2020 | Faculty of Medicine and Life Sciences

UHASSELT

KNOWLEDGE IN ACTION

Maastricht University

Doctoral dissertation submitted to obtain the degree of Doctor of Biomedical Sciences, to be defended by

Noémie Aubert Bonn

DOCTORAL DISSERTATION

The failure of success – Careers, cultures, and integrity in science

Promoter: Prof. Dr Wim Pinxten | UHasselt

2020 | Faculty of Medicine and Life Sciences

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Doctoral dissertation submitted to obtain the degree of
Doctor of Biomedical Sciences, to be defended by

Noémie Aubert Bonn

DOCTORAL DISSERTATION

The failure of success – Careers, cultures, and integrity in science

D/2020/2451/ 50

Promoter: Prof. Dr Wim Pinxten | UHasselt

Promoter:	Prof. dr. Wim Pinxten	Hasselt University, Belgium
Chair:	Prof. dr. Marcel Ameloot	Hasselt University, Belgium
Jury members:	Dr. Anna Hatch	American Society for Cell Biology; San Francisco Declaration on Research Assessment, USA
	Prof. dr. Sven Hendrix	Hasselt University, Belgium
	Prof. dr. Ivo Lambrichts	Hasselt University, Belgium
	Dr. Brian C. Martinson	HealthPartners Institute; Minneapolis VA Medical Center; University of Minnesota, USA
	Prof. dr. Frank Miedema	UMC Utrecht, Utrecht University, The Netherlands
	Prof. dr. Geert Molenberghs	Hasselt University; KU Leuven, Belgium
	Ann Peters, MBA	Institute of Tropical Medicine Antwerp, Belgium

Initial submission: 6 February 2020
Printed: 1 October 2020
Defence: 14 October 2020

This project was funded by the Bijzonder Onderzoeksfonds (BOF) from Hasselt University, grant number 15NI05.

TABLE OF CONTENT

“We need less research, better research, and research done for the right reasons”

Doug Altman, 1994

Abbreviations

The following are common abbreviations I use throughout the thesis. For Dutch terms I add, where available, the official translation or, where unavailable, an unofficial translation in square brackets.

BOF: Bijzondere Onderzoeksfonds / Special research funds

DORA: San Francisco Declaration on Research Assessments

EP: Editors or publishers of scientific journals (see Chapter 3 Table 1)

FA: Funding agency (see Chapter 3 Table 1)

FWO: Fonds Wetenschappelijk Onderzoek / Research Foundation - Flanders

KU Leuven: Katholiek Universiteit Leuven [Catholic University of Leuven]

ILVO: Instituut voor Landbouw, Visserij en Voedingsonderzoek / Research institute for agriculture, fisheries, and food

IMEC: Interuniversity Microelectronics Centre

ITM: Institute of Tropical Medicine Antwerp

LT: Lab technicians (see Chapter 3 Table 1)

PhD: In this thesis, I use PhD to indicate a doctoral research degree. The PhD is the highest degree granted in a university, and depending on countries, it can be called a post-graduate degree, a doctorate, or a doctoral degree; I use PhD to refer to any of those, and to abbreviate 'PhD students' in select chapters (see Chapter 3 Table 1)

PMI: Policy makers or influencers (see Chapter 3 Table 1)

PostDoc: Post-doctoral researchers (see Chapter 3 Table 1)

RCC: Researchers who changed career, namely individuals who were involved in academia but decided to pursue a career outside academia (see Chapter 3 Table 1)

RI: Research integrity

RIL: Research institution leaders, such as deans, directors of doctoral schools, or directors of research (see Chapter 3 Table 1)

RIN: Research integrity network members, meaning researchers or other experts involved in research on research integrity (see Chapter 3 Table 1)

QRP: Questionable Research Practices

UAntwerpen: Universiteit Antwerpen / University of Antwerp

UGent: Gent Universiteit / Ghent University

UHasselt: Hasselt Universiteit / Hasselt University

VIB: Vlaams Instituut voor Biotechnologie [Flemish institute for biotechnology]

VITO: Vlaamse Instelling voor Technologisch Onderzoek [Flemish institution for technological research]

VLIZ: Vlaams Instituut voor de Zee [Flemish institute for the sea]

VLIR: Vlaamse interuniversitare Raad / Flemish Interuniversity Council

VUB: Vrije Universiteit Brussels [Free University of Brussels]

General abstract

ABSTRACT

Science is powerful and vulnerable at the same time. On the one hand science is the basis for innovation and for allowing us to understand the world we live in. On the other hand, science is also performed by humans whose behaviours may compromise integrity. Although scientists are expected to care for the integrity of science, doing so can impose difficult dilemmas.

The present thesis investigates the connections between research integrity, research cultures, and research success. Our findings are presented in three steps, each of which uses a distinct methodology.

In the first step, we analysed and compared a decade of research literature on research integrity. Our analysis revealed two important blind spots in the field of research integrity. First, although issues from the research system (e.g., pressures, perverse incentives, and competition) are most frequently identified as causes for misconduct and questionable research practices, approaches to foster integrity generally tackle researchers' knowledge and compliance rather than documented problems from the research system. Second, although past research on research integrity thoroughly captures the perspectives of researchers, it largely overlooks the perspectives of other key stakeholders.

In the second step, we addressed these two blind spots by conducting interviews and focus groups with a whole array of research stakeholders to discuss issues inherent to the research system. These discussions revealed that research assessments are an important cause for concern in current academia. Indeed, current assessments overvalue research outputs but largely ignore important research processes that are essential in protecting the integrity of science. As a result, researchers often feel the need to compromise on integrity in order to advance or simply maintain their careers in academia. Although most interviewees agreed that current research assessments are inadequate, nobody felt able to instigate a change. Instead, actor groups tended to blame one another for the inadequacies of the current system and to lose faith in the possibility for change.

In the final step of the project, we built a survey to capture the perspective of researchers on success indicators that raised disagreement in the interviews and focus groups. We found that success indicators related to openness, quality, and

innovation were considered important or even essential in advancing science. However, these indicators were often thought to be irrelevant in advancing researchers' careers. Conversely, indicators which denoted the prestige and competitiveness of researchers were considered important in advancing researchers' careers, but largely irrelevant or even detrimental in advancing science. These responses evidence an obvious need to rethink research assessments so that they can value openness, quality, and innovation. Nonetheless, responses also revealed that the resources and infrastructures necessary to support openness and quality practices are largely missing, and that such resources must become available before changes to research assessments take place.

Considering these cumulative findings, I conclude by proposing four recommendations which could help promote better science. First, I argue that approaches meant to foster research integrity should target the faulty dynamics of the research system rather than focus on individual researchers. Second, I propose that research assessments must be adapted to reflect our aspirations for high quality science. In this regard, I suggest that research assessments (i) must be based on transparent and reflective methods, (ii) must consider the value of team efforts, (iii) must recognise research processes even when those are not associated with positive outputs, (iv) must remain realistic in their demands and expectations, and v) must be addressed at all levels. Third, I argue that we must rethink the structure of academic careers and recognize the issues caused by current insecure and precarious climates. Finally, I support that we must discuss and collaborate between actor groups so that we can combine existing efforts into broad and coordinated approaches to make science better.

SAMENVATTING

Wetenschap is zowel krachtig als kwetsbaar. Ze maakt het enerzijds mogelijk om te innoveren en om de wereld waarin we leven beter te begrijpen, terwijl ze anderzijds wordt beoefend door feilbare mensen die door hun gedrag de wetenschappelijke integriteit kunnen schenden. Hoewel we mogen veronderstellen dat wetenschappers begaan zijn met de integriteit van hun werk, kunnen ze hierbij complexe dilemma's ondervinden.

In deze dissertatie worden de verbanden tussen wetenschappelijke integriteit, onderzoekscultuur en succes in de wetenschap onderzocht. De resultaten van dit onderzoek worden in drie delen voorgesteld, waarbij telkens een specifieke methodologie wordt gebruikt. In een eerste deel wordt de vakliteratuur in het domein van de wetenschappelijke integriteit die het voorbije decennium werd gepubliceerd geanalyseerd. Dit bracht twee belangrijke blinde vlekken in het onderzoeksveld aan de oppervlakte. Ten eerste worden de oorzaken van wetenschappelijke wanpraktijken vaak gezocht in het 'systeem' (zoals druk op onderzoekers, de competitie en het beloningssysteem), terwijl de aanpak van zulke wanpraktijken zich niet op dit systeem, maar op de individuele onderzoekers richt (bijvoorbeeld hoe hun kennis over integriteit verhoogd kan worden en hoe onderzoekers de regels beter kunnen naleven). Ten tweede heeft het onderzoek in het domein zich voornamelijk op onderzoekers gericht, en werden anderen spelers in het domein veelal genegeerd.

In een tweede deel werden deze blinde vlekken nader geanalyseerd. In een reeks interviews en focusgroepen waarin een grote diversiteit aan actoren uit de wetenschappelijke wereld werd betrokken, werd verkend hoe aspecten van het 'systeem' het gedrag van wetenschappers beïnvloeden, met mogelijk ongunstige gevolgen voor de wetenschappelijke integriteit. In dit deel van het onderzoek werd duidelijk dat met name de evaluatie van onderzoekers een belangrijke bron van bezorgdheid vormt, omdat zulke evaluaties zodanig focussen op productiviteit en publicatie dat andere belangrijke aspecten van het onderzoek, die vaak essentieel zijn voor de wetenschappelijke integriteit, buiten beschouwing blijven. Bijgevolg zet het nastreven van (het behoud van) een academische loopbaan de wetenschappelijke integriteit vaak onder druk. Terwijl de meerderheid van de geïnterviewden het er over eens waren dat de huidige manier van evalueren niet

voldoet, voelde niemand zich in staat om hierin verandering te brengen. Integendeel: verschillende actoren verweten elkaar voor de deficiënties in het huidige systeem en het verlies van vertrouwen in mogelijkheden tot verandering.

In het derde en laatste deel van dit project hebben we middels een vragenlijst onderzoekers bevraagd over hun visie op wetenschappelijk succes. Hierbij werd uitgegaan van de discussies die hierover tijdens de focusgroepen en interviews in het tweede deel van dit onderzoek gevoerd werden. Uit de respons op de vragenlijst bleek dat inzetten op openheid, kwaliteit en innovatie belangrijk of zelfs essentieel zijn voor succes, terwijl investeringen hierin niet bijdragen tot de uitbouw van een wetenschappelijke carrière. Omgekeerd werden prestige en competitiviteit belangrijk geacht voor de wetenschappelijke carrière, terwijl deze aspecten irrelevant of zelfs schadelijk voor goede wetenschap zijn. Deze respons suggereert een duidelijke nood om de evaluatie van onderzoekers te herzien zodat openheid, kwaliteit en innovatie geherwaardeerd kunnen worden. Tegelijkertijd gaven de respondenten ook aan dat de middelen en infrastructuur om openheid en kwaliteit in praktijk te bewerkstelligen grotendeels ontbreken, en dat hierin verandering gebracht moet worden vooraleer de evaluatie herzien kan worden.

Vanuit deze bevindingen worden vier aanbevelingen gemaakt, die goede wetenschapsvoering kunnen bevorderen. Vooreerst moeten inspanningen om de wetenschappelijke integriteit te versterken zich richten op de dynamiek van het systeem, eerder dan op individuele onderzoekers. Ten tweede moet de evaluatie van onderzoekers onze verwachtingen van goede wetenschap reflecteren, en daarom moet deze evaluatie zelf (i) zich baseren op transparante en reflectieve methodes, (ii) ook het werk dat gezamenlijk in onderzoeksgroepen wordt geleverd correct waarderen, (iii) goed onderzoekswerk (als proces) waarderen, zelfs als het niet tot gepubliceerde output leidt, (iv) realistische verwachtingen aan onderzoekers stellen, en (v) op alle niveaus worden aangepakt. Ten derde moet de opbouw van academische loopbanen herzien worden, met aanacht voor onzekerheid en kwetsbaarheid. Tot slot is er meer dialoog en samenwerking nodig tussen de diverse actoren in het onderzoekslandschap om tot een sterk, omvattend en goed gecoördineerd plan van aanpak voor betere wetenschap te komen.

