdisciplinary research, which many institutions regard as paramount to confront global challenges in theory, but in practice remains very hard to assess.

- The recognition of Open Science as an important phenomenon and attempts to instigate a culture change within academic institutions. Such attempts can take substantive forms, such as modification of criteria for hiring or promoting academics so that they include efforts in data curation, reproduction of other studies and quality assessment.

Tensions among stakeholders

- Tension between publicly funded researchers, mainly in universities, who are expected to be fully transparent on the one hand, and researchers working in industry on the other, who receive scrutiny from regulatory bodies but are not required to undergo scrutiny by academic peers.
- Tensions between researchers at different levels of seniority, mainly because the additional cost and effort involved in conducting research to a higher criterion of reproducibility is borne by junior researchers who are more likely to have acquired relevant skills.
- Key perceived problem in the pursuit of reproducibility the institutional reward and incentive structures are lagging behind and continue to revolve around conventional metrics.
- Even when institutions complement performance metrics with qualitative evaluation, which provides rewards for reproducibility efforts, the emphasis continues to be on novelty over replication or quality.
- Reviewing is a labour-intensive activity but is voluntary and almost completely unrewarded.

Science communication

- Misconception of the so-called "deficit model"; i.e. the public is lacking knowledge that science communicators can provide, thus remedying the public's information deficit
- For publics to be able to assess the value of scientific claims, it is imperative to understand science as a self-correcting process of inquiry, rather than as a set of infallible truths.
- View of scientific communication as 'information provision' does not provide any avenue for non-researchers to enter in a dialogue and to contribute their own expertise (citizen science).
- Majority of public challenges to data by regulated industries, lobbyists, and trade organizations.
- The blanket use of "open data" when research involves sensitive information is problematic and in conflict with the goals of diversity and inclusivity.
- The trade-offs between transparency and privacy and between reproducibility, resilience, trust, and respect for privacy, must be examined frankly and thoroughly.

Observations for research in public and private interest

- The stakes are higher for commercially exposed research, and particularly research devoted to transforming scientific insights into commodities for global markets, characterized by huge pressure to reproduce results and use very precise and standardised quality checks.
- Both publicly and privately funded researchers are exposed to incentives and demands towards producing outputs as fast as possible, with little time devoted to exploring implications and to documenting and properly scrutinizing research processes.
- For privately funded researchers there are severe obstacles to the free sharing and wide-ranging scrutiny - of data.
- For publicly funded research there is a tendency to privilege secrecy over open collaboration, especially in crowded domains such as biomedicine; and to overly rely on automated systems
- For Research Institutes at the Translational Cutting Edge there is no institutional body in charge of overseeing whether their transnational research carried is socially responsible.

Voorwoord

De reeks Standpunten van de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten draagt bij tot een wetenschappelijk onderbouwd debat over actuele maatschappelijke en artistieke thema's. De auteurs, leden en werkgroepen van de Academie, schrijven in eigen naam, onafhankelijk en met volledige intellectuele vrijheid. De goedkeuring voor publicatie door één of meerdere Klassen van de Academie waarborgt de kwaliteit van de tekst. Dit Standpunt werd goedgekeurd voor publicatie door de Klasse van de Technische Wetenschappen op 17 september 2022 en de Menswetenschappen op 19 november 2022.

De organisatie van het Denkersprogramma 2022 door de KVAB in samenwerking met de Koninklijke Academie voor Geneeskunde van België (KAGB) en De Jonge Academie

Patrick Onghena (KMW), Joos Vandewalle (KTW) and Inez Dua (KVAB-medewerker)

Elk jaar organiseert de KVAB twee Denkersprogramma's op initiatief van één van haar klassen en/of reflectiegroepen. Het programma "Reproduceerbaarheid en repliceerbaarheid van onderzoek" werd voorgesteld door de permanente reflectiegroep Responsible Research & Innovation (RRI) en Wetenschapsethiek tijdens zijn online vergadering in maart 2021 in een pandemische situatie.

Context: In een typisch Denkersprogramma wordt een centrale rol toegekend aan één of twee vooraanstaande internationale deskundigen. Het voorgestelde thema "Reproduceerbaarheid en repliceerbaarheid van onderzoek" is gekoppeld aan een goed ontvangen KVAB Standpunt "Repliceerbaarheid in de empirische menswetenschappen", geschreven door één van onze academieleden, Patrick Onghena, en gebaseerd op zijn academische presentatie van 2018. Bovendien zijn er recent veel rapporten van academies verschenen over dit belangrijke thema: het rapport van de Koninklijke Nederlandse Academie van Wetenschappen (KNAW) inzake replicatiestudies "Improving reproducibility in the empirical sciences" (2018), het rapport van de Amerikaanse National Academies of Sciences, Engineering, and Medicine "Reproducibility and Replicability in Science" (2019), het rapport van de Britse Academy of Medical Sciences "Reproducibility and reliability of biomedical research: improving research practice" (2015), en het overzichtsrapport van de EU "Reproducibility of scientific results in the EU: Scoping Report" (november 2020).

Thema: Het voorgestelde thema "Reproduceerbaarheid en repliceerbaarheid van onderzoek" is vrij breed, maar het is zeker bedoeld voor de hele verzameling van wetenschappelijke disciplines, elk met hun eigen specificiteit. Op 18 oktober 2018 publiceerde Nature een speciaal nummer over de uitdagingen van niet-reproduceerbaar onderzoek, waarin werd gesteld: "Wetenschap gaat vooruit door corroboratie – wanneer onderzoekers de resultaten van anderen verifiëren. De wetenschap gaat sneller vooruit als mensen minder tijd verspillen aan het volgen van valse aanwijzingen. Geen enkel onderzoeksdocument kan ooit worden beschouwd als het laatste woord, maar er zijn er te veel die verdere studie niet doorstaan."

Uiteindelijk is het vertrouwen van de samenleving en het bredere publiek in de wetenschap en de maatschappelijke waarde van wetenschappelijke resultaten deels gebaseerd op de reproduceerbaarheid en repliceerbaarheid van de resultaten. Het probleem doet zich in verschillende vormen voor in de verschillende wetenschappelijke domeinen. Transparantie en openheid van de gegevens,

de verwerkingsmethoden, een betere statistische verwerking en openheid van de software kunnen hier zeker toe bijdragen. Bovendien kunnen succesvolle benaderingen uit zwaar getroffen gebieden zoals de psychologie een inspiratiebron vormen voor andere gebieden. Om de reikwijdte te beperken moest de stuurgroep in interactie met de Denkers een focus bespreken. Centrale thema's zijn zeker de "aanpak van meer transparantie, de verbetering van de kwaliteit van het onderzoek, goede praktijken, ethische kwesties en vertrouwen van het publiek in het onderzoek", en niet zozeer de repressieve methoden van slecht of nauwelijks replicerbaar onderzoek.

Activiteiten: Het voorstel is in mei 2021 door de Raad van Bestuur van de KVAB aanvaard en in juli 2021 is een startnota uitgewerkt. Vervolgens werd de samenwerking met de KAGB en de Jonge Academie veilig gesteld. In juli 2021 werd een stuurgroep samengesteld met leden van de KVAB, de Jonge Academie, de KAGB en verschillende externe deskundigen (zie bijlage 3). De rol van de stuurgroep bestond erin de activiteiten van de Denkers goed te onderbouwen en de nodige input te leveren. Er werd dus geen directe sturing verwacht, maar eerder ondersteuning. In de zomer van 2021 zijn de twee Denkers geselecteerd: *Sabina Leonelli*, wetenschapsfilosofe aan de Universiteit van Exeter, Turing Institute, Global Young Academy, Open Science Group, en *Stephan Lewandowsky*, professor cognitieve psychologie aan de Universiteit van Bristol (zie bijlage 2 voor hun cv). De Denkers kregen alle vrijheid en bleven volledig onafhankelijk bij het schrijven van hun rapport met aanbevelingen. Het was de bedoeling dat de Denkers samen met de stuurgroep en tal van partners en belanghebbenden rond het gekozen thema zouden werken en een belangrijke bijdrage zouden leveren aan de verdere wetenschappelijke strategie van Vlaanderen door een langetermijnvisie te ontwikkelen en zo bij te dragen tot de beleidsvorming. Door middel van talrijke activiteiten kunnen concrete observaties, ideeën en aanbevelingen worden ontwikkeld.

De stuurgroep had een eerste online vergadering op 15 oktober 2021, gevolgd door een online kick-off meeting met de Denkers op 29 november 2021, waarin de verdere planning van de bijeenkomsten met belanghebbenden in open of gesloten workshops en bezoeken werd voorbereid. De Denkers hebben relevante vragen opgesteld en met de stuurgroep gedeeld. Deze zijn vooraf aan de belanghebbenden gegeven en in het voorjaar van 2022 in verschillende online en fysieke bijeenkomsten met de Denkers besproken. Op 1 juni 2022 werd een drukbezocht slotsymposium gehouden met de bevindingen van de Denkers. De presentaties van de Denkers zijn toegankelijk gemaakt op YouTube.

<https://kvab.be/nl/activiteiten/final-symposium-reproducibility-and-replicability-science>

Tijdens het symposium hadden het publiek en andere sprekers verschillende interessante interacties met de Denkers. Op basis van deze wederzijdse communicatie hebben de Denkers hun rapport opgesteld.

Het is een genoegen iedereen te bedanken die dit programma tot een succes heeft gemaakt: de stuurgroep, de belanghebbenden, het publiek en de medewerkers van de KVAB. In het bijzonder willen we de twee toegewijde en actieve Denkers, Sabina en Stephan, bedanken voor een uitstekend rapport. Zij zijn erin geslaagd de informatie uit de initiële interviews en de discussies tijdens het symposium van juni op een gedachtenprikkelende manier samen te brengen en te integreren. Het rapport vormt een uitstekende basis voor de Vlaamse wetenschappelijke gemeenschap om de discussie uit te breiden en de focus te houden op het verbeteren en verdedigen van de wetenschap.

Inleiding en overzicht

Reproduceerbaarheid is fundamenteel voor de wetenschap als een manier om de geloofwaardigheid van resultaten en procedures te verifiëren. Bevindingen die niet kunnen worden gerepliceerd en wetenschappelijke procedures die niet kunnen worden gereproduceerd kunnen niet bijdragen tot kennis en verspillen in het slechtste geval de tijd van onderzoekers wanneer ze een doodlopende weg volgen op basis van onbetrouwbare resultaten. Zoals een editorial van het tijdschrift *Nature* het verwoordde op 18 oktober 2018: "Wetenschap gaat vooruit door corroboratie – wanneer onderzoekers de resultaten van anderen verifiëren." Reproduceerbaarheid is ook lang beschouwd als een pijler van de scheidslijn tussen wetenschap en pseudowetenschap. Horoscopen, homeopathie, de theorie van de holle aarde en andere pseudowetenschappelijke claims bieden geen reproduceerbare inzichten, maar hooguit vermakelijke verhalen (Zwaan, Etz, Lucas, & Donnellan, 2017). Bovendien is het vertrouwen van de samenleving, beleidsmakers en het bredere publiek in de wetenschap en wetenschappelijke resultaten verbonden met de reproduceerbaarheid en replicerbaarheid van de resultaten. Wetenschappelijk advies kan legitiem veranderen door nieuwe kennis, vooral tijdens een snel evoluerende crisis zoals het begin van de COVID-19-pandemie, en het is belangrijk dat het publiek begrijpt dat wetenschappelijke kennis gepaard gaat met onzekerheidsgrenzen (Joslyn & LeClerc, 2016); niettemin is wetenschappelijke kennis alleen betrouwbaar voor zover hij het resultaat is van betrouwbare en goed onderzochte onderzoeksprocessen, en reproduceerbaarheid is een belangrijk criterium om de kwaliteit en de legitimiteit van wetenschappelijke beweringen te evalueren.

Er moet daarom groot belang worden gehecht aan het feit dat talrijke experimentele bevindingen in uiteenlopende disciplines de laatste tijd niet reproduceerbaar bleken te zijn. In de psychologie bleek uit een grootschalige poging om 100 gepubliceerde bevindingen te repliceren dat de effectgroottes in de replicaties gemiddeld slechts de helft bedroegen van die van de oorspronkelijke publicaties en dat slechts 25 % en 50 % van de bevindingen gerepliceerd werden in respectievelijk de sociale psychologie en de cognitieve psychologie (Open Science-samenwerking, 2015). Hoewel een groot deel van de discussie over reproduceerbaarheid gericht was op de psychologie, hebben talrijke andere disciplines soortgelijke problemen ondervonden. In een parallelle poging om de replicerbaarheid in de economie vast te stellen, bleek bijvoorbeeld minder dan de helft van de resultaten van 60 gepubliceerde artikelen reproduceerbaar te zijn (Chang & Li, 2015); en pogingen om gepubliceerde biomedische bevindingen onafhankelijk te verifiëren vonden significante fouten in de gegevensrapportage en -analyse (Allison et al., 2016). Bovendien bevatten wetenschappelijke publicaties vaak onvoldoende details (over bijvoorbeeld methoden, data en codes) om pogingen tot reproductie van de resultaten te ondersteunen, waardoor het onmogelijk is om de betrouwbaarheid

van die studies op deze manier te controleren (National Academies of Sciences, Engineering, and Medicine, 2019).

Het ontstaan van de “replicatiecrisis” heeft zowel binnen de wetenschappelijke gemeenschap als bij het publiek en de belanghebbenden tot heftige discussies geleid. Er is een grote verscheidenheid aan standpunten naar voren gekomen, en hoewel de kwestie op het moment van dit schrijven nog lang niet is opgelost, lijken er drie consensuele standpunten te zijn ontstaan:

1. Replicatie op zich is slechts één aspect van reproduceerbaarheid, en om reproduceerbaarheid grondig te begrijpen moet een bredere definitie worden gehanteerd die componenten omvat zoals transparantie, onderzoekspraktijken en de rol van theorie in de wetenschap.
2. Het verband tussen deze componenten en reproduceerbaarheid kan complex en contra-intuïtief zijn.
3. Wil de wetenschap robuuste, reproduceerbare en geloofwaardige kennis opleveren, dan moeten we de onderzoeksinstellingen, de evaluatie en de praktijken zo veranderen dat er uiteindelijk niet meer over reproduceerbaarheid gepraat hoeft te worden. We zullen reproduceerbaarheid hebben bereikt wanneer we er niet meer over discussiëren.

De rest van dit rapport onderzoekt deze drie aspecten van reproduceerbaarheid met als doel aanbevelingen te doen aan belanghebbenden over hoe de reproduceerbaarheid kan worden gewaarborgd. Op basis van de instructies die we ontvingen, was het onze bedoeling een brede reeks aanbevelingen te doen die relevant zijn voor alle disciplines en de deur openzetten voor verdere discussie, eerder dan specifieke acties voor te schrijven.

Methoden

Het rapport is gebaseerd op verschillende rondes van intensieve bijeenkomsten met belanghebbenden in heel Vlaanderen. De volgende belanghebbenden hebben aan deze gesprekken bijgedragen:

- Universiteit Gent
- VUB
- Universiteit Antwerpen
- Universiteit Hasselt
- Onderzoekscentra (vertegenwoordigers van Imec, VITO, Flanders Make, VLIZ en ILVO)
- Koepelorganisaties (De Jonge Academie, ReproducibiliTea journal club, Open Access Belgium)
- De farmaceutische industrie (vertegenwoordigers van Janssen Pharmaceutica, onderdeel van Johnson & Johnson, en GSK)
- Financiers (vertegenwoordigers van FWO, VLAIO)

Hoewel de gesprekken met de belanghebbenden flexibel en open waren, werden ze gevoed met een lijst van vragen die vóór elke bijeenkomst werd verspreid. De volgende vragen werden aan de belanghebbenden voorgelegd:

- Hoe interpreteert u de begrippen reproduceerbaarheid en repliceerbaarheid? Beschouwt u ze als relevant voor de huidige praktijken in uw organisatie/projecten (zo ja, hoe; zo nee, waarom niet)?
- Is reproduceerbaarheid altijd een noodzakelijk ingrediënt van een betrouwbare en robuuste kennisproductie?
- Wat is de relatie tussen reproduceerbaarheid en een bevestigingsvooroordeel (indien aanwezig)?
- Wat denkt u over de relatie tussen reproduceerbaarheid en Open Science?
- Welke institutionele structuren en/of stimulansen zijn er om de reproduceerbaarheid te ondersteunen?
- Verwacht u negatieve gevolgen van de invoering van deze structuren?
- Kunt u goede praktijken inzake reproduceerbaarheid en/of rolmodellen aanwijzen (zowel op het niveau van individuen als van instellingen)?
- Met welke discipline(s) werkt u? Gebruikt u kwalitatieve, kwantitatieve of gemengde methoden?
- Wat is eventueel de rol van de theorie bij het tot stand brengen van reproduceerbaar onderzoek?
- Wordt in de communicatie met het publiek en de pers rekening gehouden met reproduceerbaarheid?
- Hoe wegen we reproduceerbaarheid/repliceerbaarheid af tegen actualiteit in een crisissituatie, zoals de huidige pandemie?
- Is er behoefte aan een speciale opleiding voor jonge onderzoekers, promovendi en masterstudenten?

- Welke stimulansen zijn er nodig om een dergelijke opleiding uitvoerbaar te maken?
- Is disciplinair onderzoek meer of minder reproduceerbaar dan interdisciplinair onderzoek, en waarom?
- Kunt u één of meer concrete scenario's beschrijven die uw opvatting hierover illustreren?
- Denkt u dat kunstmatige intelligentie (AI) het streven naar reproduceerbaarheid kan ondersteunen of belemmeren, en zo ja, hoe?
- Kunt u voorbeelden geven van bekende bevindingen in uw vakgebied die later niet reproduceerbaar of niet replicerbaar bleken te zijn?

De gesprekken met de belanghebbenden zijn uitgebreid genoteerd en besproken door de auteurs, en de input is in dit verslag verwerkt. Waar relevant belichten we in het rapport specifieke punten van zorg die door belanghebbenden naar voren zijn gebracht.

DEEL 1. De vele gezichten van reproduceerbaarheid

Reproduceerbaarheid kan ruim worden gedefinieerd als een overkoepelende wetenschappelijke waarde, die vaak wordt genoemd in discussies over wat "goede praktijken" of "valide onderzoek" zijn, en die de mate aangeeft waarin consistente resultaten worden verkregen wanneer een onderzoek wordt herhaald. Het idee van reproduceerbaarheid is nauw verbonden met debatten over Open Science, vanwege de gemeenschappelijke nadruk op het delen van informatie die kan worden gebruikt om na te gaan of een onderzoek al dan niet reproduceerbaar is.