Preface

The control tower

I often receive funny reactions when presenting my work in conferences. "*You do research on research?!*" Hence, before diving into a thesis entirely dedicated to 'research on research', I believe that it would be interesting to give a bit of the human context behind the project, a bit of the reasons that motivated my participation in this PhD.

When I was very young, I remember spending a fair bit of time at the airport with my family. My dad was a researcher in Earth remote sensing and he travelled quite often for his work. Driving him to the airport was one of our favourite family activity. We were three kids of similar age, so for us the futuristic, vast, and quiet building was a marvellous playground. But what we liked most of all was watching the airplanes take off. My dad had found this big window that overlooked the runway without needing to pass the customs. We would sit there and watch the planes, making bets on their destination and trying to guess their Boeing numbers. Even though there was not much to it, no promises of travel, no extraordinary difference between take offs, not even the loud roar and wind from the engine, we could enjoy that activity for hours without getting tired. We would always stay until my dad had almost missed his flight. He would give us a kiss — one kiss for each day he would be away — and we would go back home amazed and fascinated.

Doing research on research can feel like watching planes take off. You look at them go, you admire their strength and power, you wonder where they are headed, you try to understand what enables them to fly, but you are not part of the trip. So why do research on research? Why observe rather than board?

Everyone has a different story. Mine goes like this. I have known I wanted to be a researcher for a very long time. Almost as long as my first guesses at Boeing numbers (although at the time I probably still hoped to become a professional hockey player, an astronaut, or a cat-sitter). In any case, while I was growing up I did everything it takes to become a researcher. Early on, my interest focalised on cognitive neuroscience and psychiatry. The mysteries of the brain fascinated me. I undertook a Bachelor's degree and aimed for a Master's on the topic, hoping to add to our understanding of how the brain works. But as soon as I started my Master's, the objective shifted. What mattered then was not so much the added knowledge, but rather the added lines on my CV, the publications, and the connections I built. The objective shifted from the common intellectual goal to my individual career. It became normal for me to think in terms of CV lines, to make project plans starting at the number of papers I wanted to achieve, and to attend conferences simply to make connections that could help me publish more. *That* became science to me. I was still very driven to become a researcher, and near the end of my Master's my supervisor offered that I help him build a tenure application to learn how research careers work. This is where things changed. After months of hard work retrieving decade-old documents to prove all the small activities described on his CV, we finally had a dossier put together. But even then, more publications were asked before tenure could be granted. This outcome changed my vision of science. What I considered to be an activity meant to advance knowledge suddenly appeared like a production line in which trust — the foundation of science — was absent. In the months that followed, I stumbled across an old funding application my dad had submitted just a few months before he passed away. In the application, my father takes time to detail the following:

> *Le principal frein à la productivité scientifique pour la majorité des chercheurs se situe au niveau de la lourdeur administrative des universités. En 2004, j'ai fait le décompte du nombre de pages écrites dans l'année. Pour environ 200 pages scientifiques (articles, chapitres de livres, notes de cours) il y a eu environ 1000 pages administratives (rapports, budgets, réquisitions, bilans, justificatifs divers, lettres d'appui). Les chercheurs ne sont évalués que sur leur production scientifique, et les tentatives d'allègement de leur tâche administrative ne se traduisent souvent que par des formulaires additionnels, malgré toute la bonne volonté déployée par les administrateurs de la recherche.*

Which translates to:

> *For most researchers, the main obstacle to scientific productivity comes from the administrative burden within universities. In 2004, I counted the pages I wrote throughout the year. For about 200 scientific pages (articles, book chapters, course notes), there have been about 1000 administrative pages (reports, budgets, requisitions, balance sheets, supporting documents, reference letters). Researchers are only assessed on their scientific production, and attempts at reducing the administrative load often simply lead to additional forms, despite the good will of research administrators.*

For him to have used 102 precious application words simply to describe the problem of the system confirmed to me that something was off in science. At this point, I decided that, before becoming a researcher in any specialty I would try to solve this problem. This is how my journey began, from publication ethics to research integrity and eventually to research assessment.

For researchers like my father, for great thinkers who feel the need to compromise their creativity and their integrity to remain successful, for science, for knowledge… I feel lucky to be the kid who watches the planes take off. During the years of my doctoral studies, I have spoken with so many people who share the same worries and who want to make science better. I have seen such enthusiasm, passion, and mobilization for change. And I am confident that if enough of us get together and if enough pilots, cabin crew, and gate agents raise their voices we will change how things are done in the control tower.

General introduction

Science progresses by creating, accumulating, expanding, and refining knowledge. Yet science is also essentially human. It is built by individuals who have personalities, ambitions, and who aspire to improve our understanding of the world. As any human activity, science is thus subject to mistakes, inaccuracies, carelessness, and deception, all of which may threaten its integrity and have disastrous implications (e.g. Flaherty, 2011; Teixeira da Silva & Bornemann-Cimenti, 2017). Integrity and accuracy are thus paramount to scientific progress.

Current academia most often depicts research integrity as a responsibility of researchers. Codes of conduct delineate acceptable practices, mandatory trainings ensure that these practices are understood, oversight ensures compliance, and allegations ensure accountability. Consequently, researchers' knowledge, actions, and decisions are currently the pillars upon which research integrity is ensured. While it may be easiest to blame individual researchers for losing track and deviating from integrity, it is highly debatable whether the root of integrity failures can be reduced to scientists behaving badly. Indeed, intentional misconduct — although extensively mediatized in response to notorious scandals — is relatively rare and only seems to be the tip of the iceberg (Fanelli, 2009; Martinson, Anderson, & De Vries, 2005; Pupovac & Fanelli, 2014). Questionable (or detrimental[1]) research practices, on the other hand, are more prevalent (Fanelli, 2009) and are believed to have a greater cumulative impact on the scientific record (Bouter, Tijdink, Axelsen, Martinson, & ter Riet, 2016). These practices encompass the way in which the data is reported (e.g., selective reporting), the way in which it is published (e.g., salami slicing), the completeness and representation of the findings (e.g., undisclosed limitations, inaccessible data, incomplete methods, spin), the choice of methods and analysis (e.g., p-hacking, undocumented clearance of outliers), or simply openness, transparency, and collegiality. Although we cannot exclude the responsibility of researchers for misconduct and questionable research practices, past research has shown that it

[1] In 2017, the US National Academies of Sciences, Engineering, and Medicine issued a report entitled 'Fostering Integrity in Research' (NASEM, 2017). In the report, the term 'detrimental research practice' replaced the previously common term of 'questionable research practice' as a way to emphasise the detrimental effect that such practices can have on science. In the current thesis, I use the term 'questionable research practices' solely for simplicity and to avoid confusion.

is at best a partial explanation of what is going on, and that the causes of integrity breaches are most often highly multifactorial. Pressure to publish (e.g., Anderson, Ronning, De Vries, & Martinson, 2007; Fanelli, 2010; Singh & Guram, 2014; Tijdink, Verbeke, & Smulders, 2014; Wester, Willse, & Davis, 2010), temptations from inadequate incentives and conflicting interests (e.g., DuBois et al., 2013; Kaiser et al., 2012; Lundh, Krogsbøll, & Gøtzsche, 2012; Shrader-Frechette, 2011), and excessive competition (e.g., Anderson et al., 2007), for example, have all been described as possible threats for the integrity of science. For early career researchers, insecurity and imbalanced rivalry are added to the lot, threatening not only the science itself, but also researchers' wellbeing and mental health (Evans, Bira, Gastelum, Weiss, & Vanderford, 2018; Levecque, Anseel, De Beuckelaer, Van der Heyden, & Gisle, 2017; "The mental health of PhD," 2019; Pain, 2017; Powell, 2016). In turn, pressures, incentives, competition, and career instability are all related to the way in which researchers are assessed and promoted. To be successful in current academia, researchers must be excellent, productive, fast, impactful, and competitive. Yet, such professional demands conflict with the more traditional scientific demands for quality, honesty, transparency, rigorousness, and openness, and the conflict may force researchers to adapt their behaviours and compromise on the integrity to survive in academia.

Following such concerns, a growing dissatisfaction towards research assessments has emerged among scientists and research stakeholders. Most critics assert that current assessments over-rely on reductionistic metrics which are not fit for what they are used for (American Society for Cell Biology, 2013; Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015; Moher et al., 2019; Wilsdon et al., 2015), and that the scope and the areas assessed are too rigid and limited (Alperin et al., 2019; "A kinder research culture," 2019; Winker, 2017). Ill-fitted assessments are worrisome because they are known to influence researchers' decisions and practices. Performance-informed research assessments (e.g., publication counts) have been shown to influence publication patterns and to potentially displace the goal of researchers from producing good quality science to producing a high quantity of outputs (Butler, 2003; Moed, 2008). Yet, impact-informed research assessments (e.g., citation counts, impact factors, etc.) which were introduced to balance quantity with impact[2] also generate problems by increasing pressures and by inciting strategic responses, many of which can be categorized as questionable research practices (de Rijcke, Wouters, Rushforth, Franssen, & Hammarfelt, 2015; Gingras, 2016; Larivière & Sugimoto, 2018; Wouters, 2014).

Despite the growing concern and dissatisfaction with current research assessments, concrete changes remain slow and unassuming. Three reasons can help explain part of this resistance for change. First, despite knowing that current metrics are inadequate, introducing more fitting methods to assess researchers is difficult. This difficulty is in part due to a delay in the acceptability of new methods and a lingering endorsement of old ones, but in other part due to a lack of agreement on what we really want to measure. Indeed, despite the sophistication and the number of available metrics being in constant expansion, new methods are rarely implemented given a lack of understanding of what research assessments should really look at (Van Noorden, 2010). Second, success in science is also implicitly dependent on cultures and personal views. Even though metrics are a big part of research assessments, human input and peer review retain an important role in assessing research. Adding a human factor to research assessments is essential to contextualise and accurately interpret research metrics (see for e.g., Council of Canadian Academies, 2012; Hicks et al., 2015; Holtrop, 29 November 2018; Moher et al., 2019). Nonetheless, peer assessors are often asked to build their interpretation on broad and undefined concepts, such as 'excellence', 'innovation', and 'impact'. In failing to define these key terms, current research assessments are vulnerable to implicit biases and differential views (Hatch, 2019). A clearer understanding of the meaning of these concepts is needed to reduce the ambiguity and interpretability of research assessments while preserving the richness of qualitative input. A third reason for the lagging changes in research assessment can be attached to the fact that changing systems is strenuous and requires coordinated actions from multiple

[2] 'Impact' in research assessments is frequently conflated with research quality. High impact factor journals are often considered to be of higher quality than low impact factor journals, and citations are often believed to act as a stamp of approval from the research community which certifies the quality of the scientific work. We will see in upcoming chapters that this confusion is still frequent (Chapters 3 and 6 in particular), but I found important to specify that this thesis considers impact indicators to reflect the attention that a piece of work has received, not its quality (Sugimoto & Larivière, 2018).

actors. Research institutions, funders, publishers, policy makers, but also researchers and research students all need to be involved and to agree to take part in the change. Unfortunately, current dialogues and decisions are often driven by single actor groups or, at the very best, still forget important actors. Consequently, current recommendations may fail to take into account the obligations and constraints of forgotten actors and may thus lack realism. Obtaining a deep understanding of the issues and perspectives faced by all research actors involved is thus crucial to allow concrete changes to take place.

The current thesis is an effort to address these three drawbacks. Through our findings, we help understand the failures of current success indicators (Chapter 6) and question the profound motivations for assessing researchers (Chapter 3). We also shed light on the interpretations of different concepts of success in science and locate where incompatible interpretations jeopardize the integrity of science (Chapter 3, 4, and 5). Finally, we capture the perspectives of a broad array of research actors to uncover forgotten dependencies, pressures, and disagreements between actor groups (Chapter 4). In light of our overall findings, I then propose four recommendations for change which I believe could make a foundation for cultivating better integrity in science.

OUTLINE OF THE THESIS

The present thesis is built as a collection of manuscripts organised in different chapters. The chapters are all interconnected and reflect the progression of our project and its findings.

Analysis of existing literature. In order to capture the landscape of research on research integrity and to find the gaps that currently exist, the first step of our project consisted of a broad review of scientific literature in the field. Instead of conducting a traditional narrative review of the literature, we systematically retrieved ten years of scientific articles on research integrity and compared different features of each article. In doing so, we identified two blind spots we believed should be addressed in future research. These blind spots will become the basis of our research questions in the following chapters.

First, we found that when describing why misconduct happens, most empirical findings identify problems from the research system as threats to the integrity of science, while much fewer articles inculpate issues related to the awareness or the personality of researchers. Paradoxically, we found that most empirical works proposing or testing approaches to promote integrity focused on increasing researchers' awareness and compliance (e.g., training, guidelines, etc.) rather than on approaches that change how the scientific system works. In light of this first finding, we decided to focus our project on issues from within the system, most particularly on research success, reward systems, and research assessments, which embedded the big issues of pressures, incentives, and competition that were predominant in past literature.

Second, we found that empirical research on research integrity disproportionately targeted researchers and research students, but largely ignored other key actors within the research system, such as funders, science policy makers, institution leaders, etc. To counter this one-sided perspective, we involve an array of different research actors in our project. The literature analysis and its findings are presented in Chapter 1.

Interviews and focus groups. The second step of the project consisted of a series of interviews and focus groups with different research actors. Since our project used Flemish biomedical research as a sample of study, we added a short chapter to describe research and integrity in Flemish research and to allow readers to place our findings into context (Chapter 2). Our interviews and focus groups explored perspectives of success in research, threats to research integrity, and responsibilities towards research integrity. Given the breadth of answers and topics covered by our respondents, we separated our findings in three different chapters, namely in Chapters 3, 4, and 5. Chapter 3 and 4 are highly descriptive, with the former detailing our findings about success and research assessments, and the latter detailing our findings about the problems that plague the research system. Finally, Chapter 5 takes the findings to a higher level and looks at the place of the scientific community and the norms of disinterestedness and communalism in current science. For this chapter, we compared our qualitative data to descriptions that Robert K Merton and Warren O. Hagstrom shared almost half a century ago.

Survey on research assessment. After having synthesised and summarized the findings from our interviews and focus groups, we used the indicators of success that raised conflicting opinions to build a short survey. The survey looks at specific indicators of success in order to determine which indicators are most important to the advancement of science, which are most important for the personal satisfaction of researchers, and which are exaggeratedly important in advancing one's career without contributing to science and personal satisfaction. The survey was shared primarily with Flemish researchers in the end of 2019, and our findings are detailed in Chapter 6.

Recommendations for change. Since most chapters are intended to be published as standalone scientific publications, they each contain an introduction, a discussion, and a conclusion, sometimes feeling redundant but also limited in scope. In the final *Discussion* chapter, I reflect on our overall findings and formulate four recommendations I believe would help foster scientific integrity and healthy research climates.

PRACTICE WHAT WE PREACH

Meta-research (i.e., research on research) aims to understand how science works, how research systems function, and how researchers behave. Researchers who conduct meta-research feel the challenging situation of being both the subject and the object of study. With a thesis focusing on research integrity, I often asked myself which of my actions best embraced integrity. I experienced how difficult it can be to find the best way forward, and also often felt that acting in favour of integrity tangled up with advancing my professional career. Although like many researchers in my position I am yearning for a future career in the field, I still used this thesis as an occasion to commit to my personal convictions about research integrity. I will describe below how these convictions influenced my publications and reporting strategies, my description of authorship, and my data accessibility.

I want to stress that, although my unconventional choices are essentially personal decisions, I would not have been able to make them without external support. The funding by the Bijzondere Onderzoeksfonds (BOF [Special research funds]) of Hasselt University did not impose formal requirements which would force me to give up my convictions on research integrity. Instead, its open approach allowed me to express these convictions throughout my thesis, an opportunity which would not have been possible in many other funding schemes. My supervisor, Wim Pinxten, also allowed me to follow my convictions. Despite the potential consequences for his own track record, Wim not only accepted my unconventional choices, but also openly encouraged my activism. It is thus together with the supporting system around me that I decided on the following points:

Publication choices. At first glance, it may seem odd that most of the chapters are organised like separate manuscripts and the thesis looks like a paper-based dissertation, while most of these chapters are still unpublished at the time of submission. In our project, the different chapters profoundly influenced each other, feeding forward and backwards to one another. To enable this important feature, we decided to write the papers as parallel and not consecutive efforts. Consequently, most manuscripts were ready for submission only by the time the full thesis was finalized.

Given the importance of both publishing and peer-review, the papers will obviously also be submitted for publication. Chapters 3 and 4 will be submitted as preprint and for publication simultaneously with the submission of this thesis, and Chapters 5 and 6 will be submitted shortly after[3]. In the choice of journals, we will not prioritize journals with high impact factors, but rather favour audience, openness, and accessibility.

[3] By the time of printing this thesis, Chapters 3 and 4 have now undergone review at *Research Integrity and Peer Review*, while Chapter 6 is under review at *PLOS One*.

Writing style. In the current thesis, I also strived to increase transparency and reproducibility. I wanted to ensure that our methods and potential influence on the views expressed were available to the readers and, for qualitative parts, I wanted to provide access to numerous quotes so that readers get a genuine opportunity to endorse or challenge our interpretations. This results in papers which are extensive and long. To increase accessibility of our findings in a period where research time is so limited, we have already agreed with a few organisations to provide summaries and short communications to summarise our work and bring it to the attention of key stakeholders for systemic change.

Investing beyond outputs. Another decision of the current thesis was to invest not only in the measurable outputs, but also in activities which are often overlooked by current assessments (as we will see in Chapter 3 and 6). As a result, we decided to invest in teaching and in activism by attending key events and by communicating about research integrity outside the experts' circle. At the end of the thesis, I include a section which describes teaching activities, conference presentations, and popular communication which resulted from the present thesis (section entitled 'Additional outcomes from the current thesis').
As I will mention in the recommendations (see 'General discussion'), I believe that research on research integrity needs to extend outside its expert circle to reach those who are least aware and concerned about the issues of the current system. We found important to invest in these activities to promote discussions about the implication of our findings and to increase awareness on the topics with actors who are often underexposed.

Contributorship and authorship. In the current thesis, I also adopted contributorship as an alternative for traditional conceptions of authorship to recognize the collaborative efforts of all involved in the creation of research results and the report thereof. We will see in Chapters 3 and 4 that authorship is a frequent cause of dispute among researchers and that it often triggers undesirable competition and individualism in research (see Chapter 5). Indeed, the concept of authorship fails to do justice to modern, highly collaborative science. Consequently, I detail the specific contributions which permitted each chapter using the contributor roles taxonomy (CRediT) proposed by the Consortia

Advancing Standards in Research Administration Information (CASRAI). The definition of each contributor roles is available in Appendix 13.

It is important to understand that contributors are distinct from authors and that their responsibility differs accordingly. The International Committee of Medical Journal Editors (ICMJE) defines authors as contributors who have provided:

> "• *Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND*
> • *Drafting the work or revising it critically for important intellectual content; AND*
> • *Final approval of the version to be published; AND*
> • *Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved." (International Committee of Medical Journal Editors (ICMJE))*

In line with this perspective, to be considered authors, contributors would need to (i) have contributed substantially to the 'conceptualization' or the 'investigation', 'formal analysis', or 'validation' of data for the work; (ii) have contributed to 'writing – original draft' or 'writing – review & editing' in a manner that was critical for important intellectual content; (iii) to have approved the final version of the manuscript; and (iv) to agree to be accountable for all aspects of the work.

In the current state of this thesis, only Wim Pinxten and I (Noémie Aubert Bonn) would fulfill the criteria for authorship of most chapters, and therefore only we are accountable and responsible for any issue in the present work and its findings.

Data availability. Finally, in order to increase transparency and encourage reuse, I provide, where possible, links to our full data. The data files will remain available in our Open Science Framework registration at https://osf.io/ap4kn/ within the folder entitled 'Thesis Online Material'. I do not provide full transcripts for interviews and focus groups to avoid jeopardizing confidentiality (Chapters 3, 4, and 5) but I included numerous quotes which I extend when possible in tables and appendices.

REFERENCES

Alperin, J. P., Muñoz Nieves, C., Schimanski, L. A., Fischman, G. E., Niles, M. T., & McKiernan, E. C. (2019). How significant are the public dimensions of faculty work in review, promotion and tenure documents? *eLife, 8*, e42254. doi:10.7554/eLife.42254

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics, 13*(4), 437-461. doi:10.1007/s11948-007-9042-5

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review, 1*(1), 17. doi:10.1186/s41073-016-0024-5

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation, 12*(1), 39-46. doi:10.3152/147154403781776780

CASRAI. *CRediT – Contributor Roles Taxonomy*. Retrieved from https://casrai.org/credit/

Council of Canadian Academies. (2012). *The Expert Panel on Science Performance and Research Funding.* [ISBN: 978-1-926558-42-4]. Retrieved from https://cca-reports.ca/wp-content/uploads/2018/10/scienceperformance_fullreport_en_web.pdf

de Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2015). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation, 25*(2), 161-169. doi:10.1093/reseval/rvv038

DuBois, J. M., Anderson, E. E., Chibnall, J., Carroll, K., Gibb, T., Ogbuka, C., & Rubbelke, T. (2013). Understanding Research Misconduct: A Comparative Analysis of 120 Cases of Professional Wrongdoing. *Accountability in Research, 20*(5-6), 320-338. doi:10.1080/08989621.2013.822248

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*, 282. doi:10.1038/nbt.4089

Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE, 4*(5), e5738 (5731–5711). doi:10.1371/journal.pone.0005738

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE, 5*(4), e10271. doi:10.1371/journal.pone.0010271

Flaherty, D. K. (2011). The vaccine-autism connection: A public health crisis caused by unethical medical practices and fraudulent science. *Annals of Pharmacotherapy, 45*(10), 1302-1304. doi:10.1345/aph.1Q318

Gingras, Y. (2016). *Bibliometrics and Research Evaluation: Use and Abuses*. Cambridge (MA): MIT Press.