Een belangrijke uitdaging die onmiddellijk naar voren komt in discussies over reproduceerbaarheid in alle groepen van belanghebbenden is die van de diversiteit. Wetenschap is een zeer gefragmenteerde onderneming, met ontelbare deelgebieden die methoden en instrumenten gebruiken die aanzienlijk van elkaar verschillen, omdat ze allemaal zijn toegesneden op specifieke materialen, problemen en doelstellingen. De belangrijkste benaderingen van onderzoek kunnen zowel kwalitatieve als kwantitatieve methoden omvatten, zowel experimentele benaderingen als observationele studies en computersimulaties, en zowel theoriegerichte als gegevensintensieve onderzoeksmethoden. Deze wetenschappelijke diversiteit wordt versterkt door verschillende factoren, waaronder: de kenmerken van de door de onderzoekers onderzochte entiteiten, waarbij studies van menselijke populaties aanzienlijk verschillen van studies van biologische monsters of nanodeeltjes; de verschillen tussen de onderzoeksculturen, middelen en doelstellingen die ten grondslag liggen aan onderzoek op verschillende locaties; de stimulansen en prioriteiten die zijn vastgesteld door de instellingen die het onderzoek ondersteunen; en de taakverdeling die in verschillende wetenschappelijke omgevingen wordt gehanteerd, met specifieke en uiteenlopende verantwoordelijkheden die worden toegewezen aan beginnende onderzoekers, senior academici, ondersteunend personeel, bibliothecarissen, technici, gegevenswetenschappers en gegevensbeheerders. Zoals de filosoof Hans Radder al in 1996 opmerkte, moeten we, gezien deze overweldigende diversiteit, "om het begrip reproduceerbaarheid te verduidelijken de volgende vraag beantwoorden: reproduceerbaarheid van wat en door wie?" (Radder, 1996, p. 16).

Algemeen wordt erkend dat er een verschil bestaat tussen pogingen om onderzoeksprocessen te reproduceren en pogingen om de resultaten van dergelijke processen te reproduceren. De Amerikaanse National Academy of Sciences (NAS) heeft onlangs naar dit verschil verwezen door onderscheid te maken tussen replicerbaarheid en reproduceerbaarheid, waarbij de eerste werd gedefinieerd als "het verkrijgen van consistente resultaten tussen studies gericht op het beantwoorden van dezelfde wetenschappelijke vraag, die elk hun eigen gegevens hebben verkregen" (National Academies of Sciences, Engineering, and Medicine,

2019), en de tweede als “het verkrijgen van consistente berekeningsresultaten met gebruikmaking van dezelfde inputgegevens, berekeningsstappen, methoden, code en analysevoorwaarden” (National Academies of Sciences, Engineering, and Medicine, 2019; zie ook commentaar in Onghena 2020). Hoewel we de waarde van dit fundamentele onderscheid erkennen, wijzen we in wat volgt op een meer complexe en gediversifieerde manier waarop wetenschappelijke domeinen tegen het idee van reproduceerbaarheid aankijken (zie kader 1).

KADER 1: Reproduceerbaarheid vereist niet noodzakelijk replicatie

Bezorgdheid over reproduceerbaarheid en robuuste wetenschap is gewoonlijk nauw verbonden met de veronderstelde noodzaak van replicatie. Replicatie mainstream maken (“Making replication mainstream”, Zwaan et al., 2017) is een verklaard doel van veel wetenschappers die zich zorgen maken over reproduceerbaarheid en de huidige “replicatiecrisis” die veel disciplines in zijn greep heeft. Er is echter betrekkelijk weinig onderzoek gedaan naar de omstandigheden waaronder replicaties al dan niet raadzaam zijn. Moet elke studie worden gerepliceerd? Zo ja, door wie? Als er een beslissing over replicatie moet worden genomen, waarop moet deze dan gebaseerd zijn? Moeten we altijd repliceren voordat we een nieuwe voorspelling testen? Het blijkt dat het venijn in de details zit als het gaat om beslissingen over replicatie. Soms zijn die details opvallend contra-intuïtief. We beschouwen twee recente voorbeelden.

Het eerste voorbeeld betrof de vraag of onderzoekers een bevinding altijd moeten repliceren voordat ze deze publiceren (Lewandowsky & Oberauer, 2020). Om deze vraag te beantwoorden, hebben Lewandowsky en Oberauer de totale prestaties van een hypothetische wetenschappelijke gemeenschap gemodelleerd, waarbij ze zich concentreerden op de kosten en baten van het genereren van kennis onder verschillende replicatieregimes. De boodschap van die modelstudie was dat de publicatie van potentieel niet-repliceerbare studies de kosten minimaliseert en de efficiëntie van de kennisverwerving voor de wetenschappelijke gemeenschap in het algemeen maximaliseert. Deze contra-intuïtieve conclusie was het onvermijdelijke gevolg van één cruciale veronderstelling, namelijk dat de wetenschappelijke gemeenschap niet alle gepubliceerde resultaten even belangrijk vindt. Wetenschappers publiceren hun bevindingen omdat ze die interessant en opwindend vinden, en ze publiceren in de hoop dat de wetenschappelijke gemeenschap die opwindende zal delen en op de gepubliceerde bevindingen zal voortbouwen met interessanter onderzoek. In werkelijkheid worden die hoop en verwachtingen zelden vervuld: de meeste gepubliceerde artikelen verdwijnen spoorloos en slechts enkele bevindingen slaan aan en iedereen begint erover te praten – op conferenties, op Twitter, of door de bevindingen in de literatuur aan te halen.

Zodra men zich bewust wordt van deze scheve verdeling van de belangstelling in de gemeenschap, levert de vraag wie bevindingen moet repliceren, en wanneer,

contra-intuïtieve antwoorden op. Het model van Lewandowsky en Oberauer vergeleek twee replicatieregimes: in het "particuliere" regime werden alle bevindingen vóór publicatie door de auteur gerepliceerd om te voorkomen dat de replicatie achteraf zou mislukken. Bevindingen die niet werden gerepliceerd, werden nooit gepubliceerd. In het alternatieve "openbare" regime werden alle belangrijke bevindingen gepubliceerd, ook al was de repliceerbaarheid ervan onzeker. Pogingen om die gepubliceerde bevindingen te repliceren bleven beperkt tot de (weinige) verschijnselen die de wetenschappelijke gemeenschap interessant vond. Uit de vergelijking van deze twee regimes kwam één algemene bevinding naar voren: het particuliere regime maakte aanzienlijk meer kosten, in termen van het totale aantal uitgevoerde experimenten, dan het openbare regime om dezelfde hoeveelheid reproduceerbare nieuwe kennis aan het licht te brengen. De hogere kosten ontstonden doordat veel replicaties vóór publicatie bevindingen betroffen waaraan andere wetenschappers uiteindelijk geen belang hechtten. En elke replicatie die voor niets wordt uitgevoerd, betekent verspilde moeite die aan een productiever doel had kunnen worden besteed. Het lage aantal replicaties van gepubliceerde bevindingen kan derhalve gunstig zijn voor de wetenschappelijke productiviteit in het algemeen, omdat minder moeite wordt verspild aan irrelevante replicaties.

Het tweede voorbeeld betreft de wisselwerking tussen gegevens en theorie. Als een bevinding met succes door een theorie is voorspeld, moet die dan worden gerepliceerd om ons vertrouwen in de theorie te vergroten, of kunnen we beter een nieuwe voorspelling door de theorie testen? Het antwoord is opnieuw genuanceerd en hangt af van hoe strak de voorspellingen gekoppeld zijn aan een theorie (Oberauer & Lewandowsky, 2019). Wanneer theorieën sterk beperkt en zeer gedetailleerd gespecificeerd zijn (bv. in een computersimulatie), dan blijkt dat replicatie van een succesvol voorspelde bevinding minder informatief is dan een test van een nieuwe voorspelling. Met de wiskunde van de Bayesiaanse inferentie toonden Oberauer en Lewandowsky aan dat ons vertrouwen in een theorie meer zou moeten toenemen na een succesvolle test van een tweede, andere voorspelling dan na replicatie van de eerste succesvolle test. Het omgekeerde gold voor minder specifieke theorieën waarvan de voorspellingen niet nauw gekoppeld waren aan de axioma's van de theorie – zoals het geval is bij de meeste verbaal geformuleerde theorieën. Voor het ontdekkingsgerichte onderzoek dat typisch is voor die theorieën, vergroten replicaties ons vertrouwen meer dan tests van nieuwe voorspellingen.

Deze twee voorbeelden illustreren dat er geen eenvoudig één-op-één verband bestaat tussen reproduceerbaarheid en robuuste wetenschap enerzijds en replicatie anderzijds. Het oplossen van de replicatiecrisis vereist een cultuuromslag en vereist dat goed gerichte replicaties mainstream worden. Maar een robotachtige replicatie voor en na publicatie is niet de beste oplossing.

Het volstaat niet een onderscheid te maken tussen enerzijds het op een computersysteem gebaseerde idee om elk onderdeel van een studie, inclusief de resultaten, te reproduceren, en anderzijds de nadruk op het reproduceren van resultaten, waarbij variaties in methoden worden toegestaan, een benadering die populairder is binnen de experimentele wetenschappen. Over de onderzoeksdomeinen heen, en vaak zelfs daarbinnen, worden verschillende veronderstellingen gemaakt over: of en in hoeverre onderzoekers de onderzoeksomstandigheden kunnen controleren; wat variatie kan aangeven over de onderzochte verschijnselen, waarbij het kan gaan om een fout of een poging tot fraude, maar ook om een nieuwe ontdekking of een artefact van de onderzoekopzet; welke methoden het meest geschikt zijn om kennis uit gegevens af te leiden, en met name de mate waarin het onderzoek afhankelijk is van statistische en berekeningsinstrumenten (die soms eenvoudigweg niet geschikt zijn voor de beoogde doelen, zoals in het geval van de interpretatie van kwalitatieve gegevens); welke onderzoeksdoelen het belangrijkste zijn en in hoeverre die kunnen worden gespecificeerd voordat het onderzoek daadwerkelijk wordt uitgevoerd; wat een robuuste bewijsbasis is, en of er vaste criteria zijn om die beoordeling te maken; en of het subjectieve oordeel van onderzoekers al dan niet betrouwbaar kan zijn, met name in situaties waarin het onderzoek niet in hoge mate kan worden gestandaardiseerd (zie tabel 1).

Tabel 1. Factoren die van invloed zijn op de opvattingen van onderzoekers over reproduceerbaarheid

Veronderstelde mate van controle over de onderzoeksvoorwaarden
Inzicht in variatie
Afhankelijkheid van statistieken en berekeningen als afleidingsinstrumenten
Precisering van de onderzoeksdoelstellingen
Stabiliteit van achtergrondkennis en feitenmateriaal
Afhankelijkheid van het deskundigenoordeel van de onderzoekers

Gezien deze diversiteit stellen we een taxonomie van zes vormen van reproduceerbaarheid voor die nauw aansluit bij de betekenis die reproduceerbaarheid onder verschillende onderzoeksomstandigheden krijgt (zie ook Leonelli, 2018).

Zes vormen van reproduceerbaarheid

1. Reproduceerbaarheid van berekeningen

Bij computeronderzoek en digitale experimenten is de reproduceerbaarheid van berekeningen meestal gericht op het opnieuw doorlopen van een bepaalde reeks

algoritmen met dezelfde gegevens. Discrepanties tussen de resultaten worden opgevat als een teken dat er problemen zijn in het computerapparaat dat wordt gebruikt om de gegevens te analyseren, en worden dus gebruikt om fouten en bugs in de gegevensanalyse en de programmering op te sporen en op te lossen. Een belangrijke voorwaarde voor deze benadering is de beschikbaarheid van herbruikbare code en gegevens, waardoor de analyse kan worden herhaald.

2. Directe experimentele reproduceerbaarheid

Directe experimentele reproduceerbaarheid wordt meestal aangetroffen in sterk gecontroleerde experimenten die erop gericht zijn een welomschreven hypothese te toetsen, zoals bijvoorbeeld klinische proeven, en bestaat uit het vermogen om dezelfde resultaten te verkrijgen door de herhaalde toepassing van dezelfde onderzoeksmethoden/-processen. Een belangrijke voorwaarde voor deze benadering is de aanwezigheid en publieke beschikbaarheid van gestandaardiseerde protocollen, materialen en omgevingen, die zorgvuldig zijn ontwikkeld en gekalibreerd om met grote precisie zeer specifieke resultaten op te leveren, en waardoor de juiste mate van experimentele controle kan worden bereikt.

3. Indirecte reproduceerbaarheid

Wanneer wordt erkend dat aanzienlijke delen van een proefopstelling aan de controle van de experimentatoren ontsnappen (een situatie die we zullen aanduiden als "semi-gestandaardiseerde experimenten"), kunnen onderzoekers een beroep doen op indirecte reproduceerbaarheid, d.w.z. nagaan of de resultaten van verschillende experimenten voldoende op elkaar lijken. Het vergelijken van gegevens over hetzelfde verschijnsel en/of dezelfde hypothese uit verschillende experimentele opstellingen wordt vaak beschouwd als een nuttig valideringsinstrument om na te gaan of de onder variabele omstandigheden bekomen resultaten al dan niet overeenkomen, waardoor de robuustheid van de betrokken bevindingen kan worden geverifieerd (Radder, 1996; Derksen & Morawski, 2022). Deze benadering van reproduceerbaarheid is typisch voor verkennend onderzoek, waarbij de experimentele omstandigheden niet in hoge mate gestandaardiseerd kunnen worden omdat de onderzoekers nog bezig zijn met het definiëren van de methoden en/of de doelstellingen van hun onderzoek. Ze kenmerkt ook studies van verschijnselen die voortdurend veranderen in reactie op hun ruimere omgeving, zoals evoluerende organismen of sociaal gedrag. Voorbeelden zijn preklinisch onderzoek (Lowe, Leonelli, & Davies, 2019), verkennende experimenten op modelorganismen (Ankeny & Leonelli, 2020) of studies van ontwikkelingsprocessen (Love, 2020; Weber, 2022), en sommige psychologische experimenten waarvan de bredere context specifiek is voor een bepaalde plaats of historische periode (Felt, 2019; Derksen & Morawski, 2022).

4. Reproduceerbaarheid van de reikwijdte

Opnieuw met betrekking tot semi-gestandaardiseerde experimenten, is een andere kijk op reproduceerbaarheid het idee van de reproduceerbaarheid van de reikwijdte, dat wil zeggen het herhalen van experimenten met als doel na te gaan of de resultaten al dan niet onder verschillende omstandigheden worden verkregen (met andere woorden, de reikwijdte van de resultaten: bijvoorbeeld of een bepaalde bevinding over het expressiepatroon van een bepaalde genetische route alleen geldt voor de specifieke muizenstam waarop het experiment voor het eerst werd uitgevoerd, of dat dezelfde patronen worden gevonden wanneer andere muizenstammen of zelfs andere organismen worden gebruikt). De benadering van de reproduceerbaarheid van de reikwijdte bestaat erin verschillen in de resultaten van herhaalde experimenten niet als bewijs van fouten te beschouwen, maar eerder als bewijs dat de bevindingen beperkt kunnen zijn tot specifieke experimentele omstandigheden (Leonelli, 2018). De onderzoekers zijn in dit geval geïnteresseerd in het onderzoeken van de verschillen tussen de omstandigheden waarin bepaalde resultaten worden verkregen en de omstandigheden waarin dat niet het geval is, wat op zich belangrijke informatie kan verschaffen over de betrokken verschijnselen.

5. Hypothetische reproduceerbaarheid

In gevallen waarin onderzoeksmethoden en -omstandigheden niet sterk gestandaardiseerd zijn, en dus moeilijk te repliceren zonder de onderzoeksopzet ingrijpend te wijzigen, wordt vaak een beroep gedaan op het idee van hypothetische of conceptuele reproduceerbaarheid. Dit is de inspanning om resultaten te verkrijgen die overeenkomen met de resultaten die voorspeld zijn als implicaties van eerdere bevindingen, waardoor indirect de betrouwbaarheid van de eerdere bevindingen wordt bevestigd (Romero, 2017). Deze benadering van reproduceerbaarheid is gebaseerd op het bestaan van robuuste redeneerketens die aan het onderzoek ten grondslag liggen: er wordt een hypothese en/of voorspelling gedaan over de implicaties van een bepaalde reeks bevindingen, en wanneer die hypothese empirisch wordt getoetst, wordt ze beschouwd als een validatie van de bevindingen waarop ze was gebaseerd.

6. Reproduceerbare deskundigheid

Vooraf wanneer kwalitatieve onderzoeksmethoden worden toegepast in een zeer vluchtige en dynamische omgeving, maakt de veranderende aard van het doel het zeer moeilijk om na te denken over het reproduceren van onderzoeksresultaten of zelfs onderzoeksmethoden, aangezien deze voortdurend moeten worden aangepast aan de veranderende onderzoekssituatie. Er wordt echter wel verwacht dat de deskundigheid die wordt ingezet om het onderzoek uit te voeren reproduceerbaar is: met andere woorden, dat elke deskundige onderzoeker die op

een bepaalde tijd en plaats met dezelfde methoden en hetzelfde type materiaal werkt, vergelijkbare resultaten zal produceren en er zo niet dezelfde gegevens, dan toch minstens dezelfde overkoepelende patronen en inzichten uit zal pikken (Leonelli, 2018). Veldwerk in de ethologie of etnografisch onderzoek in de antropologie, bijvoorbeeld, is sterk afhankelijk van het idee van reproduceerbare deskundigheid, met methodologieën die specifiek zijn ontwikkeld om het hoofd te bieden aan de onmogelijkheid van directe reproductie van bevindingen (zoals beperkte toegang en onderzoek over verschillende monsters, of het beheer van de manier waarop onderzoek wordt uitgevoerd; bijvoorbeeld de centralisatie van onderzoek op locaties waar veel onderzoekers kunnen samenwerken, elkaars werk kunnen controleren en de betrouwbaarheid ervan kunnen garanderen voor degenen die geen toegang hebben tot dezelfde instrumenten/bronnen).