Hatch, A. (2019). To fix research assessment, swap slogans for definitions. *Nature, 576*(9), 9. doi:10.1038/d41586-019-03696-w

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature News, 520*, 429-431. doi:10.1038/520429a

Holtrop, T. (29 November 2018). The evaluative inquiry: a new approach to research evaluation. *LSE Impact Blog*. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2018/11/29/the-evaluative-inquiry-a-new-approach-to-research-evaluation/

International Committee of Medical Journal Editors (ICMJE). *Defining the Role of Authors and Contributors*. Retrieved from http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two

Kaiser, K. A., Cofield, S. S., Fontaine, K. R., Glasser, S. P., Thabane, L., Chu, R., . . . Allison, D. B. (2012). Is funding source related to study reporting quality in obesity or nutrition randomized control trials in top-tier medical journals. *International Journal of Obesity, 36*(7), 977-981. doi:10.1038/ijo.2011.207

A kinder research culture is possible. (2019). *Nature, 574*, 5–6. doi:10.1038/d41586-019-02951-4

Larivière, V., & Sugimoto, C. R. (2018). The Journal Impact Factor: A brief history, critique, and discussion of adverse effects. *arxiv*. doi:arXiv:1801.08992v2

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868-879. doi:10.1016/j.respol.2017.02.008

Lundh, A., Krogsbøll, L. T., & Gøtzsche, P. C. (2012). Sponsors' participation in conduct and reporting of industry trials: a descriptive study. *Trials, 13*, 146. doi:10.1186/1745-6215-13-146

Martinson, B. C., Anderson, M. S., & De Vries, R. (2005). Scientists behaving badly. *Nature, 435*(7043), 737-738. doi:10.1038/435737a

The mental health of PhD researchers demands urgent attention. (2019). *Nature, 575*, 257-258. doi:10.1038/d41586-019-03489-1

Moed, H. F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity? *Scientometrics, 74*(1), 153-161. doi:10.1007/s11192-008-0108-1

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2019). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *OSF Preprints*. doi:10.31219/osf.io/m9abx

NASEM. (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press.

Pain, E. (2017, April 4th). Ph.D. students face significant mental health challenges. *Science*. Retrieved from https://www.sciencemag.org/careers/2017/04/phd-students-face-significant-mental-health-challenges. doi:10.1126/science.caredit.a1700028

Powell, K. (2016). Young, talented and fed-up: scientists tell their stories. *Nature, 538*, 446-449. doi:10.1038/538446a

Pupovac, V., & Fanelli, D. (2014). Scientists Admitting to Plagiarism: A Meta-analysis of Surveys. *Science and Engineering Ethics, 21*(5), 1331-1352. doi:10.1007/s11948-014-9600-6

Shrader-Frechette, K. (2011). Climate Change, Nuclear Economics, and Conflicts of Interest. *Science and Engineering Ethics, 17*(1), 75-107. doi:10.1007/s11948-009-9181-y

Singh, H. P., & Guram, N. (2014). Knowledge and attitude of dental professionals of North India toward plagiarism. *North American Journal of Medical Sciences, 6*(1), 6-11. doi:10.4103/1947-2714.125854

Sugimoto, C. R., & Larivière, V. (2018). *Measuring research What everyone needs to know*. New York (NY): Oxford University Press.

Teixeira da Silva, J. A., & Bornemann-Cimenti, H. (2017). Why do some retracted papers continue to be cited? *Scientometrics, 110*(1), 365-370. doi:10.1007/s11192-016-2178-9

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication Pressure and Scientific Misconduct in Medical Scientists. *Journal of Empirical Research on Human Research Ethics, 9*(5), 64-71. doi:10.1177/1556264614552421

Van Noorden, R. (2010). Metrics: A profusion of measures. *Nature, 465*(7300), 864-866. doi:10.1038/465864a

Wester, K., Willse, J., & Davis, M. (2010). Psychological climate, stress, and research integrity among research counselor educators: A preliminary study. *Counselor Education and Supervision, 50*(1), 39-55. doi:10.1002/j.1556-6978.2010.tb00107.x

Wilsdon, J., Liz Allen, Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from https://re.ukri.org/documents/hefce-documents/metric-tide-2015-pdf/

Winker, K. (2017). Eyeballs on science: Impact is not just citations, but how big is readership? *bioRxiv*. doi:10.1101/136689

Wouters, P. (2014). The Citation: From Culture to Infrastructure. In B. Cronin & C. R. Sugimoto (Eds.), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (pp. 47-66). Cambridge (MA): The MIT Press.

Chapter 1

A decade of empirical research on research integrity: what have we (not) looked at?

An adaptation of this chapter is currently published as:

Aubert Bonn, N., & Pinxten, W. (2019). A Decade of Empirical Research on Research Integrity: What Have We (Not) Looked At? *Journal of Empirical Research on Human Research Ethics, 14*(4), 338-352. doi:10.1177/1556264619858534

CONTRIBUTIONS

Conceptualization: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries
Funding acquisition: Wim Pinxten. Funding granted by the Bijzonder Onderzoeksfonds (BOF) 15NI05
Project administration: Noémie Aubert Bonn, Wim Pinxten
Methodology: Noémie Aubert Bonn
Investigation: Noémie Aubert Bonn
Data curation: Noémie Aubert Bonn
Formal analysis: Noémie Aubert Bonn
Visualization: Noémie Aubert Bonn
Validation: Noémie Aubert Bonn, Wim Pinxten; Additional thanks to the organisers of the Doctoral Forum of the 5th World Conference of Research Integrity: Nicholas H. Steneck, Elizabeth Heitman, and Nils Holger Axelsen as well as its participants for their comments and recommendations regarding this work.
Supervision: Wim Pinxten
Writing – original draft: Noémie Aubert Bonn
Writing – review & editing: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries[1]; additional thanks to anonymous peer-reviewers from *PLOS One* and *Science and Engineering Ethics* for their useful comments which helped improve earlier version of this manuscript.

1. Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor (MI), USA

KEYWORDS

research integrity, research misconduct, questionable research practices, detrimental research practices, meta-research, research fraud

ABSTRACT

Research on research integrity has become a field of its own, yet a comprehensive overview of the field is still missing. We systematically searched SCOPUS, Web of Science, and PubMed for relevant articles published between 2005 and 2015. We extracted the topic, methodology, focus, and citations from each article. From the 986 articles included, only 342 report empirical data. Empirical papers predominantly targeted researchers and students. Although empirical articles questioning causes for misconduct mostly blamed research systems (e.g., pressure, competition) for detrimental research practices, articles proposing approaches to foster integrity focused on researchers' awareness and compliance rather than on system changes. Involving non-researchers and reconnecting what is known to what is proposed may help research on research integrity move forward.

INTRODUCTION

Research integrity (RI) has been part of the scientific discourse for many years and has evolved to a topic of research itself over the past 20 years. Research on RI highlighted that research misconduct comes in many forms (De Vries, Anderson, & Martinson, 2006), occurs more often than was initially thought, and that questionable research practices (QRP) — also referred to as *detrimental research practices*, practices outside the realm of misconduct which still risk damaging the scientific output — are far from rare (Fanelli, 2015; Pupovac & Fanelli, 2014).

In 1999, one of the first paper setting the agenda for research on RI concluded with the following words:

> *Over the last decade, researchers and research institution have made significant strides toward restoring [trust in science] by actively confronting misconduct. […] With so much accomplished, the time is right to see whether the policies we have put in place, the funds and time spent, have made a difference. Have we achieved levels of integrity in research that are acceptable? (Steneck, 1999, p. 173)*

Now nearly two decades later, this call for research on RI seems to have been heard. Scientific literature on RI and research misconduct increased exponentially, broad scale funding and consortiums have been established to enable more research on the topic (e.g., the European Commission Horizon 2020 contributed well above 20 million euros in projects on RI since 2015), attendance to the last *World Conference on Research Integrity* exceeded 900 participants, and some institutions are starting to build departments with PhD students specializing on the topic.

Notwithstanding this growing interest for research on RI and misconduct, it is unclear how the potential to identify and quantify the problems, to highlight and understand determinants of bad science, and to assess and propose approaches that foster integrity and prevent misconduct have been employed. To provide better insights in the field, we analysed published research on RI. The goal of this analysis was twofold. On the one hand, we aimed to understand how researchers focusing on RI perform research (i.e.: which methods are used, which stakeholders are studied, and which topics are most investigated). On the other hand, we aimed to document gaps of knowledge to inform future research endeavours.

METHODS

Studying research on RI is methodologically challenging. Researchers from many different fields address the topic in different ways. There is poor consistency in how the scope of RI is delimited (e.g., Is research ethics part of integrity? Is academic integrity only targeting students?) and in the choice of journals or article formats. For example, the empirical piece of Brian Martinson and colleagues (2005) — widely recognised as a cornerstone in research on RI — was published as a 'Commentary' in Nature, is currently being classified as a 'Note' in Scopus and as 'Editorial material' in the Web of Science. Consequently, systematic searches for relevant empirical works on research integrity have serious blind spots if the sample is kept manageable.

We are aware that, despite all efforts to gather a manageable sample of the highest possible relevance, the choices we made towards our search strategy unavoidably come at a cost (e.g., not including the Martinson et al. paper, and unavoidably several other important pieces of research). We ensured that such costs are transparently reflected throughout this paper.

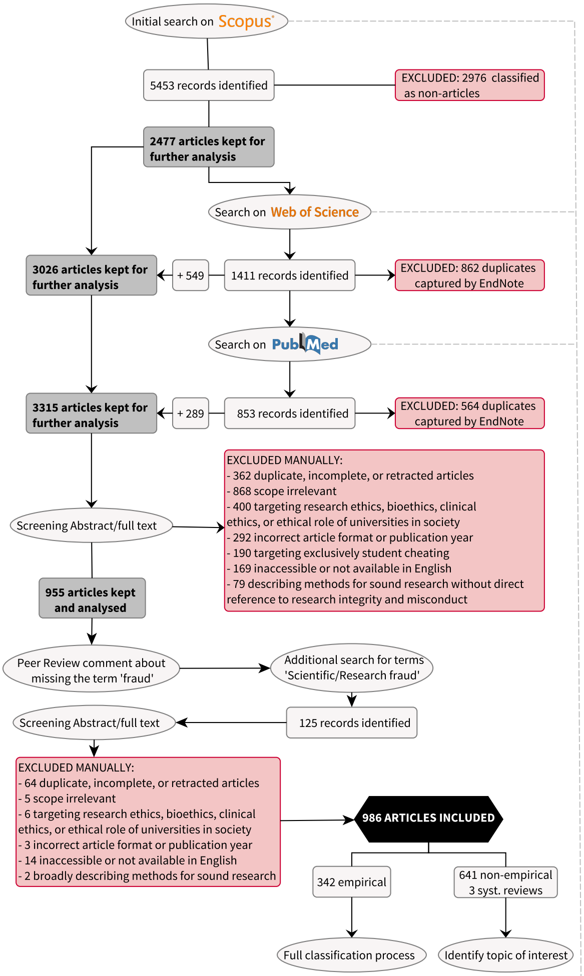
To characterize the broad spectrum of research on research integrity, we performed an analysis of the literature on RI published in English between 2005 and 2015. Our analysis differs from a typical literature review: we classified several variables beyond the findings of the included articles (e.g., publication year, impact metrics, geographical distribution, and several methodology characteristics) and analyzed the relationship between such variables. Consequently, our findings do not describe what is known about specific aspects of RI, but rather provide an overview of how research on RI is performed and published in order to highlight the areas or actors where/with whom most research is done and areas or actors current research might have overlooked.

We used three major bibliographic databases to find relevant literature on RI: SCOPUS, Web of Science, and PubMed. We performed and adapted our search between February and April 2017 for SCOPUS, between October and November 2017 for Web of Science and PubMed, and in February 2019 to add the terms ‘scientific fraud’ and ‘research fraud’ as recommended by peer-reviewers. We extracted all results in an Excel sheet, which is available as tab delimited in the folder ‘Thesis Online Material’ in our OSF Registration (https://osf.io/ap4kn/). We only kept the records present on the sheet for further analyses. The complete study flow diagram with inclusion/exclusion counts and search queries may be seen in Figure 1.

In summary, our queries screened the titles, abstracts, and keywords of published literature for any mention of 'academic misconduct', 'academic integrity', 'research misconduct', 'research integrity' (or any expression of six words or less containing such terms), 'responsible research' (or any expression of four words or less containing these two terms), or the exact expressions ‘scientific integrity’, ‘scientific misconduct’, ‘scientific fraud’, and ‘research fraud’. We chose these keywords after a few adaptations as we believed that they would provide a broad and yet specific enough overview of works that have been published on RI. Having worked in the field of RI in non-English speaking countries for some time, we purposively included the expression ‘academic misconduct’ despite its more direct relationship to student cheating to allow capturing articles which might have used the term differently to refer to research misconduct (Aubert Bonn, Godecharle, & Dierickx, 2017).

We did not include papers relating to the ethical care of animals in research. Beyond papers whose scope was directly irrelevant, we also excluded several themes which were related, but not directly linked to RI, namely (i) academic integrity or cheating limited to undergraduate students, or with no apparent extension to RI in the discussion and the abstract of the paper (ii) research ethics looking at the protection of human participants; (iii) clinical ethics, bioethics, or responsible research innovations focusing on societal concerns of research discovery; and (iv) techniques meant to improve the validity of research, but devoid of direct reference to QRP, misconduct, or integrity. In our findings, we respectively refer to those exclusions as (i) *Cheating (exclusively)*, (ii) *Research ethics*; *(iii) bioethics, clinical ethics, or ethical role of universities in society*; and (iv) *Methods and tools*.

Figure 1. Study Flow Diagram



SEARCH TERMS:

Scopus

(TITLE-ABS-KEY ((academic OR research) W/4 (misconduct OR integrity)) OR TITLE-ABS-KEY ((responsible W/2 research)) OR TITLE-ABS-KEY ("Scientific fraud' OR "Research fraud") OR KEY ("Scientific misconduct") OR KEY ("Scientific integrity")) AND PUBYEAR > 2004 AND PUBYEAR < 2016

Web of Science

(TS=((academic OR research) NEAR/4 (misconduct OR integrity)) OR TS=((responsible NEAR/2 research)) OR TS=("Scientific misconduct") OR TS=("Scientific integrity") OR TS=("Scientific fraud") OR TS=("Research fraud")

Document types: ARTICLE; Timespan: 2005-2015.

PubMed

(((((research integrity[Text Word]) OR research misconduct[Text Word]) OR academic integrity[Text Word]) OR academic misconduct[Text Word]) OR (responsible research[Text Word]) OR (("scientific misconduct"[Other Term]) OR ("scientific fraud"[Other Term]) OR ("research fraud "[Other Term]) OR ("scientific integrity"[Other Term]) AND ("2005"[Date - Publication] : "2015"[Date - Publication]))

Article types: Clinical Trial/Journal Article

CLASSIFICATION PROCESS

To build the classifications for our research, we used an inductive process based on the findings from the first set of papers retrieved (i.e., the SCOPUS search). An inductive process means that we started with the general goal of *describing research*, and that we decided on which categories and classification options we should include based on what we found in the abstracts and papers assessed. For this analysis, NAB built the search, retrieved the literature, selected articles to be included, and inductively classified the articles in categories. WP helped refine and simplify the categories, revised individual papers which were ambiguous, and provided assistance on the specific wording used for the categories.

A full description of the inductive process that led to our final categories is available in Appendix 1. The final categories and classification options and their definitions are listed in Table 1.

Except for the 'determinants' and the 'approaches' subcategories which were weighed, each article was fitted only once in each category. In case of ambiguity, we revised the papers further to determine what the authors highlighted most in the title and abstract, and we decided the classification based on their emphasis. For example, if a paper looked at guidelines and policies for plagiarism, the paper would obviously be a good fit for both topic groups of 'guidelines and policy', and of 'plagiarism'. In such a case, we decided according to the terminology used by the authors in the abstract and title. We did not assess the quality of included literature.

We first classified all relevant papers according to their *topic of interest*. We then used the abstract and full text to determine whether the article was empirical or not. For the purpose of our analysis, we considered anything that included a minimal description of data collection and analysis, from qualitative research to bibliometric studies or textual analyses, as 'empirical'.

We further classified each empirical article according to (i) the general methodology; (ii) the studied population; (iii) the source of data collection; (iv) the focus of interest; and (iv) the research objective.

In addition, for papers in which the focus of interest was 'determinants' of misconduct and QRP, we extracted the specific determinants found in the empirical work and classified them between *personal* issues, *systemic* issues, or

Table 1. Classification categories found inductively and used in our analyses

TOPIC OF INTEREST	TYPE OF RESEARCH	GENERAL METHODOLOGY	STUDIED POPULATION	SOURCE OF DATA	FOCUS	OBJECTIVE
• QRP and misconduct (general) • RI (general) • RCR training, education and mentorship • Cheating and academic misconduct • Publication ethics • Authorship and collaborations • Plagiarism • Conflicts of interest • Guidelines and policies • Peer Review • Research infrastructures and environments • REC/IRB • Allegations, sanctions, disclosure of cases • Whistleblowing • CV and application misrepresentation • Research on RI	• Empirical	• Surveys, Interviews or focus groups • Content and textual analysis • Bibliometric study • Investigation or forensic analysis • Tool building or validation • Combined methods • Other	• Researchers • Students • Mentor, RCR Instructor • Institution • Industry • Editors • Participants, Public, Media • Peer-reviewers • Policy makers • REC or IRB • Clinical professional • Research integrity officer • General	• Survey or assessment • Interviews • Focus group/expert panel • Published material • Retractions, notes, or bibliometric data • Industry document • Guidelines, policies, or university requirements • Teaching material • CV and application forms • Case reports • Combined sources • Other	• Determinants - System - Personal - Awareness and compliance • Problem/State of affair • Approaches - System - Awareness and compliance • Consequences • Research on RI	• Describe, explore, or quantify • Test a hypothesis • Denounce or detect QRP and misconduct • Build approach • Assess approach efficacy • Build or validate research tool
	• Non Empirical • Systematic reviews and meta analyses	—	—	—	—	—
To identify the *topic of interest*, we looked more particularly at the title and abstract to highlight what the authors appeared most interested in discussing. There could be overlap or different interpretations in this category, but it provides a general overviews of the themes most addressed in the field.	We classified included papers according to whether or not they reported empirical work, as defined earlier. A few systematic reviews were also found but we did not include them in further classification processes.	This category shows the general methodology used in the investigated paper. Naturally, it will be tightly linked with the *source of data*	Here we show who participated or was studied in the work. Depending on the *source of data*, these actors might not have been investigated directly, but by a proxy (e.g., Looking at researchers' misconduct through retractions	In the source of data, we showcase the type of data that was gathered to perform the research by showing from which source it was acquired.	Here, we identified which particular step of the integrity problem is studied (i.e., the determinants or causes for misconduct, the problem itself, approaches to fight the problem, consequences of the problem, or research on RI).	For each focus, we described what we considered appeared to be the general objective of researchers.

A more detailed description of the classification process and how we developed the categories inductively is available in Appendix 1. For each relevant paper, we first noted the topic of interest. We then identified the paper as either empirical or not empirical. For empirical papers, we further identified the general methodology used, the studied population, the source where the data came from, the focus on research integrity issues, and the general research objective.

issues related to *researcher's awareness and compliance*. Likewise, in papers in which the *focus of interest* was 'approaches' to misconduct and QRP, we classified the approach as either targeting the *system*, or targeting *researchers' awareness or compliance* (note: we luckily did not find any approaches that proposed to change personal characteristics such as gender or personality, so we did not include *personal* approaches).

After completing the classification, we analyzed our data in Excel to observe ongoing trends. We then used the data visualization program Tableau Software 10.4 to build figures that illustrate our findings. Given the inductive approach and the lack of predefined hypotheses from our review, we do not include expression of precisions such as confidence intervals and p values in our findings.

DATA AVAILABILITY

The full dataset, with both included and excluded records and full classification categories, as well as the data on determinants and approaches are available in the 'Thesis Online Material' folder in our Open Science Framework registration at osf.io/ap4kn/ in the files 'Chapter 1 – 2. General Data', and 'Chapter 1 – 3. Determinants and Approaches', respectively.

STUDY LIMITATIONS AND OTHER CONSIDERATIONS

Given the current lack of a comprehensive review in the field of RI, we consider our work to be a first step to expose what has been done and how it has been done in research on RI. That being said, as in any research project, several limitations were inevitable to allow us to manage the amount of data gathered with the resources at hand.

First, we decided on a cut off of 2005–2015 to grasp the bulk of research on RI that happened after the impactful Nature paper *Scientists behaving badly* (Martinson et al., 2005), widely recognized as a milestone in the field. As this review was the first step of a bigger project, we had to set a cut off to achieve a realistic record sample. Starting this extractions in 2016, we chose not to include literature published after 2015 since it might not be fully archived on databases at the time where we performed the search. We invite follow ups on our study as it would be very interesting to see what has happened in the most recent years.

Furthermore, we limited our search to records classified as ‘articles’ to obtain a more manageable and relevant subset of research to include in our analyses. Although we are aware that this automatic classification is not flawless (i.e., it sometimes includes editorials, news pieces, etc., and it might overlook a few research articles), we considered this automatic classification to be the best way to obtain a manageable sample of papers in which the bulk of empirical research on RI should be present. During the manual screening of the papers, we further excluded papers that were evidently not ‘articles’ (e.g., labeled editorials, labeled news reports, short conference abstracts, and letters to the editor). Nonetheless, to avoid biasing our inclusions to the terminology used by journals to distinguish article categories, and because we noticed that empirical data were sometimes reported in differently labeled records, we kept other papers with a more substantial format (e.g., opinions, commentary, viewpoint, ethics corner, correspondence, etc.) when they were automatically classified under the ‘article’ category.

In light of the two former points, and added to the fact that we lacked a reference point, it was difficult to evaluate the completeness of our sample and the sensitivity of our search strategy. Our findings should thus not be considered in isolation of the methods we have used (e.g., search terms included and not included, the way we defined integrity for the purpose of this research, etc.) and choices we have made to reach a manageable sample of papers (e.g., document type, years included, etc.).

It is also essential to note that a certain level of subjectivity cannot be fully excluded from the classification of the included papers. For example, when looking at the topic of interest, many papers could fit in several topics — a paper on ghost authorship with the pharmaceutical industry would inevitably fit into ‘authorship’, but also concern ‘conflicts of interest’, ‘QRP and misconduct’, ‘reporting and publishing’, and so forth. We were careful to select the categories and classifications we considered most appropriate based on what was highlighted by the authors in the abstract. Although the classification process was not triangulated by individual reviewers, uncertainties were marked and discussed between authors until a common agreement could be reached. Oftentimes, we reflected upon, corrected, and revisited our categories to strengthen the fit, but we did so without consideration of trends or hypotheses.

Classifications in one category were also often linked to classifications in another category. For example, papers on 'RCR training and mentoring' will often involve 'approaches' to deter misconduct, have the objective to 'assess' a method, and study researchers, students, or RCR educators. We tried to remain as neutral as possible when classifying our articles by building our classification from the content of the paper rather than from expected trends. Nonetheless, we believe that our results should not be considered in isolation but as a whole in which each category may intertwine with another.

Finally, because we decided not to assess the quality of included papers, we included a wide range of journals and paper standards. Within our inclusions, ten articles were published in journals present on Beall's list of predatory publishers (note however that five records come from the publisher Frontiers, whose status as predatory publisher is now mostly refuted). Given that Beall estimates that predatory publications accounts for 5-10% of all open access articles (Butler, 2013), ten papers in 986 is a small proportion. Nonetheless, given the topic of our review, the fact that not all included articles were open access, and the fact that we conducted our search using databases which already screen for journal quality, we considered that one percent worthy of mention.

RESULTS

INCLUSIONS AND EXCLUSIONS

After screening titles and abstracts for relevance to the topic, we included 986 articles. Table 2 highlights the number of inclusions and manual exclusions (i.e., manually excluded after the initial Excel sheet has been compiled). The complete dataset, with both included and excluded papers and full classification categories, is available in the in our Open Science Framework registration at osf.io/ap4kn/ under 'Chapter 1 – 2. General Data'.