Enge en ruime interpretaties van reproduceerbaarheid

De zes verschillende vormen van reproduceerbaarheid die we hier hebben voorgesteld zijn geïdealiseerd en beogen niet volledig te zijn voor alle vormen van reproduceerbaarheid in het hedendaagse onderzoek. Hun nut ligt in het feit dat ze een voorbeeld zijn van de mate van diversiteit tussen benaderingen van een goede onderzoekspraktijk, en het belang van het vinden van manieren om die diversiteit te erkennen en te ondersteunen in elk initiatief ter ondersteuning van de reproduceerbaarheid.

Wat leren we dus van het beschouwen van meerdere interpretaties van reproduceerbaarheid? Het belangrijkste is wellicht dat we erkennen dat reproduceerbaarheid belangrijk is, maar dat ze verschillende vormen aanneemt, aangepast aan verschillende onderzoekssituaties. Verschillen doen zich niet alleen voor op het niveau van de disciplines, maar kunnen veel fijnmaziger zijn, afhankelijk van de specifieke doelstellingen, methoden, materialen en onderzoeksomstandigheden van de betrokken onderzoeksgroep. Gezien deze noodzakelijke variatie moet reproduceerbaarheid worden geconceptualiseerd en toegepast op een manier die niet te eng is, zodat het onderzoek niet onbedoeld wordt beperkt op een manier die niet nuttig is, noch te ruim om zinvol te zijn.

Denk eerst aan de problemen met een te enge interpretatie van reproduceerbaarheid. Dit gebeurt soms door institutioneel onderzoeksbeleid, waarbij universiteiten zich richten op de reproduceerbaarheid van berekeningen of directe reproduceerbaarheid als de belangrijkste en soms enige manier om dit concept te begrijpen. Wanneer dit gebeurt, worden sterk gecontroleerde experimenten met vooraf gespecificeerde doelstellingen een voorbeeld van "goede praktijk" en "rigoreus onderzoek", waarbij geen recht wordt gedaan aan andere onderzoeksmethoden, zoals kwalitatieve tradities gericht op de analyse van de gesitueerdheid, en verkennend kwantitatief onderzoek (bv. datamining). En toch is een enge interpretatie van reproduceerbaarheid eenvoudigweg

ongepast in het geval van vele onderzoeksomgevingen, en met name die welke worden gekenmerkt door een gebrek aan standaardisatie (misschien omdat het bestudeerde verschijnsel uniek is of voortdurend verandert) en/of afhankelijkheid van deskundige kennis van de betrokken onderzoekers. Ter illustratie: een etnografische studie over hoe de inheemse bevolking op een Polynesisch eiland haar fysieke en sociale structuren na een verwoestende overstroming opnieuw opbouwt, kan op een later tijdstip op geen enkele zinvolle manier worden gerepliceerd.

Door dergelijke onderzoekssituaties te devalueren, verliest de enge benadering van reproduceerbaarheid de belangrijke rol die deskundigheid en belichaamde kennis spelen bij de productie, verwerking en beoordeling van gegevens, evenals het belang van de sociale context bij het definiëren van specifieke onderzoeksomstandigheden. De enge benadering heeft ook de neiging de nadruk te leggen op een valse dichotomie tussen "hermeneutische" en/of "subjectieve" onderzoeksmethoden enerzijds en "kwantitatieve" en/of "objectieve" methoden anderzijds, waarbij deze laatste doorgaans worden ondersteund. Dit onderscheid is echter niet nuttig bij de beoordeling van goede praktijken in verschillende soorten onderzoeksomgevingen, en met name in situaties van fundamenteel onderzoek waarin onderzoekers de doelstellingen, methoden, instrumenten en materialen voor hun werk gaandeweg ontwikkelen. Het minst nuttige is dat de enge interpretatie van reproduceerbaarheid niet helpt bij het maken van een onderscheid tussen onopzettelijke fouten, bedrog en fraude, en variatie als gevolg van verschillen in onderzoeksomstandigheden; evenmin helpt ze bij het maken van een onderscheid tussen oprechte pogingen om geaccepteerde "feiten" in twijfel te trekken en te verifiëren, en het kwaadwillig in twijfel trekken van wetenschappelijke resultaten op grond van gevestigde belangen (waardoor reproduceerbaarheid als "wapen" wordt gebruikt, zie hieronder).

Denk nu aan de problemen die ontstaan wanneer een te **ruime** interpretatie van reproduceerbaarheid wordt gehanteerd. Ter illustratie: er kan een te ruime interpretatie ontstaan wanneer wordt afgezien van eisen inzake replicatie ten gunste van meer flexibele "conceptuele replicaties" (d.w.z. andere, gewoonlijk onverwachte verschijningsvormen van een resultaat worden geherinterpreteerd als "conceptueel" identiek of gelijkaardig en aldus ogenschijnlijk een "replicatie" vormend). In de eerste plaats stelt een te ruime opvatting van reproduceerbaarheid onderzoek bloot aan een ongepaste kritische beoordeling en mogelijk misbruik. Terwijl het geen zin heeft een strikt begrip van directe reproduceerbaarheid toe te passen op etnografisch onderzoek, is het evenzeer problematisch het niet reproduceren van goed gespecificeerde experimentele omstandigheden te rechtvaardigen door een beroep te doen op conceptuele reproduceerbaarheid en eenvoudigweg over te stappen op een andere methode. Het is misschien niet verrassend dat er bewijs is dat interne "conceptuele replicaties" – d.w.z. conceptuele replicaties gerapporteerd door dezelfde oorspronkelijke auteurs –

latere directe replicaties door andere onderzoekers niet waarschijnlijker maken (Kunert, 2016).

Zelfs iets schijnbaar zinnigs als het verfijnen van het onderzoeksdoel tijdens een onderzoek, in reactie op wat men tijdens het onderzoek over dat doel leert, kan worden misbruikt. Dergelijke strategieën kunnen worden gebruikt om te verhullen dat een bepaalde onderzoeksbevinding te idiosyncratisch is om geloofwaardig te zijn, en creëren een hellend vlak voor de rechtvaardiging post hoc van welk resultaat dan ook. Dit kan met name het geval zijn wanneer dezelfde onderzoeksgroep betrokken is bij het uitvoeren van een studie en het uitvoeren van conceptuele replicaties, aangezien die groep een duidelijk voorbeeld heeft, en dit is vooral betoogd wanneer deze door dezelfde onderzoeksgroep wordt uitgevoerd (die een gevestigd belang heeft bij het bevestigen van de eigen resultaten), waardoor sommigen stellen dat conceptuele replicatie de directe reproduceerbaarheid niet kan en mag vervangen (Kunert, 2016).

Ten tweede wordt een ruime opvatting van reproduceerbaarheid gewoonlijk geassocieerd met begrippen als generaliseerbaarheid en transparantie van een bepaald onderzoek, en deze associatie is problematisch. Zoals besproken in het geval van de reproduceerbaarheid van de reikwijdte, is de generaliseerbaarheid van onderzoek niet hetzelfde als de kwaliteit van onderzoek: een bepaalde reeks bevindingen kan zeer betrouwbaar zijn wanneer ze wordt toegepast op een specifiek gebied, en helemaal niet nuttig wanneer ze wordt toegepast op een ander gebied. Inzichten in hoe celwanden de celsignalering filteren om de reacties van planten op pathogenen te vergemakkelijken, bijvoorbeeld, kunnen zeer betrouwbaar zijn en van groot belang voor de productie van pesticiden voor gewassen; maar ze zullen niet generaliseerbaar zijn naar de celsignalering bij dieren, en dit gebrek aan generaliseerbaarheid heeft niets te maken met hun kwaliteit als onderzoeksresultaten. Een soortgelijke verwarring kan ontstaan wanneer reproduceerbaarheid wordt gekoppeld aan het idee van transparantie. Hoewel het delen van informatie over een experiment het ongetwijfeld gemakkelijker maakt om het te reproduceren en de kwaliteit ervan onafhankelijk te verifiëren, kan te veel transparantie de verificatie soms moeilijk of zelfs onmogelijk maken (bijvoorbeeld wanneer de belangrijkste elementen van een onderzoeksmethode worden versluierd doordat er te veel niet-essentiële details worden verstrekt, waardoor het voor anderen moeilijk wordt de methode te begrijpen en opnieuw te gebruiken). Wie precies toegang moet hebben tot informatie over een bepaald experiment is ook een belangrijke kwestie voor de reproduceerbaarheid: is het relevant dat iedereen die geïnteresseerd is in het repliceren van een studie toegang heeft tot informatie over hoe dat moet, of is het aanvaardbaar om dat voorrecht te beperken tot vertrouwde groepen? Deze vraag wordt cruciaal in het geval van commercieel gevoelig onderzoek, waar reproduceerbaarheid doorgaans geen transparantie inhoudt, maar veeleer de garantie dat regelgevers de integriteit en de geldigheid van een bepaald experiment kunnen controleren ten behoeve van toekomstige consumenten.

Ten derde bestaat de neiging te denken dat door onderzoek reproduceerbaar te maken ook problemen worden opgelost in verband met de betrouwbaarheid van onderzoeksresultaten op schaal, de blootstelling van onderzoeksresultaten aan een transdisciplinaire kritische beoordeling en het gemak waarmee onderzoek in producten kan worden omgezet. Dit is echter een misverstand over wat reproduceerbaarheid voor onderzoek kan betekenen. Reproduceerbaarheid kan een nuttig vereiste zijn om de betrouwbaarheid van de resultaten onder gespecificeerde onderzoeksomstandigheden te controleren. Ze helpt niet om te beoordelen of een bepaald inzicht kan worden toegepast op een groter of ander doel (bijvoorbeeld bij pogingen om technische oplossingen, zoals een bepaalde software, op te schalen en te optimaliseren voor een veel grotere groep gebruikers dan oorspronkelijk voorgesteld, zoals een grote groep patiënten); ze houdt niet in dat een bepaald onderzoek wordt blootgesteld aan feedback van een grote verscheidenheid van belanghebbenden, hetgeen een degelijke kritische beoordeling zou opleveren; en ze garandeert niet dat de inzichten op correcte wijze zullen worden gebruikt voor de ontwikkeling van producten die op de markt zullen worden gebracht. Voor optimalisering, transdisciplinariteit en vertaling zijn aantoonbaar veel meer middelen en verschillende vormen van arbeid nodig dan alleen maar reproduceerbaarheid.

Het belang om te balanceren op de grens tussen een al te enge en een al te ruime interpretatie van reproduceerbaarheid wordt duidelijk wanneer men kijkt naar het mogelijke gebruik van pre-registratie als strategie om de reproduceerbaarheid van onderzoeksresultaten te bevorderen. Pre-registratie verwijst naar het idee dat de experimentatoren a priori, voordat de gegevensverzameling begint, belangrijke details van de methodologie en de analyse specificeren. Het vooraf registreren van de intentie beschermt tegen een aantal onbedoelde twijfelachtige onderzoekspraktijken (bv. "p-hacking" door het experiment voort te zetten totdat een gewenst resultaat wordt verkregen) en wordt genoemd als een belangrijk aspect van reproduceerbare wetenschap (Strømmland, 2019). Naar onze mening kan pre-registratie een uitstekend instrument zijn wanneer het wordt gebruikt als een manier om methodologische redeneringen en beslissingen in specifieke stadia van het onderzoeksproces vast te leggen, zodat de geschiedenis van een bepaald onderzoeksproject gemakkelijk kan worden gereconstrueerd, geëvalueerd en gereproduceerd. Er moet echter niet worden verwacht dat onderzoekers zich zullen houden aan de plannen die bij de pre-registratie van een studie zijn gemaakt: goed gemotiveerde variaties in een onderzoeksplan zijn vaak nodig naarmate het project vordert en de onderzoekers meer te weten komen over de onderwerpen van hun studie, dus als een plan niet wordt nageleefd, wijst dat niet noodzakelijk op problematisch en onbetrouwbaar onderzoek (zie kader 2).

KADER 2: Pre-registratie

DE "PRE-REGISTRATIEREVOLUTIE"

Het vermogen om resultaten te voorspellen is fundamenteel voor de wetenschap. Hoewel wetenschappers hypothesen kunnen formuleren op basis van ongeleide observaties van natuurlijke gebeurtenissen, wordt veel rigouros onderzoek gekenmerkt door het testen van hypothesen over wat er in een experiment zal gebeuren. Het testen van voorspellingen wordt algemeen beschouwd als essentieel voor het vaststellen van bewijs voor verklarende claims (Nosek, Ebersole, DeHaven, & Mellor, 2018). Tests van voorspellingen verschillen van "post hoc" verklaringen, dat wil zeggen pogingen om – soms onverwachte – bevindingen te verklaren nadat de gegevens zijn verkregen.

Er ontstaan problemen wanneer voorspelling en post hoc verklaring worden verward of door elkaar gehaald. Als bijvoorbeeld een onverwachte bevinding wordt gepresenteerd alsof ze al was voorspeld voordat de gegevens werden verkregen, kan een dergelijke post hoc redenering leiden tot een te groot vertrouwen in de interpretatie en ongegronde verwachtingen wekken over de reproduceerbaarheid van een bevinding. Dit maken van hypothesen nadat de resultaten bekend zijn (in het Engels aangeduid als "Hypothesising After the Results are Known" of "HARKing"; Kerr, 1998) wordt algemeen beschouwd als een twijfelachtige onderzoekspraktijk die heeft bijgedragen tot de replicatiecrisis.

Pre-registratie biedt een oplossing voor dit probleem omdat het vereist dat onderzoekers hun hypothesen vooraf bepalen, voordat de gegevens zijn verzameld of geanalyseerd. Pre-registraties worden online gearchiveerd en van een datumstempel voorzien en kunnen na indiening niet meer worden gewijzigd. De pre-registratie van hypothesen voorkomt onbedoelde of heimelijke HARKing omdat lezers (en beoordelaars) de interpretatie van de resultaten door de auteur kunnen toetsen aan de pre-registratie.

POSITIEVE ASPECTEN VAN PRE-REGISTRATIE

Pre-registratie biedt tal van voordelen en waarborgen. Zo beschermt een gedetailleerde voorafgaande registratie van de methode en het analyseplan niet alleen tegen HARKing, maar ook tegen verschillende andere twijfelachtige onderzoekspraktijken, zoals het elimineren van afwijkende waarnemingen om de resultaten van een analyse gunstig te beïnvloeden, of het aanpassen van een steekproefplan op basis van gedeeltelijke resultaten (bijvoorbeeld deelnemers blijven testen totdat een resultaat significant is). In het algemeen beperkt pre-registratie de "vrijheidsgraden" van onderzoekers die anders het genereren en interpreteren van gegevens ten onrechte zouden kunnen beïnvloeden (Oberauer & Lewandowsky, 2019).

Een bijzonder krachtige uitbreiding van pre-registratie zijn de zogenaamde “geregistreerde rapporten”, waarbij de methode en het analyseplan door vakgenoten worden beoordeeld en een voorlopig besluit over de publicatie wordt genomen voordat de gegevens zijn verzameld. Het resultaat is dan niet van invloed op de beslissing tot publicatie, hetgeen bescherming biedt tegen publicatiebias ten gunste van significante resultaten. Bovendien geven geregistreerde rapporten terugkoppeling aan de onderzoeker wanneer dat aantoonbaar het belangrijkste is, namelijk voordat het onderzoek wordt uitgevoerd en terwijl er nog wijzigingen in de methodologie kunnen worden aangebracht.

Het opstellen van dergelijke rapporten geeft een nuttig, expliciet beeld van de veronderstellingen die bij de start van het onderzoek zijn gemaakt, en vormt tevens een stimulans om negatieve resultaten te publiceren voor het geval het onderzoek er niet in slaagt positieve bevindingen op te leveren (aangezien de onderzoekers al een peer assessment hebben doorlopen om te bewijzen dat hun onderzoeksopzet significant en geloofwaardig was, waardoor negatieve bevindingen op zichzelf een significant resultaat worden; zie <https://journal.trialanderror.org/>).

VEEL VOORKOMENDE MISVERSTANDEN OVER PRE-REGISTRATIE

1. Pre-registratie is geen blauwdruk voor onderzoek

Er zijn verschillende misverstanden over pre-registraties geuit die moeten worden rechtgezet. Het belangrijkste is wellicht de misvatting dat een pre-registratie een onwrikbare blauwdruk voor onderzoek is die koste wat kost moet worden nageleefd en die onderzoekers ervan weerhoudt onverwachte aspecten van hun gegevens te onderzoeken. Deze bezorgdheid is misplaatst omdat pre-registraties afwijkingen van een analyseplan niet verhinderen of verbieden. In plaats daarvan bepalen pre-registraties alleen dat niet-geplande verkennende analyses in het rapport van een studie als zodanig moeten worden vermeld. Dit beperkt geenszins de creativiteit of de verkenning; het vereist alleen een duidelijk onderscheid tussen geplande en verkennende analyses.

2. Pre-registratie is geen instrument voor de beoordeling van de kwaliteit van het onderzoek

Pre-registratie beschermt tegen het gebruik van hetzelfde bewijs om zowel een hypothese te genereren als deze vervolgens te testen, wat kan gebeuren wanneer onderzoekers post hoc verklaringen creëren voor onverwachte resultaten met behulp van een theorie die in hetzelfde experiment werd getest. Deze redenering sluit noodzakelijkerwijs de vervalsing van een theorie uit, waardoor

onderzoekers worden verleid tot een te groot vertrouwen in de theorie (Rubin, 2017). Pre-registratie voorkomt het gebruik van resultaten bij het formuleren van een hypothese door een strikte tijdsvolgorde af te dwingen: de hypothese moet worden geformuleerd en geregistreerd voordat de gegevens bekend zijn.

De tijdsvolgorde is echter slechts één manier om ervoor te zorgen dat hypothesen uit de theorie worden afgeleid zonder dat ze op dezelfde gegevens steunen (omdat de gegevens nog niet bestaan). De tijdsvolgorde op zich is echter niet noodzakelijk irrelevant om ervoor te zorgen dat de resultaten geen onrechtmatige invloed hebben op de hypothesen die uit een theorie worden afgeleid zonder op dezelfde gegevens te zijn gebaseerd. Hetzelfde doel kan worden bereikt door ervoor te zorgen dat de bestaande gegevens niet bijdragen tot de formulering van een hypothese. Een beroemd voorbeeld is Einsteins algemene relativiteitstheorie, waaraan algemeen de voorspelling van de waargenomen precessie van het perihelium van Mercurius wordt toegeschreven, hoewel die waarnemingen tientallen jaren vóór de algemene relativiteitstheorie plaatsvonden. Omdat die waarnemingen geen rol speelden bij de ontwikkeling van de theorie, voorspelde de relativiteit de precessie van het perihelium van Mercurius en ondersteunt het succes van die voorspelling de theorie "net zo sterk als wanneer die feiten pas na de formulering van de theorie aan het licht waren gekomen" (Worrall, 2014, p. 55).

Evenzo, als een hypothese wordt afgeleid nadat de resultaten van een experiment bekend zijn, maar de afleiding volledig berust op ante hoc (a priori) theorie en bewijs, dan is de legitimiteit ervan identiek aan die van een hypothese die werd afgeleid voordat de gegevens bekend waren (Oberauer & Lewandowsky, 2019; Rubin & Donkin, 2022). Bovendien is pre-registratie niet vereist om HARKing te voorkomen in dergelijke omstandigheden, mits deze afdoende wordt gedocumenteerd in het onderzoeksrapport.

Er is echter één belangrijke voorwaarde voor de legitimiteit van hypothesen die worden afgeleid nadat de gegevens bekend zijn, en die heeft betrekking op de precisie en de conceptuele duidelijkheid van de theorie. In de psychologie zijn veel theoretische constructen vaag en ontbreekt het vaak aan conceptuele strengheid (Bringmann, Elmer, & Eronen, 2022). Dit kan het moeilijk maken om na te gaan of een hypothese werkelijk is afgeleid van een theorie zonder te zijn gebaseerd op de gegevens die ze beweert te voorspellen. Om dit probleem te vermijden moeten theorieën worden geformuleerd op een computationeel of kwantitatief niveau, waarbij alle bijkomende aannames worden gespecificeerd (Oberauer & Lewandowsky, 2019).

DEEL 2. Implementatie van reproduceerbaarheid binnen academische instellingen

Welke vorm reproduceerbaarheid uiteindelijk ook aanneemt, ze vereist institutionele steun en inbedding in de onderzoekscultuur om succesvol te zijn. Het uitgebreide overleg met belanghebbenden leverde ons inzichten op over de huidige praktijk en de mogelijke toekomst van het Vlaamse onderzoek.

Implementatiepraktijken

Uit ons overleg met Vlaamse belanghebbenden zijn verschillende belangrijke veranderingen naar voren gekomen in de manier waarop onderzoek wordt geïnstitutionaliseerd, beloond en gefinancierd. De meest opvallende zijn:

- Een erkenning van de rol van data stewards als integraal onderdeel van het onderzoeksecosysteem, waarbij veel universiteiten investeren in permanente posities voor data stewards om ondersteuning te bieden bij dataverzameling, -conservering en -beheer. Na bespreking van de rol met verschillende data stewards die deelnamen aan onze overlegsessie, merkten we de positieve impact op van deze professionele figuren op het academische landschap, zowel op het vlak van ondersteuning van onderzoek als op het vlak van het helpen bemiddelen bij conflicten en spanningen tussen belanghebbenden (zie punt hieronder). Tegelijkertijd stelden we vast dat deze rollen snel door de werklast werden overspoeld en dat er nog steeds een gebrek aan erkenning is voor de deskundigheid die met gegevensconservering gepaard gaat, waarbij dergelijke rollen vaak worden gedelegeerd aan een profiel van technicus in plaats van dat van onderzoeker, en het promotie-/groeitraject dienovereenkomstig wordt afgebakend.
- Een groter bewustzijn van het belang van integriteit binnen het onderzoek, met als gevolg een toename van ethische commissies en integriteitsfunctionarissen aan verschillende universiteiten. Deze ontwikkelingen zijn zeer welkom, mits ze voldoende gestroomlijnd zijn om de onderzoekers niet met een extra werklast op te zadelen. We hebben echter soortgelijke problemen ontdekt als die welke door data stewards worden ervaren met betrekking tot de erkenning en mogelijke onderwaardering van dergelijke rollen binnen instellingen.
- Nieuwe opleidingsprogramma's voor onderzoekers en professioneel ondersteunend personeel om te leren hoe zij hun werk methodologisch meer verantwoord en robuuster kunnen maken. Deze werden zowel door onderzoekers en instellingen als door particuliere medewerkers toegejuicht, hoewel ook werd opgemerkt dat opleidingsprogramma's een beperkte maatregel vormen omdat ze de last van de naleving van de reproduceerbaarheidsnormen volledig op de schouders van de onderzoekers leggen en niet op die van de

institutionele infrastructuur. Opleiding moet daarom worden aangevuld met een werklastverdeling waarbij rekening wordt gehouden met de tijd en moeite die onderzoekers en ondersteunend personeel moeten besteden aan het verkrijgen en uitvoeren van een dergelijke opleiding, evenals met de alternatieve resultaten die uit dergelijke inspanningen kunnen voortvloeien (zoals datasets, infrastructuur en andere middelen die met de bredere gemeenschap worden gedeeld en een grote impact hebben, maar die door academische werkgevers en financiers doorgaans niet als een teken van excellentie worden beschouwd).

- De bereidheid van financierende instellingen zoals het FWO om rekening te houden met integriteit en Open Science-praktijken, waarbij negatieve resultaten momenteel worden erkend als een legitieme output voor onderzoeksevaluaties, en overwegingen rond onderzoeksintegriteit (zoals de vraag of het betrokken onderzoek op verantwoorde wijze en in samenwerking met de juiste belanghebbenden wordt uitgevoerd) nu een zogenaamde "tweede as" vormen voor onderzoeksevaluatie. Dit zijn zeer bemoedigende tekenen, aangezien financiers een enorme macht hebben om het systeem van stimulansen te veranderen door onderzoekers en hun instellingen ertoe aan te zetten meer aandacht aan deze kwesties te besteden. Tegelijkertijd bleek uit discussies onder belanghebbenden dat onderzoekers zich ongemakkelijk voelen bij het etiket "tweede as" voor onderzoek: er werd opgemerkt dat een dergelijk label kan worden geïnterpreteerd als een bevestiging dat kwesties van integriteit, kwaliteit en reproduceerbaarheid geen prioriteit hebben bij de beoordeling van onderzoek en altijd een ondergeschikte rol zullen spelen ten opzichte van het aantal en de plaats van de publicaties die zijn gepland of uit een bepaald project voortvloeien. Het FWO was zich terdege bewust van deze problemen en bezorgdheden en wees op de moeilijkheden om dergelijke nieuwe beoordelingscriteria toe te passen in de Vlaamse context, waar een groot deel van de beoordelingen wordt uitgevoerd door internationale deskundigenpanels die zeer uiteenlopende perspectieven hebben op wat als waardevol moet worden beschouwd in onderzoek. Daarbij wordt in het bijzonder gewezen op de moeilijkheden bij de beoordeling van transdisciplinair onderzoek, dat veel instellingen in theorie van het grootste belang achten om mondiale uitdagingen aan te gaan, maar dat in de praktijk zeer moeilijk te beoordelen blijft aan de hand van de traditionele onderzoeksmaatstaven en op één discipline gerichte beoordelingspanels.
- De erkenning van Open Science als een belangrijk fenomeen en pogingen om een cultuurverandering teweeg te brengen binnen academische instellingen. Dergelijke pogingen kunnen concrete vormen aannemen, zoals de wijziging van de criteria voor de aanwerving of promotie van academici zodat ze inspanningen omvatten op het gebied van dataconservering, reproductie van andere studies en kwaliteitsbeoordeling. Sommige universiteiten proberen

dergelijke overwegingen op professoraal niveau in te voeren, bijvoorbeeld door de invoering van beschrijvende cv's en eigen verklaringen over de naleving van de Open Science-regels, maar het is nog niet duidelijk of het invoeren van dergelijke aanvullende criteria een verschil maakt voor wie wordt aangenomen als of bevorderd tot professor. Deze initiatieven werden door de onderzoekers verwelkomd als uiterst belangrijk om het academische kredietsysteem zodanig te veranderen dat er beter rekening wordt gehouden met goede onderzoekspraktijken. Een andere, minder veeleisende verandering bestaat in het erkennen van goed wetenschappelijk werk door middel van eenmalige onderscheidingen, zoals Open Science-prijzen en "badge"-systemen voor de erkenning van OS-activiteiten. Deze onderscheidingen werden met meer ambivalentie behandeld door de onderzoekers, die vreesden dat dit misschien gemakkelijke vormen van erkenning zijn die geen enkel verschil maken zolang ze niet in aanmerking worden genomen door aanwervings- en promotiecomités en door financieringsinstanties bij de toekenning van subsidies. Deze vormen van erkenning lopen ook het risico gemakkelijk te worden "gemanipuleerd", hetgeen aanleiding geeft tot "open-wassen" (d.w.z. het cosmetisch opsmukken van onderzoekspraktijken als zijnde in overeenstemming met Open Science-regels, terwijl er in feite weinig is veranderd en onderzoekers gewoon hebben geleerd te "lijken" alsof ze aan de eisen voldoen; bijvoorbeeld wanneer een onderzoekster haar gegevens "open" maakt door ze in een database op te laden, maar niet echt de moeite neemt om die gegevens te formatteren voor hergebruik door anderen, geen passende metadata verstrekt en/of alleen gegevens oplaadt waarvan ze denkt dat ze niet erg belangrijk zijn).

Naast institutionele initiatieven vonden we een overvloed aan bottom-up activiteiten en opleidingsprogramma's die door onderzoekers in het veld werden opgezet. Bijzondere vermelding verdienen hier initiatieven van beginnende onderzoekers, zoals ReproducibiliTea.¹ ReproducibiliTea begon als een kleine journal club aan de Universiteit van Oxford die gewijd was aan het bespreken van papers met betrekking tot reproduceerbaarheid. Het is nu een internationaal netwerk van journal clubs verspreid over meer dan 100 instellingen in 25 verschillende landen. In Vlaanderen is de KU Leuven² een belangrijk knooppunt van het netwerk. Het doel van ReproducibiliTea is het creëren van een veilige haven voor interdisciplinaire discussie over wetenschap in het algemeen, en meer specifiek over Open Science en reproduceerbaarheid. De deelnemers lezen vóór elke bijeenkomst relevante artikelen en bespreken die in kleine groepen, met een plenaire inleiding en een afsluitende reflectie. De bijeenkomsten zijn informeel en interdisciplinair, en worden gecoördineerd door promovendi en postdoctorale onderzoekers, die hun

¹ <https://reproducibiliTea.org/>

² <https://kuleuvenopenscienceday.pubpub.org/pub:1479tyu8/release/1>

tijd aan deze initiatieven besteden omdat ze menen dat deze een waardevolle bijdrage leveren aan hun eigen opleiding en aan de onderzoeksexcellentie van hun gastinstelling.

Deze bottom-up initiatieven hebben een echt verschil gemaakt door opleiding en mogelijkheden voor discussie tussen onderzoekers aan te bieden, evenals een forum om zo nodig externe deskundigheid in te schakelen. Aan de KU Leuven zijn er contacten met de overheid via de taskforce Open Science van de universiteit. Er dient opgemerkt te worden dat dergelijke initiatieven niet altijd goed gesteund en erkend lijken te worden door de instellingen die ze organiseren, wat jammer is (zie aanbevelingen). Het blijft een open vraag of de promovendi en postdoctorale onderzoekers die zich voor deze initiatieven inzetten, hun werk voor Open Science beloofd zullen zien in hun toekomstige carrière.

Spanningen tussen belanghebbenden

De roep om reproduceerbaarheid heeft geleid tot allerlei spanningen tussen verschillende belanghebbenden en verschillende gemeenschappen, die moeten worden onderzocht en opgelost. Misschien de belangrijkste hiervan is de spanning tussen enerzijds door de overheid gefinancierde onderzoekers, voornamelijk aan universiteiten, die geacht worden volledig transparant te zijn, en anderzijds onderzoekers die in het bedrijfsleven werken en die door regelgevende instanties worden gecontroleerd, maar niet door academische vakgenoten, in het bijzonder wat betreft hun stroomopwaartse onderzoeksinspanningen.

Er zijn bijkomende spanningen vastgesteld tussen onderzoekers met verschillende niveaus van anciënniteit, voornamelijk omdat de extra kosten en inspanningen die nodig zijn voor het uitvoeren van onderzoek volgens een hoger criterium van reproduceerbaarheid worden gedragen door junioronderzoekers, die waarschijnlijk meer relevante vaardigheden hebben verworven. Dataconservatie wordt bijvoorbeeld steeds meer erkend als een hooggekwalificeerde taak, die erkenning verdient in de vorm van co-auteurschap van papers en een academische status (Europese Commissie 2018, Open Science Policy Platform 2020). De huidige hiërarchie van onderzoekers, waarbij professoren die ideeën hebben en zoeken naar financiering bovenaan staan en postdoctorale onderzoekers die experimenten uitvoeren en de gegevens verwerken onderaan staan, weerspiegelt niet de expertise en senioriteit die nodig zijn voor een effectieve afstemming op het complexe internationale data-ecosysteem (Leonelli et al., 2016; Leonelli et al., 2017).

In dit verband zijn er aanzienlijke spanningen tussen werk dat als "academisch" wordt beschouwd en werk dat als "technisch" of als "ondersteuning" van onderzoek wordt beschouwd. De rol van professioneel ondersteunend personeel blijft vaak onduidelijk. Vaak worden degenen die inspanningen op het gebied van reproduceerbaarheid ondersteunen en zelf geen academische onderzoekers

zijn, ondergewaardeerd, ondanks het hoge niveau van de vereiste vaardigheden. Het is bijvoorbeeld geweldig om de opkomst te zien van data stewards als een aparte en breed erkende professionele rol, en om te zien dat de technische en professionele vaardigheden die nodig zijn voor zo'n veeleisende rol worden erkend. De relatie van dergelijke professionals tot het onderzoek (wordt bijvoorbeeld verwacht dat ze auteurs zijn?) blijft echter onopgelost, en de promotie-/groeimogelijkheden voor dergelijk gespecialiseerd personeel zijn vaak te weinig gespecificeerd of beperkt. Nog minder duidelijk is de relatie tussen personeel dat ethische en communicatiestrategieën ondersteunt en inspanningen gericht op reproduceerbaarheid. Dit personeel toonde zich tijdens de interviews duidelijk bewust van de sleutelrol die het naast de onderzoekers kan en moet spelen bij de inspanningen op het gebied van reproduceerbaarheid; maar dit besef kwam niet tot uiting in de institutionele regelingen en vaak ook niet in de interviews met onderzoekers die dergelijke steun als bijkomstig aan "echt" onderzoek beschouwden. Deze barrières helpen niet.

Uit onze raadplegingen kwam een belangrijk ervaren probleem bij het streven naar reproduceerbaarheid naar voren, namelijk dat de institutionele belonings- en stimuleringsmechanismen achterblijven en nog steeds draaien om conventionele maatstaven (d.w.z. de hoeveelheid publicaties en citaties zonder directe weergave van reproduceerbaarheidsindices; zie de kritiek in Europese Commissie 2018). Dit leidt tot een nadruk op snel behaalbare resultaten op korte termijn, met weinig erkenning voor de langere aanlooptijd die nodig is om een laboratorium- en onderzoekscultuur op te bouwen die gericht is op reproduceerbaarheid. We merkten dat zelfs wanneer instellingen de prestatiecijfers aanvullen met een kwalitatieve evaluatie, wat een mogelijke weg biedt naar de beloning van inspanningen op het gebied van reproduceerbaarheid, de nadruk blijft liggen op nieuwigheid boven replicatie of kwaliteit. Hoewel de financiers zich bewust zijn van het probleem, vinden ze het moeilijk om tijdens de beoordeling het accent te verleggen. Het probleem is bijzonder acuut voor de onderzoekscentra die aangaven dat ze reproduceerbaarheid weliswaar zeer serieus nemen, maar het vanwege hun directe band met het bedrijfsleven moeilijk vinden om de niveaus van transparantie te evenaren die beschikbaar zijn voor universitaire onderzoekers. Ze zijn meer gericht op produceren. Een activiteit die cruciaal is voor het bereiken van reproduceerbaarheid, maar die grotendeels onzichtbaar blijft, is het reviewen van ingediende tijdschriftartikelen of subsidieaanvragen. Reviewen of beoordelen is een arbeidsintensieve activiteit, maar is geheel vrijwillig en bijna volledig onbeloond.

Last but not least vormden vertrouwenskwesaties een belangrijk en terugkerend thema dat in de meeste discussies met belanghebbenden aanwezig was. Hoe kan een onderzoekscultuur worden opgebouwd die berust op een open en eerlijke discussie, wanneer iedereen elkaar controleert op tekenen van "kwade trouw" of niet-naleving van de reproduceerbaarheidsprocedures? Op deze bredere kwestie wordt in het volgende hoofdstuk ingegaan.

DEEL 3. De spanning tussen veerkracht en reproduceerbaarheid

Wetenschap wordt zelden in een vacuüm bedreven en de interactie tussen het publiek en de wetenschappelijke gemeenschap kan vele vormen aannemen (zie kader 3). Wanneer wetenschappers in onze Melkweg een planeet ontdekken die van diamanten is gemaakt (Bailes et al. 2011), zijn publieke fascinatie en applaus vrijwel verzekerd. Wanneer wetenschappers echter ontdekken dat de grootschalige verbranding van fossiele brandstoffen klimaatverandering veroorzaakt, of dat een dodelijk virus in de lucht het dragen van een masker en social distancing vereist om onder controle te worden gehouden, dan kan de publieke en politieke reactie minder gunstig zijn. In veel gevallen kan publieke of politieke afkeuring zich vertalen in aanvallen op individuele wetenschappers of groepen wetenschappers die het wetenschappelijke proces of het resultaat ervan in gevaar kunnen brengen (Bryse, Oreskes, O'Reilly, & Oppenheimer, 2013; Freudenburg & Muselli, 2010; Lewandowsky, Oreskes, Risbey, Newell, & Smithson, 2015). Ter illustratie: er is overweldigend bewijs dat de wetenschap van de klimaatverandering onderhevig is geweest aan een goed gefinancierde en goed georganiseerde oppositie die niet alleen met succes beleidsmaatregelen heeft vertraagd, maar ook de perceptie van het publiek van de kracht van de wetenschappelijke consensus heeft ondermijnd (Brulle, 2013, 2018; Brulle, Hall, Loy, & Schell-Smith, 2021; Dunlap & McCright, 2011; Lewandowsky, 2021; Oreskes & Conway, 2010).