Within the 5453 publications yielded by the initial search in SCOPUS (i.e., our first search strategy, see Figure 1), 2477 records (44.4%) were classified as ‘articles’.

EMPIRICAL COVERAGE

Around a third of the included publications described empirical work (n=342; see Table 2). Within our inclusions, theoretical approaches, narrative reviews, recommendations, and opinions were most common (all of which are classified as non-empirical hereafter).

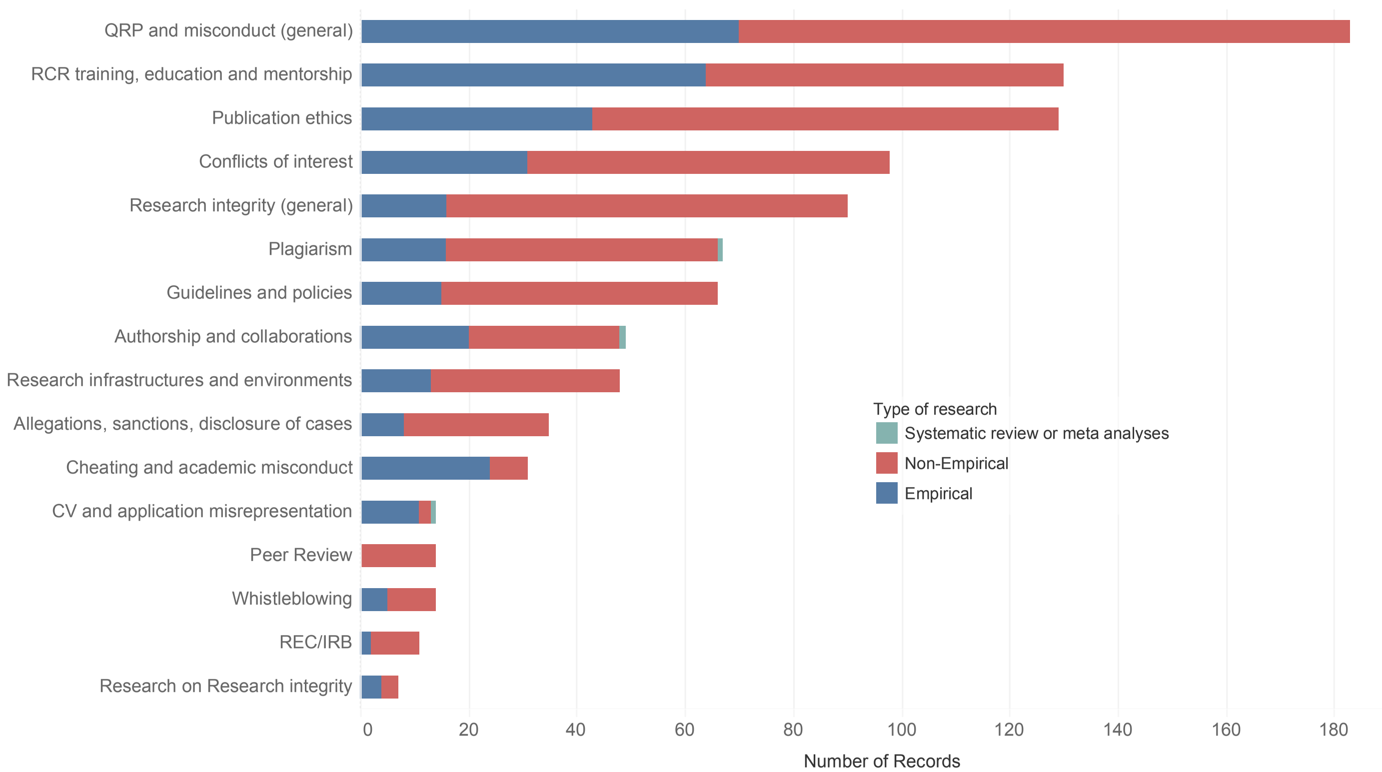
Table 2. Number of included and excluded records

	INCLUSIONS	EXCLUSIONS	REASON FOR EXCLUSION
Total	**986**	**2454**	
Empirical	342		
Non Empirical	621		
Systematic review or meta analysis	3		
		183	Accessibility and language
		295	Article format or year
		190	Cheating (exclusively)
		426	Duplicate/incomplete/retracted
		81	Methods and tools
		406	Research ethics, bioethics, clinical ethics, or ethical role of universities in society
		873	Scope irrelevant

TOPICS OF INTEREST

We extracted the topics of interest of all included papers and grouped them in categories. When papers were not clearly targeting a specific topic, we classified them in the more general categories of ‘QRP and misconduct’ or ‘Research integrity’, accordingly. Most papers targeted ‘QRP and misconduct’, but a substantial proportion of papers also targeted ‘RCR training, education, and mentoring’; ‘Publication ethics’; and ‘Conflicts of interests’ (Figure 2). The proportion of empirical articles was highest for topics of ‘Cheating and academic misconduct’ (77%), ‘CV and application misrepresentation’ (79%), 'Research on research integrity' (57%), and ‘RCR training, education, and mentoring’ (49%).

Figure 2. Topics of interest of included papers and corresponding research type



Most papers targeted 'Questionable research practices (QRP) and misconduct', but a substantial proportion of papers also targeted, 'Publication ethics', 'RCR training, education, and mentoring', and 'Conflicts of interests'. The proportion of empirical articles was higher for certain topics, for example for 'Cheating and academic misconduct' and for 'RCR training, education, and mentoring'

METHODOLOGIES

Over half of empirical papers used direct approaches, such as surveys, questionnaires, interviews, and focus groups (n=175) to obtain their data. Bibliometric studies (n=58) and content and textual analyses (e.g., policy documents, case studies; n=50) were also frequent. The distribution of methodologies alongside more specific research objectives can be seen in Table 3.

Table 3. Distribution of methodologies alongside more specific research objectives of empirical papers

GENERAL METHODOLOGY	OBJECTIVE	NUMBER OF RECORDS
Surveys, interviews or focus groups	Describe, explore, or quantify	113
	Assess approach efficacy	41
	Test a hypothesis	18
	Build or validate research tool	1
	Build approach	2
	Total	**175**
Content and textual analysis	Describe, explore, or quantify	35
	Assess approach efficacy	11
	Test a hypothesis	2
	Denounce or detect misconduct and QRP	1
	Build approach	1
	Total	**50**
Bibliometric study	Describe, explore, or quantify	40
	Assess approach efficacy	1
	Test a hypothesis	14
	Denounce of detect misconduct and QRP	2
	Build approach	1
	Total	**58**
Investigation or forensic analysis	Describe, explore, or quantify	18
	Test a hypothesis	4
	Denounce or detect misconduct and QRP	8
	Total	**30**
Tool building or validation	Describe, explore, or quantify	6
	Total	**6**
Other	Describe, explore, or quantify	3
	Assess approach efficacy	3
	Test a hypothesis	1
	Denounce or detect misconduct and QRP	2
	Build or validate research tool	1
	Build approach	4
	Total	**14**
Combined methods	Describe, explore, or quantify	8
	Test a hypothesis	1
	Total	**9**
	Total number empirical papers	**342**

About half of all included empirical papers used direct approaches to describe or quantify issues related to research integrity. A fair proportion also used content and textual analyses and bibliometric studies, also mostly to describe or quantify integrity issues

STUDIED POPULATION

Over 60% of empirical papers study researchers and students, while fewer articles involved actors other than researchers (Figure 3). Researchers and students further account for over 75% of articles that used *direct approaches* — approaches in which investigators directly addressed the studied population, such as interviews, survey, focus groups, and direct observation. Other research actors were most often studied by proxy through documents, reports, or published material.

FOCUS

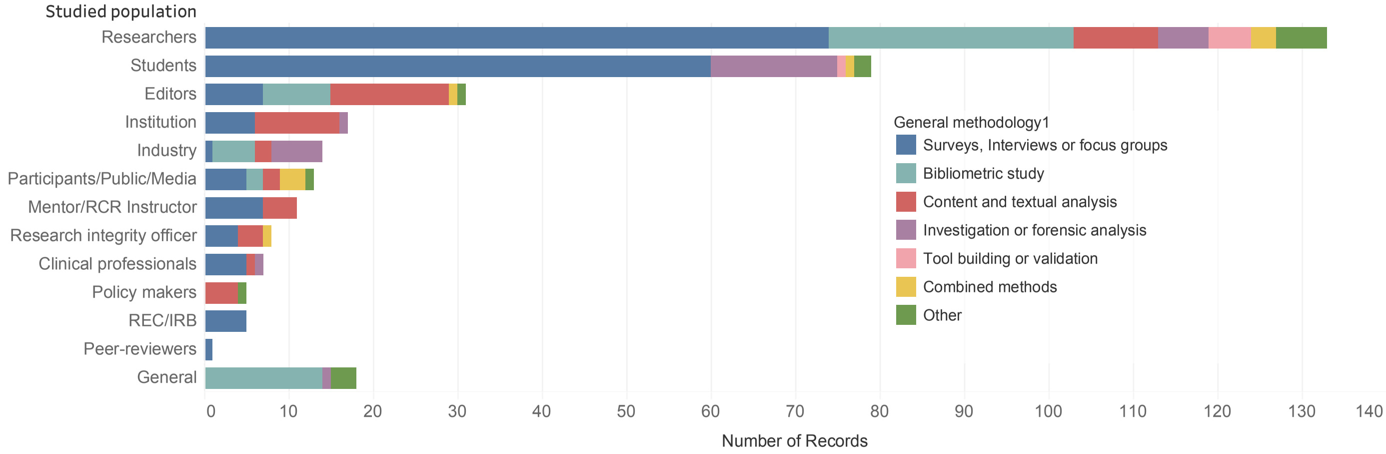
To map the most studied aspects of RI, we classified all empirical papers according to their focus on the RI issue (i.e., which particular step of the integrity problem they looked at). We defined five general focuses, namely, (i) the 'determinants' of misconduct, (ii) the 'problem/state of affair' of issues of research integrity, (iii) the 'approaches' meant to deter misconduct and promote integrity, (iv) the 'consequences' of misconduct and QRP, and (iv) tools and approaches specific to 'research on RI'. We then further classified the specific research objective we could grasp from the methodology of the paper (see Table 1).

Figure 4 shows that over 45% of empirical work on RI focused on the problem, generally with the objective to describe, quantify, or explore the issue. About a third of the articles focused on approaches to promote RI or deter misconduct, with over half of those assessing the efficacy of an approach. Only 13% of the papers focused on determinants of research misconduct and QRP, generally attempting to test the relationship between a hypothesized determinant and reported misconduct or QRP. Finally, very few articles focused on the consequences of misconduct and QRP (e.g., loss of public trust, risks to research participants, financial waste), or on research on RI.