KADER 3. Wetenschapscommunicatie

Het publiek ontvangt zelden wetenschappelijke informatie rechtstreeks van wetenschappelijke bronnen. In plaats daarvan zijn gewoonlijk verschillende communicatoren betrokken bij de vertaling van wetenschappelijke bevindingen en resultaten in artikels in de media of andere voor het publiek toegankelijke bronnen. Op zijn best biedt wetenschapscommunicatie het publiek de mogelijkheid om inzicht te krijgen in en bij te dragen tot het proces van wetenschappelijke kennisproductie. Lokale organisaties en niet-academische groepen en instellingen (bv. musea) hebben een belangrijke rol te spelen wanneer het erom gaat onderzoek op een toegankelijke manier te communiceren, waarbij de nadruk wordt gelegd op de redenen voor het vertrouwen in specifieke wetenschappelijke resultaten en deelname aan onderzoek wordt bevorderd bij degenen die bereid zijn meer betrokken te raken.

Wanneer wetenschapscommunicatie wordt uitgebreid om de gemeenschap bij het wetenschappelijke proces zelf te betrekken, kan het resultaat zeer productief zijn en voorzien in behoeften van de gemeenschap die anders misschien onvold zouden blijven. Een bekend voorbeeld is de stad Pickering in Yorkshire, die begin jaren 2000 herhaaldelijk te kampen had met verwoestende overstromingen.

De eerste oplossingen die door de regering werden voorgesteld, stuitten op hevig verzet van de plaatselijke gemeenschap, die er uiteindelijk in slaagde rechtstreeks betrokken te raken bij het uitwerken van alternatieve oplossingen die op brede steun van de gemeenschap konden rekenen en de historische stad beschermden (Garvey & Paavola, 2021; Whatmore & Landström, 2011).

Niet alle wetenschapscommunicatie is echter succesvol, en er zijn verschillende veel voorkomende valkuilen. Misschien de meest voorkomende misvatting is die van het zogenaamde "tekortmodel", d.w.z. het idee dat het publiek kennis mist die wetenschapsvoorlichters kunnen verschaffen, zodat het informatietekort bij het publiek wordt verholpen (Hornsey, 2020). Hoewel voorlichting een positief verschil kan maken (en uiteindelijk ook moet maken; anders zou alle onderwijs gedoemd zijn te mislukken), kan ze aanleiding geven tot verschillende problemen. Ten eerste bevordert ze een simplistische kijk op wetenschap als een bron van ultieme waarheid, waardoor wetenschappelijke vooruitgang (waaronder veranderingen van richting of inzicht en fouten) aangevoeld wordt als verraad van vertrouwen. De kracht van de wetenschap ligt echter eerder in haar feilbaarheid en in haar vermogen om niets als vanzelfsprekend aan te nemen: om het publiek in staat te stellen de waarde van wetenschappelijke beweringen te beoordelen, is het noodzakelijk om een begrip van wetenschap als een zelfcorrigerend proces van onderzoek te ondersteunen, in plaats van als een reeks onfeilbare waarheden (Oreskes, 2019). Ten tweede biedt de opvatting van wetenschappelijke communicatie als "informatieverstrekking" geen mogelijkheid voor niet-onderzoekers om een dialoog aan te gaan met het wetenschappelijk onderzoek, bijvoorbeeld door hun eigen expertise in te brengen (zoals vaak het geval is bij citizen science-projecten). Dit gebrek aan mogelijkheden voor betrokkenheid kan ernstige gevolgen hebben, zoals is gebleken uit de wetenschappelijke reacties op de COVID-19-pandemie, waarbij de weinige projecten waarbij burgers als deelnemers werden betrokken, veel betrouwbaardere resultaten opleverden dan projecten waarbij burgerparticipatie van meet af aan werd uitgesloten (Leonelli, 2021). Ten derde biedt de opvatting van wetenschappelijke communicatie als "informatieverstrekking" geen ruimte voor een dialoog tussen onderzoekers en publiek, waardoor elk conflict een onoverkomelijk meningsverschil lijkt. Heel vaak komen de obstakels rond de communicatie over wetenschap namelijk voort uit conflicten tussen de diepgewortelde wereldbeelden van mensen en het wetenschappelijk bewijs (Lewandowsky & Oberauer, 2016). Mensen die roken kunnen gemotiveerd zijn om bewijzen over de negatieve gezondheidseffecten van tabak te verwerpen. Mensen die een "laissez faire"-benadering van de vrijemarkteconomie onderschrijven, zijn gemotiveerd om het bewijs voor klimaatverandering en de inmenging in de markten die vereist is om deze te beperken, af te wijzen. Wetenschapscommunicatie is moeilijk in dergelijke omstandigheden, en vereist een zorgvuldige analyse en geschikte boodschappers en boodschappen. (Zie Lewandowsky, 2021, voor een overzicht van communicatie-instrumenten.)

Bij elke beschouwing over reproduceerbaarheid moet daarom ook rekening worden gehouden met de weerbaarheid van de wetenschappelijke onderneming tegen politiek gemotiveerde inmenging of andere soorten gevestigde belangen. Het blijkt dat de relatie tussen transparantie en reproduceerbaarheid enerzijds en veerkracht anderzijds gecompliceerd, genuanceerd en soms contra-intuïtief is. Zo zijn verzoeken om vrijheid van informatie onder het mom van transparantie, meestal gericht op e-mail correspondentie tussen wetenschappers, een veelgebruikt instrument geweest van politieke operatoren die het vertrouwen van het publiek in de klimaatwetenschap wilden ondermijnen (Lewandowsky & Bishop, 2016; Lewandowsky, Mann, Brown, & Friedman, 2016).

Evenzo zijn er in de VS veel politieke inspanningen geleverd om de "transparantie" van wetenschappelijke instellingen te vergroten door middel van wetgeving of regelgeving. Die inspanningen zijn vanuit de sector van de volksgezondheid gekarakteriseerd als een "paard van Troje" (Levy & Johns, 2016) omdat ze gevestigde belangen in staat stellen beleidsinitiatieven te vertragen door gegevens in twijfel te trekken waarop federale agentschappen zich baseren. De campagne voor "gezonde wetenschap" in de jaren negentig bijvoorbeeld, die erop gericht was wetten in te voeren over de toegang tot en de kwaliteit van gegevens, werd op touw gezet door Philip Morris in een poging het verband tussen passief roken en longkanker aan te vechten (Baba, Cook, McGarity, & Bero, 2005). De daaruit voortvloeiende Data Access Act en Data Quality Act hebben aantoonbaar vertraging opgeleverd voor regelgeving op het gebied van gezondheid, doordat belangengroepen uit het bedrijfsleven de onderliggende gegevens kunnen aanvechten. De overgrote meerderheid van de publieke betwistingen van gegevens was afkomstig van gereguleerde sectoren, lobbyisten en handelsorganisaties (Levy & Johns, 2016).

De gevolgen van de toegang van gevestigde belangen tot gegevens kunnen worden geïllustreerd aan de hand van het onderzoek naar het verband tussen roken en coronaire hartziekten. De tabaksindustrie begon via haar "Council for Tobacco Research" een longitudinaal onderzoek naar coronaire hartziekten te financieren nadat de federale financiering 22 jaar na de start van het project was stopgezet. De tabaksindustrie zorgde voor financiering met als voornaamste doel volledige controle te krijgen over de gegevens (inclusief de periode van federale financiering) voor heranalyses (Cataldo, Bero, & Malone, 2010). Het is wellicht niet verrassend dat deze heranalyses het verhoogde risico op coronaire hartziekten bij rokers toeschreven aan "constitutionele factoren", zoals etniciteit.

Verdere wetgevende en regelgevende inspanningen onder de vlag van transparantie en reproduceerbaarheid zijn bijna tot op heden voortgezet, met als hoogtepunt het voorstel van het Amerikaanse milieuagentschap EPA, onder de regering Trump, om alle studies waarvan de ruwe gegevens op individueel niveau niet openbaar beschikbaar zijn, uit te sluiten van overweging bij het ontwerpen

van milieunormen of -regelgeving (Schwartz, 2018). Deze eis van transparantie is problematisch omdat veel van de gegevens over milieuverontreinigende stoffen betrekking hebben op grote cohortstudies die gevoelige medische informatie bevatten en gegevens over de blootstelling aan verontreinigende stoffen op basis van de woonplaats van de deelnemers. Bijgevolg worden die gegevens gemakkelijk gedeanonimiseerd en kan de privacy van de deelnemers niet worden gewaarborgd als de gegevens openbaar worden gemaakt (Schwartz, 2018). Omdat andere wetgeving en ethische codes het behoud van de privacy voorschrijven, zou de voorgestelde transparantie-eis een groot deel van de epidemiologische gegevens in feite uitsluiten van beschouwing door het EPA. De voorgestelde regel werd op procedurele gronden door de rechtbank nietig verklaard en zal waarschijnlijk niet opnieuw worden ingevoerd door de regering Biden (Lash21). Een soortgelijke wetgevende inspanning, de Secret Science Reform Act (SSRA), die netjes parallel liep met de voorgestelde EPA-regeling, werd gesponsord door een lid van het Amerikaanse Huis van Afgevaardigden (Lamar Smith, R-TX) met een lange geschiedenis van verzet tegen klimaatwetenschap en nauwe banden met de industrie van de fossiele brandstoffen (Levy & Johns, 2016).

Kortom, sommige politieke inspanningen die op het eerste gezicht lovenswaardig lijken – wie wil er nu geen grotere transparantie en reproduceerbaarheid? – blijken bij nader inzien problematisch te zijn en kunnen het tegenovergestelde bereiken van wat ze lijken te beogen. Er zijn onvermijdbare spanningen tussen transparantie en andere belangrijke bezorgdheden, zoals privacy, en er is geen one-size-fits-all oplossing voor die spanning die noodzakelijkerwijs de wetenschap ten goede komt (Leonelli, 2023). In kader 4 wordt een specifiek geval besproken om het dilemma te illustreren.

De nadruk op reproduceerbaarheid leidt ook tot meerdere complexe spanningen met inspanningen om de diversiteit en inclusiviteit van wetenschappelijk onderzoek te vergroten. Hoewel dit vaak aan de aandacht ontsnapt, berust reproduceerbaarheid op diversiteit en inclusie. Niet-hispanic witte mensen van Europese afkomst maken bijvoorbeeld slechts 61 % van de Amerikaanse bevolking uit, maar ze vertegenwoordigen meer dan 90 % van de deelnemers aan klinische proeven (Ma, Gutiérrez, Frausto, & Al-Delaimy, 2021; Mak, Law, Alvidrez, & Pérez-Stable, 2007). Door dit gebrek aan inclusiviteit is het mogelijk dat veel medische kennis niet reproduceerbaar is in alle sectoren van de bevolking. Ter bevestiging: van de bijwerkingen van 5-fluorouracil, een veelgebruikt geneesmiddel tegen kanker, is nu bekend dat ze in hogere mate voorkomen bij ondervertegenwoordigde bevolkingsgroepen. Omdat bij de oorspronkelijke klinische proeven voornamelijk witte deelnemers betrokken waren, werden deze bijwerkingen bij minderheidsgroepen aanvankelijk over het hoofd gezien (Yates, Byrne, Donahue, McCarty, & Mathews, 2020). Om volledig reproduceerbaar te zijn, moet de wetenschap dus streven naar inclusiviteit in de steekproeven van deelnemers aan klinische proeven.

De noodzaak van inclusiviteit in steekproeven van deelnemers leidt echter tot andere spanningen met de reproduceerbaarheid. Etnische minderheden zijn historisch gediscrimineerd in de gezondheidszorg en worden tot op heden subtiel (of niet zo subtiel) gediscrimineerd. Afro-Amerikanen worden bijvoorbeeld systematisch onderbehandeld voor pijn in vergelijking met witte patiënten, ten minste gedeeltelijk gebaseerd op wijdverspreide maar ongegronde overtuigingen over biologische verschillen tussen rassen (Hoffman, Trawalter, Axt, & Oliver, 2016). In het licht van een geschiedenis van mishandeling van inheemse bevolkingsgroepen door westerse landen (Lowes & Montero, 2021) en het voortdurende hedendaagse misbruik van medische instellingen in ontwikkelingslanden (bv. door de CIA in haar jacht op Osama bin Laden; Reardon, 2011), is het niet verwonderlijk dat minderheden terughoudend zouden zijn om deel te nemen aan onderzoek door dezelfde instellingen die hen in het verleden schade hebben berokkend of oneerlijk hebben behandeld. Pas in 2021 publiceerde de American Psychological Association (APA) een "Verontschuldiging aan mensen van kleur voor de rol van de APA in het bevorderen, in stand houden en het niet in vraag stellen van racisme, rassendiscriminatie en menselijke hiërarchie in de VS"³. Tegen deze achtergrond is het algemene gebruik van "open gegevens" bij onderzoek dat gevoelige informatie omvat, problematisch en in strijd met de doelstellingen van diversiteit en inclusiviteit. Zwarte of hispanic deelnemers kunnen bereid zijn om deel te nemen aan een experiment over werkgeheugen of aandacht, maar ze kunnen minder geneigd zijn om deel te nemen als ze weten dat hun gegevens geheranalyseerd zouden kunnen worden om verbanden tussen cognitief vermogen en ras of etniciteit te onderzoeken – en toch zou een "open data" transparantieregeling een dergelijke heranalyse niet verhinderen (Fox Tree, Lleras, Thomas, & Watson, 2022; Lewandowsky & Bishop, 2016). Er is inderdaad overvloedig bewijs dat gegevens over individuen en gemeenschappen (van sociale media, medische diensten, gekwantificeerde zelftechnologieën en/of administratieve bronnen) worden gebruikt op manieren die minderheden en kwetsbare leden van de samenleving discrimineren (Leonelli et al., 2017; Taylor, 2017; Noble, 2018; Eubanks, 2018; Zuboff, 2019).

Kortom, reproduceerbaarheid is deels afhankelijk van inclusiviteit. Maar het bereiken van inclusiviteit vereist bijzondere zorg voor de bescherming van de privacy en anonimiteit van de deelnemers, hetgeen op zijn beurt spanningen kan creëren met de transparantie die velen als essentieel voor de reproduceerbaarheid beschouwen. Ook hier zijn er geen eenvoudige oplossingen voor dit spanningsveld tussen concurrerende aspecten van reproduceerbaarheid.

³ <https://www.apa.org/about/policy/racism-apology>

KADER 4: The Belfast Project en gevolgen

Een opvallend concreet voorbeeld van transparantie als tweesnijdend zwaard, en hoe deze in strijd kan zijn met het recht op privacy, is The Belfast Project, een mondelinge geschiedenis van "The Troubles" in Noord-Ierland die werd georganiseerd door het Boston College in de VS. "The Troubles" verwijst naar het gewelddadige conflict in Noord-Ierland tussen republikeinen (d.w.z. degenen die wilden dat Noord-Ierland een deel zou worden van de Republiek Ierland) en loyalisten (degenen die wilden dat Noord-Ierland deel bleef uitmaken van het Verenigd Koninkrijk), dat uitbrak in de jaren zestig en voortduurde tot 1998, waarbij ongeveer 3500 mensen om het leven kwamen.

The Belfast Project is een verzameling interviews met voormalige republikeinse en loyalistische paramilitairen die uit de eerste hand kennis hebben van bomaanslagen, ontvoeringen en moorden door beide partijen in het conflict. Aan de deelnemers werd vertrouwelijkheid beloofd en dat hun opgenomen interviews pas na hun dood zouden worden vrijgegeven.

In 2011 vaardigde de Amerikaanse procureur-generaal namens de politiedienst van Noord-Ierland dagvaardingen voor de gegevens uit aan Boston College, de beheerder van het project. Na een langdurige rechtszaak en ondanks de bezwaren van de onderzoekers slaagde de politie van Noord-Ierland erin de tapes te bemachtigen.

Kort daarna werd Gerry Adams, de leider van de republikeinse partij Sinn Féin, vier dagen vastgehouden voor verhoor in verband met bewijsmateriaal op de tapes. Het vrijgeven van de opnames van het Boston College maakte ook een einde aan een ander gepland project voor een mondelinge geschiedenis dat was gestoeld op de belofte van vertrouwelijkheid aan de deelnemers, waardoor de inspanningen om een beter historisch archief van The Troubles te creëren op de helling werden gezet (McDonald, 2014).

De spanning tussen transparantie en repliceerbaarheid enerzijds, en de realiteit van onderzoek in gewelddadige omgevingen anderzijds, is niet beperkt tot Noord-Ierland, maar is een bepalend kenmerk van veldonderzoek dat onder moeilijke omstandigheden wordt uitgevoerd (Thaler, 2021).

KADER 5: Reproduceerbaarheid en transparantie: belangrijkste uitdagingen

Er zijn goede redenen om voorzichtig te zijn bij het associëren van de vraag naar reproduceerbaarheid met een simplistisch verlangen naar meer transparantie in het onderzoek. Transparantie is niet altijd wenselijk en moet intelligent worden beheerd, zoals reeds in 2012 werd betoogd in het Royal Society-rapport "Science as an Open Enterprise" (Boulton et al., 2012). Tot de redenen voor deze voorzichtigheid behoren de volgende drie punten van zorg:

1. Privacy en toezicht: het is al duidelijk hoe met name in het geval van onderzoek naar menselijke proefpersonen transparantie kan leiden tot problematisch toezicht, inbreuken op de privacy en misbruik van persoonlijke informatie, zoals ervaren door patiënten van wie de kosten van de ziektekostenverzekering stijgen nadat ze een arts hebben geraadpleegd (Ebeling, 2016; Dennis et al., 2019; Tempini & Leonelli, 2018). Dit kan ook gebeuren bij onderzoek over landschappen en klimaat, aangezien dat onderzoek informatie kan bevatten over gemeenschappen die in de regio leven en een waar instrument kan vormen voor toezicht op die bevolkingsgroepen (Willamson & Leonelli, 2022). Duidelijke voorbeelden hiervan zijn het gebruik van satellietbeelden met hoge resolutie, die kunnen worden gebruikt om de verspreiding van plantpathogenen vast te leggen en schendingen van mensenrechten of oorlogsmisdaden te documenteren – maar die ook gemakkelijk kunnen worden ingezet voor het plannen van vijandige aanvallen of bioprospectie gericht op het uitbuiten van specifieke regio's.

2. Vertrouwen. Het is onmogelijk, zowel in theorie als in de praktijk, om elk aspect van een wetenschappelijke procedure volledig over te brengen: wat onderzoekers bijvoorbeeld moeten doen wanneer ze beslissen hoe ze een experiment moeten beschrijven, is selecteren welke aspecten van hun methoden en omstandigheden het belangrijkste zijn om dat experiment te reproduceren, zodat andere onderzoekers zich op die aspecten kunnen concentreren wanneer ze de resultaten van het experiment onderzoeken. De noodzaak om te selecteren welke aspecten worden beschreven en gedeeld, brengt vertrouwenskwesaties in het debat over reproduceerbaarheid, aangezien onderzoekers elkaar moeten begrijpen en vertrouwen om de belangrijkste aspecten van hun praktijk vast te leggen op een manier die begrijpelijk is voor hun collega's. Dit kan zeer moeilijk blijken in situaties waarin onderzoekers niet dezelfde achtergrond hebben, en dus niet in staat zijn de redenering achter specifieke keuzes inzake delen te begrijpen (Nguyen, 2021). Zonder een dergelijk vertrouwen en wederzijds begrip kan het delen van informatie – zelfs op de meest transparante manier, zoals het online zetten van informatie – nutteloos blijken voor degenen die de resultaten proberen te evalueren en te reproduceren (Leonelli, 2023).

3. Diversiteit. De in punt 1 gemaakte overwegingen over de vele betekenissen en vormen van reproduceerbaarheid binnen het hedendaagse onderzoek hebben gevolgen voor het begrip transparantie. Uit onderzoek naar de houding van wetenschappers tegenover het delen van gegevens bleek dat veel onderzoekers een groot voorstander zijn van beleid inzake Open Data en reproduceerbaarheid, maar tegelijkertijd voelden sommigen zich bedreigd of ondermijnd door OS-initiatieven (Levin et al., 2017; Bezuidenhout et al., 2017; Leonelli, 2022; State of Open Data reports, Science, 2022). Dit was met name het geval voor onderzoekers die in minder zichtbare en goed gefinancierde omgevingen werken en van wie sommigen ernstige moeilijkheden ondervonden bij het gebruik

van open data-archieven en reproduceerbaarheidsprotocollen. Sommige van deze instrumenten weerspiegelen de belangen, middelen en culturele vooronderstellingen van degenen die ze ontwikkelen (en van wie de meesten verbonden zijn aan rijke instellingen uit het Noorden). Onderzoekers melden ook moeilijkheden bij de toegang tot geavanceerde technologieën en software die vaak worden verondersteld bij initiatieven voor gegevensuitwisseling, met name wanneer ze OS-hulpmiddelen en -vaardigheden proberen te verwerven of te ontwikkelen die direct relevant zijn voor hun eigen specifieke prioriteiten en omstandigheden. Er is aantoonbaar een voortdurende discrepantie tussen universele beginselen op hoog niveau die met OS worden geassocieerd, zoals "openheid", "transparantie" en "reproduceerbaarheid", en de verschillende manieren waarop die beginselen worden geïnterpreteerd, afhankelijk van de onderzoeksomgeving in kwestie (Leonelli, 2023). Hiermee moet rekening worden gehouden bij het bepalen van het beleid inzake reproduceerbaarheid en transparantie van onderzoek.

Een ander, vaak over het hoofd gezien spanningsveld is dat de doelstellingen van reproduceerbaarheid en transparantie in conflict kunnen komen met culturele inclusiviteit en diversiteit van de onderzoeksgemeenschap. Er is bezorgdheid dat wetenschappers met een gemarginaliseerde achtergrond (bv. Gekleurde vrouwen) onbedoeld nadeel ondervinden van de doelstellingen van transparantie en openheid. Daar zijn verschillende redenen voor, die door Fox Tree et al. (2022) zijn verwoord. Ten eerste zijn er aanwijzingen uit de VS dat zwarte of "bruine" onderzoekers, ondanks gelijke kwalificaties, minder kans hebben op financiering door federale agentschappen (Ginther et al., 2011). Dit patroon van verwerping en veronachtzaming ontmoedigt het delen van gegevens uit vrees dat een onderzoeker niet de nodige eer krijgt voor zijn of haar werk en gegevens. Een tweede reden is dat onderzoekers uit historisch achtergestelde milieus bij voorkeur gemarginaliseerde gemeenschappen bestuderen die een lange geschiedenis hebben van wantrouwen tegenover de mainstream wetenschap (vaak om zeer goede maar pijnlijke redenen). We merkten al eerder op dat het delen van gegevens uit onderzoek waarbij dergelijke gemeenschappen betrokken zijn bijzonder beladen is met risico's, gezien de mogelijkheid van deanonimisering (die gemakkelijker is bij minderheidsdeelnemers dan bij witte deelnemers) of de mogelijkheid van heranalyse van gegevens om twijfelachtige politieke doelen na te streven (bijvoorbeeld om cognitieve prestaties tussen verschillende etnische groepen te vergelijken; Lewandowsky & Bishop, 2016). Naast de hierboven reeds genoemde gevolgen voor de deelnemers, mag de extra last van de verantwoordelijkheid van onderzoekers met een minderheidsachtergrond niet worden onderschat. We hebben vastgesteld dat de belanghebbenden zich niet algemeen bewust zijn van dit probleem.

Kortom, reproduceerbaarheid en transparantie zijn geen simpele, eenduidige begrippen. Ze vereisen een genuanceerde analyse en vaak niet-triviale afwegingen tussen concurrerende eisen. Transparantie kan op gespannen voet staan met de verwachtingen inzake privacy en reproduceerbaarheid kan de veerkracht van de wetenschappelijke gemeenschap in gevaar brengen. Om een onderzoekscultuur te creëren die reproduceerbaarheid combineert met veerkracht, vertrouwen en eerbiediging van de privacy, moeten deze afwegingen eerlijk en grondig worden onderzocht.

DEEL 4. Het spanningsveld tussen openbaar en particulier belang

De in het vorige punt besproken bezorgdheid over transparantie en een gebrek aan open controle is niet beperkt tot de particuliere sector en is in feite inherent aan zowel de stimuleringsmechanismen als de concurrentieregelingen die gelden voor door de overheid gefinancierd onderzoek.

Eenzijds hebben zowel publiek als privaat gefinancierde onderzoekers reden om uiterst voorzichtig te zijn wat betreft de geldigheid van hun bevindingen. Er staat ongetwijfeld meer op het spel bij commercieel blootgesteld onderzoek, en met name bij onderzoek dat erop gericht is wetenschappelijke inzichten om te zetten in producten die op de wereldmarkt beschikbaar worden gesteld. In die omgeving staan onderzoekers onder enorme druk om resultaten te reproduceren en zeer nauwkeurige en gestandaardiseerde kwaliteitscontroles toe te passen. Er zijn natuurlijk regelmatige inspecties door regelgevende instanties, waaraan particuliere bedrijven moeten voldoen. Maar zelfs naast deze standaardcontroles zijn de financiële en reputatieschade als producten aan het publiek worden vrijgegeven die niet functioneren zoals beloofd, of – erger nog – schade toebrengen aan potentiële gebruikers (zoals in het geval van een geneesmiddel dat dodelijke bijwerkingen blijkt te hebben), enorm, en er is een sterk engagement om mogelijke fouten in het onderzoek op te sporen voordat ze leiden tot producten met gebreken. De vertegenwoordigers van de farmaceutische industrie die in het kader van deze oefening verslag uitbrachten over hun werk, benadrukten inderdaad de strikte opleiding en protocollen die bestaan om goede onderzoekspraktijken te bevorderen en te controleren, waaronder specifieke richtsnoeren voor het reproduceren van bevindingen en een beleid voor het opsporen en corrigeren van fouten in zowel hypothetische als reële situaties.

Anderzijds echter worden zowel publiek als privaat gefinancierde onderzoekers blootgesteld aan stimulansen en eisen om hun activiteiten erop te richten zo snel mogelijk resultaten te produceren, waarbij weinig tijd wordt besteed aan het onderzoeken van implicaties en het documenteren en naar behoren onderzoeken van onderzoeksprocessen. Voor particulier gefinancierde onderzoekers zijn er, naast de enorme druk om snel bruikbare resultaten te produceren (vooral gezien de kosten van de uitvoering van grote experimenten zoals klinische proeven), ernstige belemmeringen voor het vrij delen – en breed onderzoeken – van gegevens. Dit heeft vooral te maken met commerciële gevoeligheid, waarbij bedrijven proberen hun investeringen in onderzoek te beschermen tegen een primeur voor concurrenten; maar er is ook bezorgdheid over de privacy (bijvoorbeeld bij de behandeling van medische gegevens) en over het mogelijke gebruik van gegevens als “wapen” (in het geval van bevindingen die machtige industriële sectoren bedreigen, zoals de industrie van de fossiele brandstoffen in het geval van klimaatverandering).

In het geval van onderzoekers die voor de overheidssector werken, wordt de commerciële gevoeligheid van de resultaten misschien niet zo sterk benadrukt, maar kan de terughoudendheid om resultaten en methoden algemeen beschikbaar te stellen voortkomen uit de stimulans om met andere groepen te concurreren bij het ontdekken van nieuwe technologieën en het maken van nieuwe claims. Het competitieve karakter van academisch onderzoek, waar bijdragen worden gewaardeerd als nieuw en belangrijk afhankelijk van of ze de eerste zijn die een bepaalde ontdekking doen, betekent dat er weinig stimulansen zijn om tijd te besteden aan het onderzoeken en rapporteren van de menselijke en milieukosten en de betekenis van een bepaalde ontdekking (Open Science Policy Platform, 2020). Er is een tendens om geheimhouding te verkiezen boven open samenwerking, vooral in drukke domeinen zoals de biogeneeskunde, en om te vertrouwen op geautomatiseerde systemen die de toegang tot en de analyse van gegevens kunnen versnellen, zoals verschillende vormen van algoritmen voor machinaal leren, maar die ook de neiging hebben om de onderliggende analytische processen te verhullen (Leonelli, 2017). Deze concurrentie is bijzonder problematisch voor onderzoekers uit historisch achtergestelde groepen, om de eerder genoemde redenen.

Deze tendensen worden nog versterkt door de vele manieren waarop particulier en openbaar onderzoek van elkaar afhankelijk zijn, gaande van financiering tot digitale infrastructuren. Er is een sterk spanningsveld tussen de beginselen van reproduceerbaarheid en concurrentie. Gezien het gebrek aan transparantie is het erg moeilijk om de kwaliteitsaanspraken van de industrie te controleren. Ons werd verteld dat ze jaarlijks audits uitvoeren en onderzoekers zeer streng opleiden, en hard werken aan "open wetenschap binnen het bedrijf", maar dit alles blijft binnenshuis (we kregen geen toegang tot opleidingsmateriaal of -programma's, ondanks de uitleg dat dit materiaal vertrouwelijk zou worden behandeld). Wetenschappelijke instituten die werken op de grens tussen openbaar en particulier onderzoek bevinden zich in een bijzonder gunstige positie om deze kwesties aan te pakken door de normen voor controle over de hele linie te verbeteren en te zorgen voor toezicht op de sociale gevolgen van innovatie die is ontwikkeld op basis van door de overheid gefinancierd onderzoek.

KADER 6: Onderzoeksinstellingen aan de translationele spits

Er zijn verschillende onderzoeksinstellingen in Vlaanderen die zich op het kruispunt van de publieke en de private sector bevinden.

- VITO: de missie van VITO is de overgang naar duurzame energie te versnellen en de totstandkoming van een circulaire economie te vergemakkelijken.

- Flanders Make: hun missie is het uitvoeren van hoogtechnologisch onderzoek ter ondersteuning van bedrijven in de maakindustrie om producten en productieprocessen te ontwikkelen en te optimaliseren.
- VIB: het VIB wil de vertaling van biologisch en biomedisch onderzoek in Vlaanderen naar producten voor patiënten, boeren en consumenten ondersteunen en tegelijkertijd nieuwe economische activiteit en nieuwe banen creëren.
- ILVO: de missie van ILVO is het verrichten van onderzoek op het gebied van landbouw, visserij en voeding, met bijzondere aandacht voor duurzaamheid.
- VLIZ: het Vlaams Instituut voor de Zee bevordert de mariene kenniscreatie en excellentie door onderzoek over de oceaan, zeeën, kust en getijdenestuaria.
- Imec: de missie van Imec is om innovatie op het gebied van nano- en digitale technologie mogelijk te maken.

Hoewel deze instituten duidelijk verschillende taken hebben en verschillende gebieden bestrijken, staan ze ook voor gemeenschappelijke uitdagingen die voortvloeien uit de noodzaak een brug te slaan tussen de openbare en de particuliere sector. De ondersteuning van deze instituten is cruciaal voor het vermogen van Vlaanderen om de zogenaamde translationele kloof te overbruggen en ervoor te zorgen dat er een duurzaam en ondersteund traject is om inzichten uit fundamenteel onderzoek naar de markt te brengen. Tegelijkertijd is een belangrijke uitdaging of de onderzoeksinstellingen kunnen zorgen voor een verantwoorde commercialisering van hun onderzoeksresultaten. In ons overleg met belanghebbenden in Vlaanderen hebben we vastgesteld dat er meer aandacht wordt besteed aan de doeltreffendheid van translationeel onderzoek dan aan de verantwoording ervan aan de bredere samenleving. Sommige instituten hebben een duidelijk mandaat om dringende maatschappelijke problemen aan te pakken, zoals afvalbeheersing (VITO), overbevissing (VLIZ) en meer in het algemeen voedselzekerheid (ILVO). Er is echter geen institutioneel orgaan dat erop toeziet dat het transnationale onderzoek dat onder deze vlaggen wordt uitgevoerd, maatschappelijk verantwoord is. Dit is een gemiste kans gezien de gunstige positie van instituten als de bovengenoemde om te zorgen voor een verantwoorde vertaling van publieke bevindingen naar innovatie.

DEEL 5. Aanbevelingen

Op basis van de uitgebreide discussies met belanghebbenden en de in dit rapport bestudeerde literatuur formuleren we een aantal aanbevelingen die de basis moeten vormen voor verdere discussie. Onze aanbevelingen vloeien voort uit onze analyse van de complexiteit van het reproduceerbaarheidsvraagstuk en de spanningen tussen concurrerende waarden en eisen die reproduceerbaarheid met zich meebrengt.

We plaatsen onze aanbevelingen binnen de context van de overkoepelende huidige beloningsstructuur die de academische wereld beheerst. In de internationale academische wereld wordt sterk de nadruk gelegd op een reeks ogenschijnlijk "objectieve" maatstaven om de prestaties te meten, zowel op individueel niveau als op het niveau van de instellingen. Wij vinden deze nadruk misplaatst en contraproductief. De meeste bestaande maatstaven, van het tellen van publicaties of citaties tot andere bibliometrische indicatoren, benadrukken kortetermijnprestaties en zijn daarom vaak tegengesteld aan strategische onderzoeksinspanningen op lange termijn. Bovendien zal elke maatstaf niet alleen de prestaties meten, maar ook het gedrag van onderzoekers en instellingen bepalen. Als bijvoorbeeld het aantal publicaties (in plaats van de kwaliteit ervan) een belangrijke maatstaf is, zullen onderzoekers worden gestimuleerd om meer papers te publiceren, mogelijk ten koste van de kwaliteit.

Naast deze principiële problemen creëert de nadruk op maatstaven een verdere uitdaging voor de reproduceerbaarheid, omdat er geen vaststaande "maatstaven" voor reproduceerbaarheid bestaan. Bovendien ligt zelfs in instellingen waar maatstaven worden aangevuld met kwalitatieve evaluatie, zoals bijvoorbeeld bij het REF-programma van het Verenigd Koninkrijk, de nadruk op het belonen van nieuwigheid boven replicatie en kwaliteit. Voor zover erkenning wordt verleend, bijvoorbeeld door tijdschriften die "badges" toekennen voor open gegevens, pre-registratie enzovoort, kunnen die vrij oppervlakkige indicatoren "openwassen" bevorderen, d.w.z. het gebruik van symbolische gebaren om te voldoen aan de vraag naar reproduceerbaarheid zonder een diepgaand engagement voor reproduceerbaarheid. Momenteel wordt reproduceerbaarheid dus niet beloond met algemeen gebruikte prestatiecriteria. Eerder merkten we al op dat dit een hardnekkige uitdaging vormt in de Vlaamse institutionele context. Het probleem wordt nog verergerd door het feit dat andere activiteiten die bijdragen tot reproduceerbaarheid en kwaliteitscontrole grotendeels onzichtbaar blijven. Reviewing voor tijdschriften of subsidiepanels wordt bijvoorbeeld grotendeels niet erkend, hoewel recente pogingen om deze activiteiten te registreren via platformen als ORCID of Publons (nu opgenomen in Web of Science) de zichtbaarheid van dergelijke essentiële activiteiten trachten te vergroten.

Aanbeveling 1. Een verbintenis tot een systematische heroverweging van de rol van onderzoeksevaluatie en hoe die verband houdt met reproduceerbaarheid.

De *Overeenkomst inzake de hervorming van de onderzoeksbeoordeling*⁴ die de Europese Commissie in januari 2022 lanceerde en die nu door meer dan 350 belanghebbende instellingen is ondertekend, biedt een eerste stap in die richting met haar visie dat “bij de beoordeling van onderzoek, onderzoekers en onderzoeksorganisaties de diverse outputs, praktijken en activiteiten worden erkend die de kwaliteit en de impact van het onderzoek maximaliseren. Daartoe moet de beoordeling hoofdzakelijk gebaseerd zijn op een kwalitatief oordeel, waarbij het nalezen door vakgenoten centraal staat, ondersteund door een verantwoord gebruik van kwantitatieve indicatoren.” Een van de vereisten voor een dergelijke heroverweging is een zorgvuldige evaluatie van de kenmerken en behoeften van de verschillende onderzoeksgebieden, en aandacht voor het vermijden van onbedoelde discriminatie tussen disciplines door het instellen van domeinspecifieke reproduceerbaarheidscriteria als overkoepelend model van “goede praktijk”. Bij de heroverweging moet ook rekening worden gehouden met het sterk verspreide karakter van de kennisproductie, door de nodige erkenning te geven aan activiteiten die soms als “zuiver technisch”, “ondersteunend” of “dienstverlenend” worden geclassificeerd (zoals gegevensbeheer en kwaliteitsbeoordeling), maar die nochtans hoge kwalificaties vereisen en cruciaal zijn voor de vooruitgang van het onderzoek.