DETERMINANTS AND APPROACHES

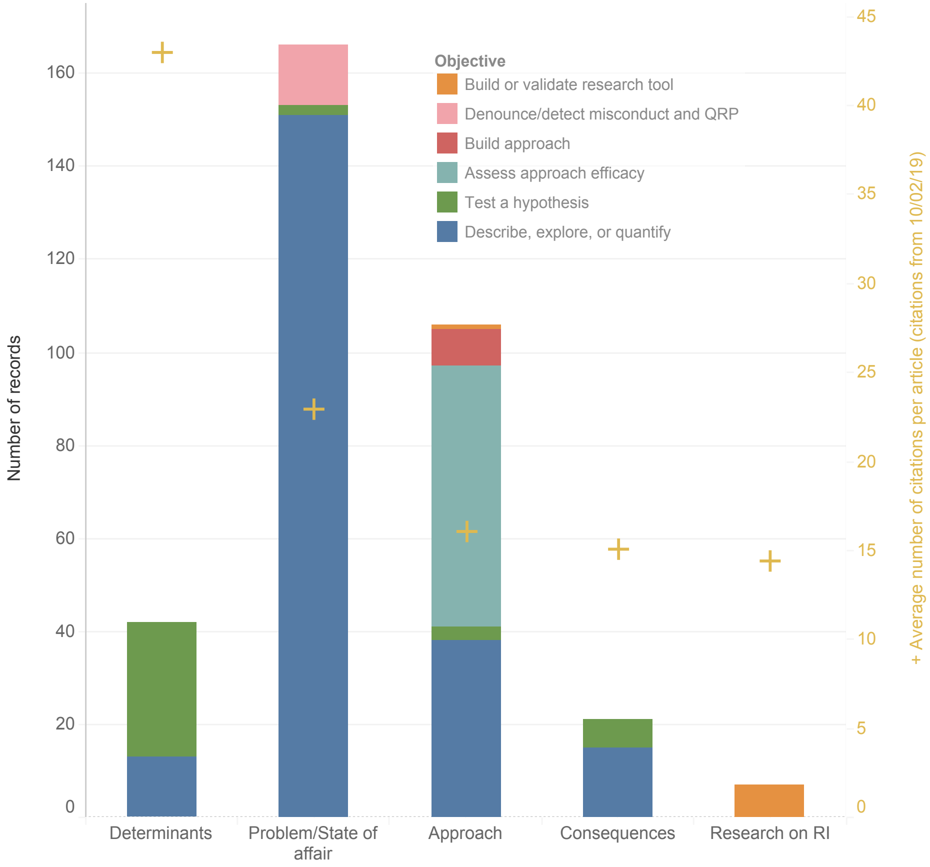
We looked in greater depth into papers focusing on *determinants* and *approaches* for misconduct and QRP.

Figure 3. Studied population and general methodologies used



Representation of the major actors studied in included empirical papers. The different colours represent the general methodology used. Over 60% of included empirical papers studied researchers and students — a percentage which rose to 75% within direct approaches such as surveys, interviews, and focus groups (blue bar)

Figure 4. Focus of included empirical work



We classified papers according to their focus (i.e., which particular step of the integrity problem they looked at), and associated research objectives. Yellow crosses show the average number of citation per article for each different focus. We can see that almost half of all empirical work targets the problem while very little articles focus on determinants and consequences. Nonetheless, determinants yielded higher average number of citations (red crosses) than other focuses.

Determinants*.* We extracted findings from the papers focusing on determinants of misconduct and QRP (40 out of the 42 papers on determinants) and grouped them into factor categories to highlight what they found as potential causes for misconduct and QRP.

We then grouped these into broader groups as either highlighting (i) *personal* issues, (ii) issues with the research *system*, or (iii) issues with researchers' *awareness and compliance*. In addition, we computed a weighed indicator for the determinant groups to ensure that regardless of the number of determinants found per paper, each article would account only for 'one' paper. (e.g., if a paper found three determinants, each determinant would have a weight of .33), 45% of the papers found that problems with the *system* played a role in misconduct and QRP, while only 16% of papers found that problems of *awareness and compliance* of researchers were at play.

Figure 5. Determinants and approaches to misconduct and Questionable Research Practices

A) Frequency findings/targets in determinants and approaches

B) Specific determinants and approaches found

			Number of mentions/ records	Weighed number of mentions*
Determinants	PERSONAL	Personality	9.0	6.2
		Career stage	4.0	1.4
		Country of affiliation	3.0	1.8
		Gender	3.0	1.2
		Rationalization	3.0	2.2
		Career Level	1.0	0.2
		Other personal determinants †	7.0	2.6
		Specific target total*	30.0	15.5
	SYSTEM	System pressure	10.0	5.3
		Perception of climate	8.0	4.5
		Financial incentives	5.0	3.6
		Disciplines	2.0	1.5
		Lack of resources	2.0	0.8
		Other system determinants ‡	9.0	2.2
		Specific target total*	36.0	18.0
	AWARENESS AND COMPLIANCE	Mentoring	4.0	1.6
		Inadequate oversight or policies	3.0	1.1
		Modeling	3.0	1.7
		Lack of awareness	3.0	1.1
		Other: Scientific expertise	1.0	1.0
		Specific target total*	14.0	6.5
		Focus total	80.0	40.0
Approaches	SYSTEM	Retractions	7.0	
		Correction and review system	2.0	
		COI management	1.0	
		Evaluation index	1.0	
		General	1.0	
		Statistical Methods	1.0	
		Specific target total*	13.0	
	AWARENESS AND COMPLIANCE	Training	49.0	
		Guidelines	27.0	
		Detection, prediction, and whistleblowing	12.0	
		Allegations	5.0	
		Specific target total*	93.0	
		Focus total	106.0	

Determinants (in red; n=79 determinants in 40 papers), and approaches (in blue; n=106 approaches in 106 papers) to misconduct and QRP found or proposed in empirical papers. We can see that most papers on determinants found that issues with the system contributed to misconduct and Questionable Research Practices (QRP), while most articles proposing or assessing approaches target researchers' awareness and compliance

Notes:
* We equally weighed the determinants to ensure that, regardless of the number of determinants found, each article would account only for 'one' paper (e.g., if a paper found three determinants, each determinant would have a weight of .33 in the paper count)
† Other personal determinants: Need for recognition, Opportunistic (Internet), Prior misconduct, Single authorship, Personal problems
‡ Other system determinants: Professional relationships, Fear of retaliation, Culture of compliance, Hampered criticism, Type of institution, Job insecurity

Two precisions are important here. First, the papers we classified in the 'determinants' categories sometimes reported direct effects on the prevalence of misconduct and QRP, but other times they reported the influence, or the perceived influence of different factors on ethical behaviours, compliance, or reporting bias. Second, even though Figure 5 only includes factors which were found to influence integrity, a few negative or integrity-promoting findings were also highlighted within the papers that looked at 'determinants', and some of those effects are not visible in Figure 5.[4]

Approaches*.* Similar to the determinants, we classified papers which targeted 'approaches' to misconduct and QRP into categories which we later grouped as either targeting the *system*, or targeting researchers' *awareness and compliance*. We did not include *personal* issues in the approaches, as we considered these to be somewhat immutable (i.e., no approach can really target or aim to change gender, seniority, discipline, or country of affiliation). As we can see in Figure 5, almost 88% of papers on *approaches* targeted researcher's *awareness and compliance*, while very few papers targeted the *system*.

Geographic distribution and citation analysis

Affiliations from the United States accounted for over half of the literature on RI captured by our sample (Figure 6B). The UK, Australia, Canada, India, Croatia, and the Netherlands followed respectively, each accounting for more than ten included articles. China, which is rapidly becoming the second most important player in scientific publication, only accounts for seven articles included (0.7%).

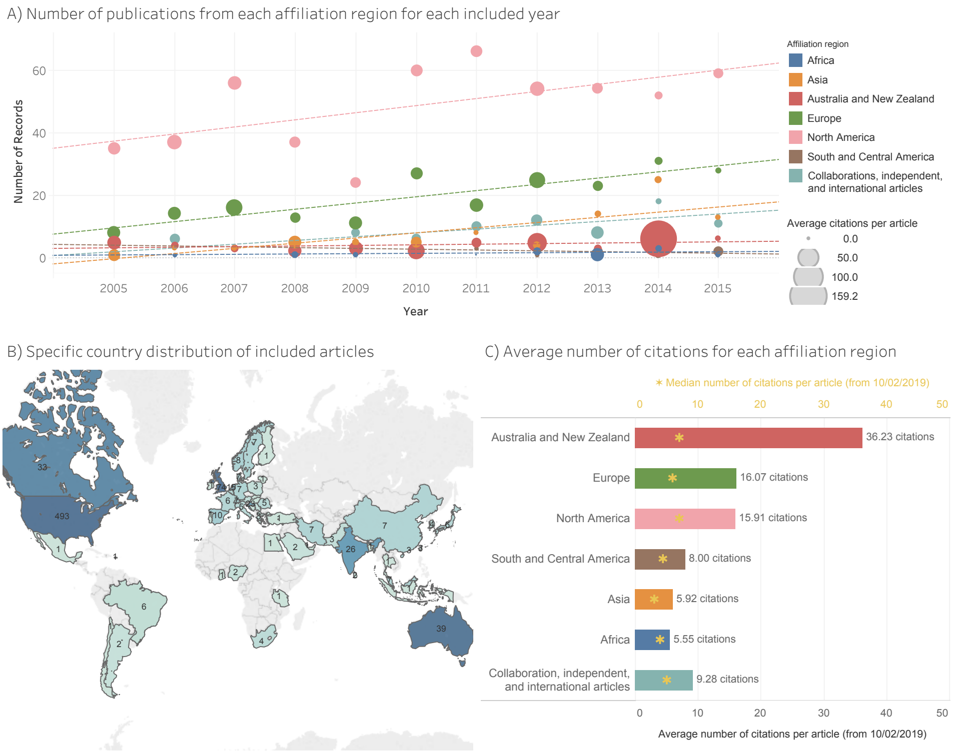
---

[4] When training was found to deter misconduct and QRP (i.e., to promote integrity), we counted it as if the paper stated that 'lack of awareness' contributed to misconduct and QRP (Geller, Boyce, Ford, & Sugarman, 2010; Kraemer Diaz, Spears Johnson, & Arcury, 2015). When papers found no effects of potential factors, we did not include those factors in the findings in Figure 5. The negative effects found were as follows: Stroebe, Postmes, and Spears (2012) found that social psychology (i.e., discipline) was *not* more prone to fraud than other disciplines. Woolley et al. (2011) found that, although country of affiliation, prior misconduct, and single authorship were related to higher misconduct-related retractions, declarations of financial incentives were not. Mumford et al. (2007) found a whole array of factors encompassing all three categories of *personal, systemic*, or *awareness and compliance*, but they also found that work commitment and limited competition did not promote unethical decisions. Finally, Fanelli, Costas, and Larivière (2015) found that inadequate oversight or policies, financial incentives, hampered mutual criticism, and career stage affected scientific integrity, but not gender and pressure.

Although we did not do a thorough citation analysis, we wanted to have an overview of the citation patterns of included articles. A few things are important to note before getting into our findings. First, we extracted citation counts directly from the databases where we extracted the record (i.e., each database counts the citations of its articles based on its pool of included material, and therefore differs), a database effect is thus possible. Additionally, when the citations were not available in the database, we looked for citations by searching for the DOI or title first in SCOPUS, then in Web of Science, and if unavailable, we grabbed the citation counts from Google Scholar. We marked the source of the citation count in the column 'Citation Source' in our online data files. Second, we looked at the total number of citations for each included paper without normalizing for the 'age' of the paper. We made this decision to avoid possible issues linked with normalization (Ioannidis, Boyack, & Wouters, 2016). Consequently, it is important to consider that reported citation means and medians may be influenced by the number of years the publications have been online, the output of the years following publication, or, on a country level, the size of the output in early years of research on research integrity; we include the average citations per paper for each publication year in Figure 6a.

We extracted citation counts from SCOPUS and Web of Science on the 10^{th} of February 2019 (older citation counts from November 2017 are also available in our online data file). On average, articles were cited 15 times, yet the distribution of citations was heavily skewed. The median number of citation was 6, and 103 articles (10%) were seemingly never cited by February 2019. Within the 10% of the literature that was never cited, only 25% were published in 2015, proposing that uncited records probably have slim chances of being taken up in the future. Over half of the total citations came from less than a tenth of the included papers (7.6%). Looking specifically within empirical papers, we further noticed that articles focusing on 'determinants' of misconduct and QRP yielded on average more citations per paper than research focusing on the 'problem', its 'approaches', or its 'consequences' (see the yellow crosses in Figure 4). When looking at highly cited papers (we selected a cut off of 30 citations; n=77), 64% were empirical, and over half (54.7%) had a main affiliation from the United States.