Aanbeveling 2. Een verbintenis om het begrip reproduceerbaarheid uit te breiden tot disciplines en benaderingen die gezien hun focus en methoden geen directe of computationele reproduceerbaarheid kunnen toepassen.

Terwijl wij een grote verscheidenheid aan opvattingen over reproduceerbaarheid hebben gesignaleerd, die elk zijn toegesneden op de specifieke kenmerken van de onderzoekspraktijk in de verschillende domeinen, bestaat er in beleidsdebatten over reproduceerbaarheid een tendens om zich alleen te richten op wat past bij sterk gecontroleerde en gestandaardiseerde experimenten. We bevelen aan het begrip reproduceerbaarheid uit te breiden tot de vormen van reproduceerbaarheid die gebruikt worden in minder gestandaardiseerde, meer exploratieve en kwalitatieve vormen van onderzoek, waarvan de nadruk op reproduceerbare expertise veel kan leren aan degenen die pleiten voor een strikte standaardisering van onderzoeksomgevingen als de belangrijkste oplossing voor de reproduceerbaarheidscrisis.⁵

Aanbeveling 3. Een verbintenis om de gevolgen van reproduceerbaarheidsinitiatieven voor de volledige diversiteit van belanghebbenden te onderzoeken, met bijzondere nadruk op jonge onderzoekers, minderheidsgroepen en vrouwen.

⁴ https://research-and-innovation.ec.europa.eu/all-research-and-innovation-news/reforming-research-assessment-agreement-now-final-2022-07-20_en

⁵ Zie ook Onghena (2020) voor aanbevelingen die specifiek zijn toegesneden op de verwezenlijking van reproduceerbaarheid binnen de geesteswetenschappen.

Een verbetering van de reproduceerbaarheid vereist institutionele hervorming en steun, en kan niet worden bereikt door enkel onderzoekers op te leiden in nieuwe methoden. Integendeel, de beschikbaarheid van opleidingsprogramma's (workshops enz.) kan, hoewel goed bedoeld, contraproductief zijn tenzij die initiatieven gepaard gaan met institutionele infrastructuur en erkenning van de deskundigheid en de inspanningen die nodig zijn om de reproduceerbaarheid te verbeteren. Duidelijke structuren voor de erkenning van de deskundigheid van data stewards, technici en beginnende onderzoekers die nauw betrokken zijn bij kwaliteitscontroles en dataconservering, bijvoorbeeld, zouden die inspanningen onmiddellijk zichtbaar maken en helpen de door ons genoemde spanningen tussen de verschillende deelnemers aan het onderzoek aan te pakken. Evenzo zal meer nadruk op een verantwoorde samenwerking met niet-academische belanghebbenden, waaronder geïnteresseerde burgers en niet-academische partners, het reactievermogen en de openheid van zowel openbaar als particulier onderzoek verbeteren. Ten slotte is het van cruciaal belang opleidingsinitiatieven binnen onderzoeksinstellingen te flankeren met correct vergoede, permanente functies ter ondersteuning van bestaande onderzoekers bij de naleving van de normen met betrekking tot Open Science en reproduceerbaarheid, die zeer arbeidsintensief zijn en niet zomaar kunnen worden toegevoegd aan de bestaande zware werklast van onderzoekers.

Aanbeveling 4. *Het verkennen van mechanismen ter ondersteuning van het opbouwen van een onderzoekscultuur van open discussie die steunt op wederzijdse verantwoordingsplicht om de opsporing en correctie van fouten te verbeteren zonder paranoia en wederzijds wantrouwen te creëren.*

Een bijna onvermijdelijk gevolg van meer transparantie is een grotere kans dat fouten worden ontdekt. Niemand is immuun voor het maken van fouten tijdens analyses of experimenten, en het opsporen van fouten is een essentieel aspect van de zelfcorrigerende aard van de wetenschap. De opsporing van fouten brengt momenteel echter reputatieschade mee voor de betrokken onderzoeker, en dus nemen onderzoekers die transparant onderzoek verrichten een risico dat hun collega's die minder transparant zijn, kunnen vermijden. Een belangrijke weg naar reproduceerbaarheid bestaat erin een positieve dialoog over fouten en de betekenis van de getrokken lessen te vergemakkelijken, zonder iets af te doen aan de autonomie van de onderzoekers en het unieke karakter van elke onderzoekssituatie. Een voorbeeld van een constructief initiatief in dit verband is het Journal of Trial and Error, dat werd gestart door een groep promovendi aan de universiteit van Leiden en snel internationaal werd erkend als een plaats waar negatieve resultaten en experimenten die niet zijn gelukt, worden gepubliceerd en een gezond debat op gang wordt gebracht over wat een "significante mislukking" is en wanneer/hoe fouten moeten worden gevierd en erkend op gelijke voet met vermeende onderzoekssuccessen (<https://journal.trialanderror.org/>). Evenzo richten veel Open Access tijdschriften zich nu minder op de beoordeling van wat kan

gelden als “nieuwe ideeën” binnen een bepaald domein (wat vaak een twistpunt is tussen beoordelaars en bevooroordeeld kan zijn ten aanzien van specifieke soorten bijdragen en bijdragers), en meer op de betrouwbaarheid en kwaliteit van de geproduceerde kennis. De nadruk die het FWO legt op negatieve resultaten en integriteit is in dit opzicht ook lovenswaardig, en moet in de toekomst worden versterkt als een manier om een cultuur van delen en leren van mislukkingen evenzeer te ondersteunen als leren van wat wel lukt.

Aanbeveling 5. *Het opzetten van strategieën voor wetenschapscommunicatie die gericht zijn op de betrokkenheid van meerdere doelgroepen bij onderzoeksprocessen, en die ervoor zorgen dat wetenschap wordt gezien als een geavanceerde maar feilbare onderneming, waarvan de resultaten geloofwaardig zijn juist omdat ze zeer nauwkeurig worden onderzocht en voortdurend verder worden geverifieerd.*

Zoals we in dit rapport hebben benadrukt, is het voor de toekomst van de wetenschap en haar rol in de samenleving van cruciaal belang dat de nadruk wordt gelegd op de betrokkenheid van het publiek bij het onderzoeksproces, op een geïnformeerde en constructieve manier en niet op een manier die schadelijk is voor de nauwkeurigheid en betrouwbaarheid van het onderzoek. Daartoe moet worden geïnvesteerd in een strategie voor wetenschapscommunicatie die zowel een sterk engagement (en plaatsen) voor de coproductie van wetenschappelijke kennis omvat, bv. Via citizen science-initiatieven en programma’s voor publieke betrokkenheid, als effectief informatie verstrekt die cruciaal is voor het openbare leven (bv. Bruikbaar advies bij wereldwijde noodsituaties zoals de COVID-19-pandemie).

Conclusies: Hoe helpt het streven naar reproduceerbaarheid ons de wetenschappelijke crisis aan te pakken?

We hebben betoogd dat het streven naar reproduceerbaarheid als overkoepelende epistemische waarde geen toverformule is voor het ontwikkelen of beoordelen van wetenschappelijk onderzoek. Op zich lost het verbeteren van de reproduceerbaarheid niet noodzakelijkerwijs de bezorgdheid over de kwaliteit van het onderzoek op, aangezien het geen oplossing biedt voor systemische problemen met beloningen en stimulansen voor onderzoek, en het geen universele oplossing biedt voor methodologische problemen, aangezien reproduceerbaarheid voor verschillende gebieden en onderzoeksbenaderingen een verschillende betekenis heeft.

Het onderzoeken van de "crisis van de reproduceerbaarheid" biedt echter een uitstekende gelegenheid om systemische problemen in het huidige systeem van wetenschappelijke kennisproductie te onderzoeken, en met name voor het aanpakken van het gebrek aan steun en middelen voor onderzoekers om expliciet en regelmatig discussies te voeren over:

1. methodologische verbintenissen binnen en tussen disciplines en buiten de academische wereld;
2. hoe leren van fouten en problemen gebeurt in de dagelijkse praktijk en wordt gedocumenteerd op een manier die voor anderen begrijpelijk en bruikbaar is (ook voor particulier gefinancierd onderzoek);
3. de strategieën die worden gebruikt om te kiezen welke onderzoekscomponenten op lange termijn behouden moeten blijven, en hoe;
4. hoe Open Science kan worden toegepast om de huidige onderzoekspraktijken te ondersteunen en te verbeteren, en kan worden gericht op deelname van de gemeenschap en maatschappelijk nut.
5. hoe kredietsystemen in de wetenschap de bijdragen van beginnende onderzoekers, technici en ondersteunend personeel moeten erkennen en de nadruk moeten leggen op langetermijnresultaten in plaats van op opbrengsten op korte termijn.

Tot de beschikbare middelen om dergelijke doelstellingen te bereiken behoort het onlangs opgerichte Belgian Reproducibility Network, dat – indien het goed wordt gebruikt – als katalysator zou kunnen dienen voor besprekingen tussen relevante belanghebbenden en derhalve voor vooruitgang op al deze fronten. Het is ook van cruciaal belang dat onderzoeksinstellingen ondersteunende structuren invoeren en hun systemen voor de toewijzing van werklast omvormen om deze af te stemmen op de opleiding en de nieuwe eisen die aan onderzoekers worden gesteld door het beleid inzake reproduceerbaarheid en Open Science. Nieuwe ontwikkelingen op het gebied van kunstmatige intelligentie (AI) zouden kunnen bijdragen tot effectievere manieren om de deugdelijkheid van onderzoek te controleren, althans op sommige gebieden. Recentelijk zijn er boeiende ontwikkelingen geweest in de

wijze waarop AI kan helpen bij wetenschappelijke taken, zoals hulp bij de synthese van bewijsmateriaal voor beleidsmakers (Porciello, Ivanina, Islam, Einarson, & Hirsh, 2020).

Het overkoepelende doel van de Vlaamse onderzoeksgemeenschap zou moeten zijn om praten over reproduceerbaarheid overbodig te maken. Wetenschappers debatteren niet langer over de continentverschuiving, het verband tussen hiv en aids of het bestaan en de oorzaak van klimaatverandering. Over deze kwesties is een consensus bereikt en daarom, misschien enigszins ironisch, debatteren wetenschappers er niet meer over, omdat het voortdurend onderzoeken van goed bevestigde ideeën een verspilling van waardevolle tijd zou zijn. Zodra we een voldoende mate van reproduceerbaarheid hebben bereikt, zou ook deze niet langer veel aandacht moeten krijgen als onderwerp van discussie.

Referenties

- Allison, D. B. et al (2016). A tragedy of errors. *Nature*, 530, 27-30.
- Baba, A., Cook, D. M., McGarity, T. O., & Bero, L. A. (2005). Legislating "sound science": The role of the tobacco industry. *American Journal of Public Health*, 95 S20-S27. doi: 10.2105/AJPH.2004.050963.
- Bailes, M., Bates, S., Bhalerao, V., Bhat, N., Burgay, M., Burke-Spolaor, S., . . . van Straten, W. (2011). Transformation of a star into a planet in a millisecond pulsar binary. *Science*, 333 1717-1720.
- Bartling S, Friesike S (eds., 2014), *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Springer.
- Beaulieu, A. and Leonelli, S. (2021) *Data and Society: A Critical Introduction*. London, UK: SAGE. <https://uk.sagepub.com/en-gb/eur/data-and-society/book/269709>
- Bezuidenhout L, Leonelli S, Kelly A and Rappert B (2017) Beyond the Digital Divide: Towards a Situated Approach to Open Data. *Science and Public Policy* 44(4): 464-475
- Boulton G et al (2012) *Science as an Open Enterprise*. 02/12. London: The Royal Society Science Policy Centre.
- Bringmann, L. F., Elmer, T., & Eronen, M. I. (2022). Back to basics: The importance of conceptual clarification in psychological science. *Current Directions in Psychological Science*. DOI: 10.1177/09637214221096485.
- Brulle, R. J. (2013). Institutionalizing delay: foundation funding and the creation of US climate change counter-movement organizations. *Climatic Change*, 122 681-694. doi: 10.1007/s10584-013-1018-7
- Brulle, R. J. (2018). The climate lobby: a sectoral analysis of lobbying spending on climate change in the USA, 2000 to 2016. *Climatic Change*. doi: 10.1007/s10584-018-2241-z
- Brulle, R. J., Hall, G., Loy, L., & Schell-Smith, K. (2021). Obstructing action: foundation funding and US climate change counter-movement organizations. *Climatic Change*, 166 17. doi: 10.1007/s10584-021-03117-w
- Brysse, K., Oreskes, N., O'Reilly, J., & Oppenheimer, M. (2013). Climate change prediction: Erring on the side of least drama? *Global Environmental Change*, 23 327-337. doi: 10.1016/j.gloenvcha.2012.10.008
- Cataldo, J. K., Bero, L. A., & Malone, R. E. (2010). "A delicate diplomatic situation": tobacco industry efforts to gain control of the Framingham study. *Journal of Clinical Epidemiology*, 63 841-853. doi: 10.1016/j.jclinepi.2010.01.021

Chang, A. C., & Li, P. (2015). Is economics research replicable? Sixty published papers from thirteen journals say "usually not". *Finance and Economics Discussion Series, 2015* (83), 1-26. doi: 10.17016/feds.2015.083

Dennis, S., Garrett, P., Yim, H., Hamm, J., Osth, A. F., Sreekumar, V., & Stone, B. (2019). Privacy versus open science. *Behavior Research Methods, 51* 1839-1848. doi: 10.3758/s13428-019-01259-5

Burgelman, J et al (2019) Open Science, Open Data, and Open Scholarship: European Policies to Make Science Fit for the Twenty-First Century. *Frontiers in Big Data 2*, 43.

Derksen and Morawski (2022) Kinds of replication.

Dunlap, R. E., & McCright, A. M. (2011). Organized climate change denial. In J. S. Dryzek, R. B. Norgaard, & D. Schlosberg (Eds.), *The Oxford handbook of climate change and society* (pp. 144-160). Oxford, UK: Oxford University Press.

Ebeling, M., 2016. *Healthcare and Big Data: Digital Spectres and Phantom Objects*. Palgrave

Edwards MA, Roy S (2017) Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science 34*, 1.

Elliott, K. (2020) A Taxonomy of Transparency in Science. *Canadian Journal of Philosophy 1-14*.

Elliott, Kevin and David B. Resnik. 2019. "Making open science work for science and society." *Environmental Health Perspectives 127*:075002.

Eubanks, V., 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. Illustrated edition, New York: St. Martin's Press.

European Commission (2018) *Open Science: Altmetrics and Rewards*. Final Report for the Mutual Learning Exercise Open Science: Altmetrics and Rewards of the European Commission. URL: <https://rio.jrc.ec.europa.eu/en/policy-support-facility/mle-open-science-altmetrics-and-rewards>

European Commission (2016) *Open Innovation, Open Science, Open to the World. A vision for Europe*, Luxembourg, Publications Office of the European Union.

European Commission (2018) *Open Science: Altmetrics and Rewards*. Final Report for the Mutual Learning Exercise Open Science: Altmetrics and Rewards of the European Commission. URL: <https://rio.jrc.ec.europa.eu/en/policy-support-facility/mle-open-science-altmetrics-and-rewards>

Fecher, B. and Friesike, S. (2014) Open Science: One Term, Five Schools of Thought. In S. Bartling, S. Friesike (eds.), *Opening Science. The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Springer, 17-47.

- Feest, U. (2019) Why Replication Is Overrated. *Philosophy of Science* 86:5, 895-905.
- Fernandez Pinto, M. (2020) Open Science for Private Interests? How the Logic of Open Science Contributes to the Commercialisation of Research. *Frontiers in Research Metrics and Analysis* 5: 588331.
- Fidler, F. and Wilcox, J. (2018) Reproducibility of Scientific Results. In Zalta E (ed) *The Stanford Encyclopedia of Philosophy*.
- Fox Tree, J. F., Lleras, A., Thomas, A., & Watson, D. (2022). *The inequitable burden of open science*. Retrieved from <https://featuredcontent.psychonomic.org/the-inequitable-burden-of-open-science/>
- Freudenburg, W. R., & Muselli, V. (2010). Global warming estimates, media expectations, and the asymmetry of scientific challenge. *Global Environmental Change*, 20 483-491. doi: 10.1016/j.gloenvcha.2010.04.003
- Garvey, A., & Paavola, J. (2021). Community action on natural flood management and the governance of a catchment-based approach in the UK. *Environmental Policy and Governance*, 32 3-16. doi: 10.1002/eet.1955
- Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, ethnicity, and NIH research awards. *Science*, 333 1015-1019. doi: 10.1126/science.1196783
- Güttinger S (2020) The Limits of Replicability. *European Journal for Philosophy of Science* 10:10.
- Hoffman, K. M., Trawalter, S., Axt, J. R., & Oliver, M. N. (2016). Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proceedings of the National Academy of Sciences*, 113 4296-4301. doi: 10.1073/pnas.1516047113
- Hornsey, M. J. (2020). Why facts are not enough: Understanding and managing the motivated rejection of science. *Current Directions in Psychological Science*. doi: 10.1177/0963721420969364
- Joslyn, S., & LeClerc, J. (2016). Climate projections and uncertainty communication. *Topics in Cognitive Science*, 8. doi: 10.1111/tops.12177
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2 196-217. doi: 10.1207/s15327957pspr0203_4
- Kunert, R. (2016). Internal conceptual replications do not increase independent replication success. *Psychonomic Bulletin & Review*, 23 1631-1638. doi: 10.3758/s13423-016-1030-9
- Lash, T. L. (2021). Getting over TOP. *Epidemiology*, 33 (1), 1-6. doi: 10.1097/ede.0000000000001424

- Leonelli S (2023) *The Philosophy of Open Science*. Cambridge University Press.
- Leonelli, S (2022) Open Science and Epistemic Pluralism: Friends or Foes? *Philosophy of Science*. <https://doi.org/10.1017/psa.2022.45>
- Leonelli, S. (2021) Data Science in Times of Pan(dem)ic. *Harvard Data Science Review* 3(1) <https://doi.org/10.1162/99608f92.fbb1bdd6>
- Leonelli S, Lovell B, Fleming L, Wheeler B and Williams H. (2021) From FAIR data to fair data use: Methodological data fairness in health-related social media research. *Big Data and Society* 8 (1) DOI: 10.1177/20539517211010310
- Leonelli S (2018) Re-Thinking Reproducibility as a Criterion for Research Quality. *Research in the History of Economic Thought and Methodology*. Volume 36B, 129-146.
- Leonelli, S. (2017) *Biomedical Knowledge Production in the Age of Big Data*. Report for the Swiss Science and Innovation Council, published online November 2017: http://www.swir.ch/images/stories/pdf/en/Exploratory_study_2_2017_Big_Data_SSIIC_EN.pdf
- Leonelli, S., Davey, R., Arnauld, E., Parry, G. and Bastow, R. (2017) Data Management and Best Practice in Plant Science. *Nature Plants* 3, 17086. doi: 10.1038/nplants.2017.86
- Levin N, Leonelli S et al (2016) How Do Scientists Understand Openness? Exploring the Relationship between Open Science Policies and Research Practice. *Bulletin for Science & Technology Studies* 36: 128-41.
- Levy, K. E., & Johns, D. M. (2016). When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data & Society*, 3 . doi: 10.1177/2053951715621568
- Lewandowsky, S. (2021). Climate change, disinformation, and how to combat it. *Annual Review of Public Health*, 42 1-21. doi: 10.1146/annurev-publhealth-090419-102409
- Lewandowsky, S., & Bishop, D. (2016). Research integrity: Don't let transparency damage science. *Nature*, 529 459-461. doi: 10.1038/529459a
- Lewandowsky, S., Mann, M. E., Brown, N. J. L., & Friedman, H. (2016). Science and the public: Debate, denial, and skepticism. *Journal of Social and Political Psychology*, 4 537-553. doi: 10.5964/jsp.p.v4i2.604
- Lewandowsky, S., & Oberauer, K. (2016). Motivated rejection of science. *Current Directions in Psychological Science*, 25 217-222. doi: 10.1177/0963721416654436
- Lewandowsky, S., & Oberauer, K. (2020). Low replicability can support robust and efficient science. *Nature Communications*, 11 358. doi: 10.1038/s41467-019-14203-0

- Lewandowsky, S., Oreskes, N., Risbey, J. S., Newell, B. R., & Smithson, M. (2015). Seepage: Climate change denial and its effect on the scientific community. *Global Environmental Change*, 33 1-13. doi: 10.1016/j.gloenvcha.2015.02.013
- Love, A. (2020) Developmental Biology. *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), Edward N. Zalta (ed.) <https://plato.stanford.edu/archives/sum2022/entries/biology-developmental/>
- Lowes, S., & Montero, E. (2021). The Legacy of Colonial Medicine in Central Africa. *American Economic Review*, 111 1284-1314. doi: 10.1257/aer.20180284
- Ma, M. A., Gutiérrez, D. E., Frausto, J. M., & Al-Delaimy, W. K. (2021). Minority Representation in Clinical Trials in the United States: Trends Over the Past 25 Years. *Mayo Clinic Proceedings*, 96 264-266. doi: 10.1016/j.mayocp.2020.10.027
- Mak, W. W. S., Law, R. W., Alvidrez, J., & Pérez-Stable, E. J. (2007). Gender and ethnic diversity in NIMH-funded clinical trials: Review of a decade of published research. *Administration and Policy in Mental Health and Mental Health Services Research*, 34 497-503. doi: 10.1007/s10488-007-0133-z
- McDonald, H. (2014). *Boston College tapes fiasco harms search for truth in Troubles*. Retrieved from <https://www.theguardian.com/politics/2014/may/06/boston-college-tapes-fiasco-northern-ireland-troubles>
- Mirowski, Philip. (2018). The future of (open) science." *Social Studies of Science* 48(2):171-203.
- National Academies of Sciences, Engineering, and Medicine 2019. Reproducibility and Replicability in Science. Washington, DC: The National Academies Press. <https://doi.org/10.17226/25303>
- Nguyen, C. Thi (2021) Transparency is Surveillance. *Philosophy and Phenomenological Research* 00: 1-31
- Noble, S.U., 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. Illustrated edition, New York: NYU Press.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115, 2600-2606. doi: 10.1073/pnas.1708274114
- Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, 26, 1596-1618. doi: 10.3758/s13423-019-01645-2
- Onghena, P. (2020) *Repliceerbaarheid in de KVAB empirische menswetenschappen*. KVAB Standpunt 64, 2020.
- Open Science Collaboration. (2015). Estimating the Reproducibility of Psychological Science. *Science*, 349(6251), aac4716. doi:10.1126/science.aac4716.

- Open Science Policy Platform (2020) *Progress on Open Science: Towards a Shared Research Knowledge System. Final Report of the Open Science Policy Platform*: European Commission, Directorate-General for Research and Innovation. URL: https://ec.europa.eu/research/openscience/pdf/ec_rtd_ospp-final-report.pdf#view=fit&pagemode=none doi: 10.2777/00139
- Romero F (2019) Philosophy of Science and the Replicability Crisis. *Philosophical Compass* 14: e12633.
- Oreskes, N. (2019) *Why Trust Science?* Princeton, NJ: Princeton University Press.
- Oreskes, N., & Conway, E. M. (2010). *Merchants of doubt*. London, UK: Bloomsbury Publishing.
- Porciello, J., Ivanina, M., Islam, M., Einarson, S., & Hirsh, H. (2020). Accelerating evidence-informed decision-making for the Sustainable Development Goals using machine learning. *Nature Machine Intelligence*, 2 559-565. doi: 10.1038/s42256-020-00235-5
- Reardon, S. (2011). Decrying CIA Vaccination Sham, Health Workers Brace for Backlash, 333 395-395. doi: 10.1126/science.333.6041.395
- Rubin, M. (2017). When does HARKing hurt? identifying when different types of undisclosed post hoc hypothesizing harm scientific progress. *Review of General Psychology*, 21 308-320. doi: 10.1037/gpr0000128
- Rubin, M., & Donkin, C. (2022). Exploratory hypothesis tests can be more compelling than confirmatory hypothesis tests. *Philosophical Psychology*, 1-29. doi: 10.1080/09515089.2022.2113771
- Schwartz, J. (2018). "transparency" as mask? the EPA's proposed rule on scientific data. *New England Journal of Medicine*. doi: 10.1056/NEJMp1807751.
- Science, Digital et al. (2022): *The State of Open Data 2022*. Digital Science. Report. <https://doi.org/10.6084/m9.figshare.21276984.v5>
- Radder (1996) *In and About the World*. New York: SUNY Press.
- Strømmland, E (2019) Preregistration and reproducibility. *Journal of Economic Psychology*, 75, Part A, 102143, <https://doi.org/10.1016/j.joep.2019.01.006>
- Taylor L., 2017. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society*, 4(2), pp.205395171773633.
- Tempini, N., & Leonelli, S. (2018). Concealment and discovery: The role of information security in biomedical data re-use. *Social Studies of Science*, 48 663-690. doi: 10.1177/0306312718804875
- Thaler, K. M. (2021). Reflexivity and Temporality in Researching Violent Settings: Problems with the Replicability and Transparency Regime. *Geopolitics*, 26 (1), 18-44. doi: 10.1080/14650045.2019.1643721

United Nations (2019) Towards Global Open Science: Core Enabler of the UN 2030 Agenda. https://research.un.org/ld.php?content_id=50970566

Weber, M. (2022) *Philosophy of Developmental Biology*. Cambridge: Cambridge University Press.

Whatmore, S. J., & Landström, C. (2011). Flood apprentices: an exercise in making things public. *Economy and Society*, 40, 582-610. doi: 10.1080/03085147.2011.602540

Worrall, J. (2014). Prediction and accommodation revisited. *Studies in History and Philosophy of Science Part A*, 45 54-61. doi: 10.1016/j.shpsa.2013.10.001

Yates, I., Byrne, J., Donahue, S., McCarty, L., & Mathews, A. (2020). Representation in clinical trials: A review on reaching underrepresented populations in research. *Clinical Researcher* 34 (7).

Zuboff, S., 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. 1st ed. New York: Public Affairs.

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2017). Making replication mainstream. *Behavioral and Brain Sciences*, 41 E120. doi: 10.1017/S0140525X17001972

Bijlage 1 Cv van de twee Denkers:

Stephan Lewandowsky, professor cognitieve psychologie aan de Universiteit van Bristol, met speciale belangstelling voor de manier waarop desinformatie in de samenleving kan blijven bestaan, voor de manier waarop mythen en desinformatie zich kunnen verspreiden, en voor de variabelen die bepalen of mensen wetenschappelijk bewijs al dan niet aanvaarden, bijvoorbeeld rond vaccinaties of klimaatwetenschap. Opmerkelijke bijdragen:

- Lezing "Being open but not naked" op de 6de World Conference on Research Integrity in Amsterdam (2017): <https://www.youtube.com/watch?v=3A8QUSSPPhk>
- Paper samen met Klaus Oberauer "Low replicability can support robust and efficient science" in Nature Communications, januari 2020 <https://www.nature.com/articles/s41467-019-14203-0>.
- Een sterke presentatie tijdens de ALLEA-workshop "Ethical Aspects of Open Access" in Brussel begin 2018

Sabina Leonelli, wetenschapsfilosofoe aan de Universiteit van Exeter, Turing Institute, Global Young Academy, Open Science Group, co-directeur van het Exeter Centre for the Study of the Life Sciences (Egenis), leidt het onderzoeksonderdeel Data Studies; Turing Fellow aan het Alan Turing Institute in Londen. Hoofdredacteur van het internationale tijdschrift History and Philosophy of the Life Sciences, samen met professor Giovanni Boniolo, en redacteur bij Harvard Data Science Review. Vice-voorzitter van de European Philosophy of Science Association, stuurgroep lid van de Society for the Philosophy of Science in Practice, Plan S-ambassadeur.

In 2021-2022, Fellow aan het Wissenschaftskolleg zu Berlin, met een project over "Excellence and Diversity in Global Scientific Practice".

<https://www.godan.info/news/making-crop-data-sharing-responsible-and-reliable>

<https://www.yumpu.com/en/document/read/8933819/sabina-leonelli-cv-esrc-genomics-network>

Bijlage 2 De stuurgroep van het Denkersprogramma

De Denkers:

Sabina Leonelli, Universiteit van Exeter
Stephan Lewandowsky, Universiteit van Bristol
Coördinatoren:
Joos Vandewalle, KTW, KU Leuven
Patrick Onghena, KMW, KU Leuven

Medewerkers:

Inez Dua, medewerker KVAB, logistieke en administratieve ondersteuning
Bert Seghers, medewerker KVAB, ondersteuning De Jonge Academie

Leden:

Freddy Dumortier, Vast Secretaris KVAB, UHasselt
Jan De Houwer, UGent
Sadia Vancauwenbergh, UHasselt
Geert Molenberghs, KU Leuven, UHasselt, KAGB
Mieke Van Houtte, UGent, KMW
Willy Verstraete, FWO, KNW, UGent
Geert Leroux-Roels, UZ Gent, KAGB
Inge Van Nieuwerburgh, UGent
Roosmarijn Vandenbroucke, De Jonge Akademie en Belgian Reproducibility Network
Joris Vandendriessche, De Jonge Akademie en Belgian Reproducibility Network

RECENTE STANDPUNTEN

55. Addendum. Jean-Pierre Henriët. – *Mijlpalen in internationale wetenschappelijke samenwerking*, KVAB/Klassen Natuurwetenschappen, 2017.
56. Piet Swerts, Piet Chielens, Lucien Posman – *A Symphony of Trees. Wereldcreatie naar aanleiding van de herdenking van de Derde Slag bij Teper, 1917*, KVAB/Klasse Kunsten, 2017.
57. Willy Van Overschée e.a. – *De mobiliteit van morgen: zijn we klaar voor een paradigmawissel?*, KVAB/Klasse Technische Wetenschappen, 2018.
58. Tinne De Laet e.a. - *"Learning Analytics" in het Vlaams hoger onderwijs*, KVAB/Klasse Technische Wetenschappen, 2018.
59. Dirk Van Dyck, Elisabeth Monard, Sylvia Wenmackers e.a. – *Onderzoeker-gedreven wetenschap. Analyse van de situatie in Vlaanderen*, KVAB/Klasse Natuurwetenschappen, 2018.
60. Liliane Schoofs – *Doctoraathouders geven het Vlaanderen van morgen vorm*, KVAB/Klasse Natuurwetenschappen, 2018.
61. Luc Bonte, Aimé Heene, Paul Verstraeten e.a. – *Verantwoordelijk omgaan met digitalisering. Een oproep naar overheden en bedrijfsleven, waar ook de burger toe kan/moet bijdragen*, KVAB/Klasse Technische Wetenschappen, 2018.
62. Jaak Billiet, Michaël Opgenhaffen, Bart Pattyn, Peter Van Aelst – *De strijd om de waarheid. Over nepnieuws en desinformatie in de digitale mediawereld*, KVAB/Klasse Menswetenschappen, 2018.
63. Christoffels Waelkens. – *De Vlaamse Wetenschapsagenda en interdisciplinariteit. Leren leven met interdisciplinaire problemen en oplossingen*, KVAB/Klasse Natuurwetenschappen, 2019.
64. Patrick Onghena – *Repliceerbaarheid in de empirische menswetenschappen*, KVAB/Klasse Menswetenschappen, 2020.
65. Mark Eyskens – *Als een virus de mensheid gijzelt. Oorzaken en gevolgen van de Coronacrisis*, KVAB/Klasse Menswetenschappen, 2020.
66. Jan Rabaey, Rinie van Est, Peter-Paul Verbeek, Joos Vandewalle - *Maatschappelijke waarden bij digitale innovatie: wie, wat en hoe?*, KVAB - Denkersprogramma 2019, KVAB/Klasse Technische Wetenschappen, 2020.
67. Oana Dima (auteur), Dirk Inzé, Hubert Bocken, Pere Puigdomènech, René Custers (eds)., *Genoombewerking voor veredeling van landbouwgewassen. Toepassingen van CRISPR-Cas9 en aanverwante technieken*, ALLEA-KVAB/Klasse Natuurwetenschappen, 2020.
68. Marie-Claire Foblets, *De multiculturele samenleving en de democratische rechtsstaat – Hoe vrijwaren we de sociale cohesie?*, KVAB/Klasse Menswetenschappen 2020
69. Joost Van Roost, Luc Van Nuffel, Pieter Vingerhoets e.a., *De rol van gas in de Belgische energietransitie – Aardgas en Waterstof*, KVAB/Klasse Technische Wetenschappen, 2020.
70. Richard Bardgett, Joke Van Wensem, *Bodem als natuurlijk kapitaal* – KVAB Denkersrapport 2020, KVAB/Klasse Technische Wetenschappen, 2021
71. Jos Smits e.a., *Multifunctionele eilanden in de Noordzee*, KVAB/Klasse Technische Wetenschappen, 2021.
72. Elisabeth Monard, red., *Kunst, Wetenschap en Technologie in Symbiose*, KVAB/Klasse Technische Wetenschappen, 2021.
73. Jan Wouters, Maaïke De Ridder, *De problematiek van de rechtsstaat en democratische legitimiteit binnen de Europese Unie*, KVAB/Klasse Menswetenschappen, 2021.
74. Hilde Heynen, Bart Verschaffel, e.a., *Architectuurkwaliteit vandaag, Reflecties over architectuur in Vlaanderen*, KVAB/Klasse Technische wetenschappen en Klasse Kunsten, 2021.
76. Bea Cantillon, *Het armoedevraagstuk en de tragiek van de welvaartsstaat, Zeven termen voor een nieuw sociaal contract*, KVAB/Klasse Menswetenschappen, 2022.
78. Jo Tollebeek, Marc Boone en Karel van Nieuwenhuyse, *Een Canon van Vlaanderen, Motieven en bezwaren*, KVAB Klasse Menswetenschappen, 2022.
79. Luc Taerwe e.a., *Duurzaam Beheer van Infrastructuur, Niet alleen een kwestie van budgetten*, KVAB/Klasse Technische Wetenschappen, 2022.
80. Willem Salet, Marleen Spiekman, Staf Roels, Tom Coppens, Ivo Van Vaerenbergh, *Naar klimaatneutrale woongebouwen in 2050*, KVAB Denkersprogramma 2022, KVAB/Klasse Technische Wetenschappen, 2022.



Om te begrijpen wat reproduceerbaarheid precies inhoudt, moet breder worden gekeken dan louter aandacht te hebben voor de empirische herhaalbaarheid van resultaten uit wetenschappelijk onderzoek. Een bredere opvatting van reproduceerbaarheid omvat componenten die verwijzen naar transparantie, onderzoekspraktijken en de rol van theorie in de wetenschap. Bovendien kan de relatie tussen deze componenten en reproduceerbaarheid complex en contra-intuïtief zijn. Reproduceerbaarheid is belangrijk, maar het neemt verschillende vormen aan die aangepast zijn aan de verschillende onderzoeksituaties, afhankelijk van de specifieke doelen, methoden, materialen en omstandigheden van onderzoek. Gezien deze noodzakelijke variatie moet reproduceerbaarheid worden geconceptualiseerd en geïmplementeerd op een manier die niet te eng is, waardoor het onderzoek onbedoeld te sterk zou worden beperkt, noch te breed, waardoor het als criterium niet meer werkbaar zou zijn.

De denkers formuleren in dit standpunt positieve acties en aandachtspunten voor onderzoeksinstellingen en onderzoeksfondsen om te komen tot robuuste, reproduceerbare en geloofwaardige kennis, waarbij een belangrijke rol is weggelegd voor een herziening van de manier waarop de kwaliteit van wetenschappelijk onderzoek wordt beoordeeld. Ze hebben daarbij aandacht voor de spanningen en belangenconflicten die leven bij de diverse betrokkenen. Ze presenteren bovendien concrete inzichten in verband met reproduceerbaarheid die nuttig zijn voor wetenschapscommunicatoren. Ten slotte formuleren zij ook concrete opmerkingen die betrekking hebben op de kwaliteit en de waarde van het onderzoek in het publieke en private domein.

De reeks Standpunten van de Academie is een bijdrage tot het wetenschappelijk onderbouwd debat over actuele maatschappelijke en artistieke thema's. De auteurs, leden en werkgroepen van de Academie schrijven in eigen naam, onafhankelijk en met volledige intellectuele vrijheid. De goedkeuring voor publicatie door een of meerdere Klassen van de Academie waarborgt de kwaliteit van de gepubliceerde studies.