Figure 6. Citation and publication distribution



(A) Number of article included per affiliation region for each included year. The size of markers represents the average number of citations per paper acquired by such region in a specific year. Trend lines illustrate the publication growth for each region. **(B) Specific country distribution of included articles.** We did not include collaborations (i.e., articles with several countries mentioned in the affiliations included in the reference record), international (i.e., articles in which the main author was represented by an international organisation), and independent articles (i.e., where the main author did not mention a located affiliation) on the map. **(C) Average number of citations for each affiliation region.** Australia/New Zealand dominates the average number of yearly citations per article, but this is due in part to one heavily cited paper. The yellow asterisks displaying the median number of citations per paper show much greater uniformity between regions.

When looking at the citation weight for different continents, it was clear that North America generates most citations in research on RI, but that this dominance of citations is partly due to the important number of publications it generates (especially the United States). In fact, North America has a lower citation average than Australia/New Zealand, and a citation average similar to European averages (Figure 6C). Australia/New Zealand has the highest citation average, but this may be due to one very highly cited paper. The median number of citations per papers (i.e., the yellow asterisks in Figure 6C) are more uniform between regions, ranging from a median of 7 citations (i.e., North America and Australia/New Zealand) to one of 3 citations (i.e., Asia).

DISCUSSION

Research on research integrity (RI) is a field that is difficult to review systematically. The lack of consistency of its key terms, the absence of a clear delimitation of its scope, the interdisciplinary nature of the journals it targets, and the inconsistency of the article formats it employs to report empirical works make research on RI a fractionated field in which systematic and comprehensive overviews are challenging. Nevertheless, our analysis of a decade (2005-2015) of scientific articles in the field of research on RI reveals a few important points which may help us define an agenda for future research in the field. We will start by describing diverse noticeable findings from our results and will end with what we consider the two main messages from our study, namely, the lack of research on a number of key actors, and the mismatch between what we know on the possible causes for misconduct and the approaches empirically assessed to promote integrity.

SELECT NOTICEABLE FINDINGS FROM OUR RESULTS

The first noticeable findings from our results are the low proportions of 'articles' and empirical works. First, the low representation of research 'articles' compared to other publication formats in our initial search (i.e., 44.4%) was atypical for scientific disciplines (e.g., for medicine and health sciences the proportion of journal article surpasses 90%; SCOPUS Content Coverage Guide of 2016). Nonetheless, similar proportions may be seen in disciplines such as politics and policy, suggesting that such disciplines may be more aligned with the type of documents published in research on RI. Second, the low proportion of empirical works among included articles was surprising given that we only included publications automatically classified as 'articles' and thus excluded most editorials, letters, and other more theoretical types of publications. Although imperfections in the automatic classification may explain parts of this finding — as we described in the methods — the broad and multidisciplinary relevance of RI may also come into play. For instance, given the fact that few of the authors of articles on RI are engaged full-time in RI, and that collaborators and target audience sometimes spread through an array of distinct disciplines, it may still be challenging to engage in empirical works on the topic. It would be interesting to see whether the proportion of empirical articles increased in recent years (i.e., 2016 on) now that research on RI is becoming a field of its own.

The second noticeable finding from our result is the distribution of topics targeted. Although we admit that our keywords may have played a role in the topics found in our results, certain topics were seldom explored in our sample despite their direct relevance towards RI. For example, Research ethics committees/Institutional review boards ('REC/IRB'), ‘Peer review’, and ‘Whistleblowing’ — which may all be considered as potential safeguards for integrity in research — were very rarely the main topics of included papers. The current focus, instead, appears to be motivated by describing the problem ('QRP and misconduct'), strengthening reporting standards ('Publication ethics', 'Conflicts of interest', 'Plagiarism', and 'Authorship'), and examining integrity training and policies ('RCR training, education and mentorship' and 'Guidelines and policies'). We also found that empirical research was more frequent in academic cheating, falsification of credentials, and integrity training. This might result from the relative ease of building empirical designs in such topics compared to other topics.

The third noticeable finding from our results is the geographic distribution of our sample. The predominance of the United States in the affiliations as the most represented country in our sample is not surprising given that they are the biggest player in published literature worldwide (see for example Phillips, 2016). Nonetheless, China, which is rapidly becoming an important player in scientific publishing worldwide, was scarcely represented in our sample. It is possible that the language limitations from our study (i.e., we only included articles in English), contributed to this disparity, but it would be interesting to extend this search to different languages to assess whether this is the case.

Finally, when looking at the distribution of citations generated by our sample, we noticed that the distribution of citations was highly skewed, and that a notable percentage of articles were never cited. The skewness of citation distributions is not specific to research on RI, and is also known occurs within single journals (see for example Larivière et al., 2016). The highly-skewed distribution of citations included in our sample may simply propose that research on RI is not immune to such dynamics. The fact that about a tenth of the included papers were never cited four years or more after publication also raised some questions. First, are there more efficient dissemination systems that could ensure utility and uptake of research on RI; and second, is it possible that published research on RI is being used but not attributed as such? It is conceivable, for example, that a significant part of the readership of research on RI uses RI literature to stay up-to-date, to gain insight, and to update training or policy rather than to conduct research, thereby using the findings without citing the articles as such? We have not conducted a deeper analysis about the source of the citations and about possible network in the citing patterns, however we assume that such analyses may also yield interesting results. In particular, investigating whether citation counts of research on RI correlate with implementation, systemic changes, and policy building may be relevant to better understand dynamics of change and impact in the field.

EMPIRICAL RESEARCH ON RESEARCH INTEGRITY OVERLOOK IMPORTANT ACTORS

As we explained in our findings, researchers and research students were the most involved in empirical works on RI. The high representation of researchers and research students is not surprising given that researchers are directly affected by and targeted in research misconduct and questionable research practice. Nonetheless, other players involved in research who also have an important role in promoting integrity appeared left out from our sample. Studies on *policy makers* and *institutions*, for example, were sparse and rarely involved direct contact with these actors, despite their crucial role in defining funding and regulations (see Figure 3). *Research ethics committees* and *peer-reviewers* were also rarely studied, despite their potentially powerful role in preventing and detecting misconduct. The *public and research participants* were only studied in a few papers that explored the consequences of misconduct (e.g., loss of public trust, risks to research participants), and they were rarely approached directly.

Altogether, this imbalance points to an important gap of knowledge in research on RI. Different members of the research community are unlikely to have the same perceptions and expectations towards research (Bird, 2010). Involving a more balanced share of diverse research actors would likely bring new perspectives to the discussion. But beyond individual actors' perspectives, the social contexts and the interaction between actors was also largely untouched by empirical works. Given the complex relationships between research actors and their interrelated dependencies, considering the broader social contexts, the conflicting perspectives, and the shared expectations of different research actors may be essential in building a realistic and comprehensive understanding of RI and misconduct.

A MISMATCH BETWEEN WHAT WE KNOW AND WHAT WE PROPOSE

At first glance, our results suggest that we know a lot about the problem of integrity, but that our understanding of why misconduct happens (determinants), what it engenders (consequences), and what can be done to promote integrity (approaches) is still limited.

The lack of research on determinants of misconduct and systemic approaches for promoting integrity is not new and has been called before (see for example Fanelli, 2015). Our findings add to this perspective by highlighting that this imbalance also reveals a mismatch between what we know may predispose to inadequate research practices and the approaches to target misconduct that are discussed in the empirical literature. Specifically, factors identified as contributing to misconduct and QRP (i.e., determinants) most often point to the system, while approaches to deter misconduct and QRP most often target researchers' awareness and compliance, rather than problems of the research system. In 2016, Ana Marušić and colleagues performed a Cochrane Review looking at the effectiveness of interventions to prevent misconduct and promote integrity in research. Most of the included interventions consisted of training or guidelines to build awareness. The review, however, concluded that "Due to the very low quality of evidence, the effects of training in responsible conduct of research on reducing research misconduct are uncertain" (Marušić, Wager, Utrobicic, Rothstein, & Sambunjak, 2016, p.2).

While we can only speculate on the reasons for the disconnect between the problems that are known to threaten integrity and the solutions proposed, a few plausible explanation may be worth mentioning.

First, the disconnect may be an artefact of our search strategy. In fact, in keeping only empirical works in our analysis of determinants and approaches we inevitably overlooked non-empirical articles that proposed approaches to target systemic changes (for example Begley & Ioannidis, 2015; Fang & Casadevall, 2015; Koole & Lakens, 2012). Nonetheless, even if our analysis overlooked these articles, our findings still suggest that more ideas for systemic changes need to be tested empirically.

Another possible explanation for this disconnect may come from our search terms. Specifically, although several scientific revolutions that propose tangible changes to research systems are currently happening (e.g., movements for open science, reproducibility crisis, evolution of metrics, etc.), it is possible that these projects do not commonly employ the research integrity jargon that we used in our search strategy. If this is the case, our findings may indicate that the field of research integrity need to ensure tight connections with disciplines in which the system is actively examined and tackled (e.g., research assessments, bibliometry and scientometry, university management, open science, responsible research innovations, research on human resources and professional wellbeing).

Finally, another more fundamental reason behind this disconnect may come from a general perspective that research integrity resides within individuals. In fact, the types of approaches identified in our review — training, codes of conduct and guidelines, monitoring, whistleblowing, and allegations — all capitalize on individual researchers. Only in our analysis of empirical articles, we show that 30 studies identified personal factors as determinants of misconduct and QRP (Figure 5), many of which concluded that psychological traits and morals were major determinants of integrity (e.g., Antes et al., 2007; Bailey, 2015; Davis, Wester, & King, 2008; Hren et al., 2006; Okonta & Rossouw, 2013). The US National Academies of Sciences, Engineering, and Medicine report '*Fostering Integrity in Research*' provides an excellent overview of the evolution of the causes that have been thought to play a role in misconduct and questionable research practices (Chapter 6; 2017). It is important to note that the initial examinations of misconduct are largely built on opinions and discussions rather than on empirical research, and that this often still applies. Early examinations of scientific misconduct were rooted in individuals characteristics, often framing misconduct as a deviance, a moral defect, or even a psychopathology. Similar perspectives are still around in current empirical works (e.g., Bailey, 2015; Davis et al., 2008; Tijdink et al., 2016), although they seem to be losing grounds in the growing body of research on research integrity. The neoliberal transformation of universities and its individualisation of responsibilities are also thought to have accentuated this deeply personal facet of research integrity (Amsler & Shore, 2017; Davies, 2019). Tying misconduct with such deeply engrained personal characteristics, however, implied that very little could be done to prevent it (Hackett, 1994). Instead, approaches aimed to minimize the damage of misconduct by strengthening obligations (codes of conduct), oversight (research integrity offices, whistleblower channels) and disciplinary actions (misconduct allegation). As the discussion on misconduct evolved, many embraced more mutable albeit still individual characteristics such as the lack of awareness and understanding of research misconduct and ethical reasoning. This new perspective led to the approaches that were most represented in our study: courses and training on research integrity. The impact of research environments on individual decisions was also discussed extensively, moving the misconduct discussion away from the individuals towards an investigation of research systems. Yet, even if the latter were already part of early discussions of misconduct (e.g., Robert K. Merton and the concept of anomie) they were seldom captured by tangible approaches to be tested empirically.

We do not undermine the role that individual researchers have in preserving integrity, nor do we downplay the need for researchers to receive appropriate training and modeling to build a solid understanding of integrity in science. Nonetheless, focalising integrity approaches exclusively on researchers risks undermining the impact that research environments, demands, climates, and cultures have on research practices. In turn, ignoring the fundamental systemic triggers behind problematic research practices is likely to yield unsustainable, short-term solutions that largely overlook the grey area of questionable research practices. To use the words of the National Academies of Sciences, Engineering, and Medicine report, "*Choosing to stick with assumptions that are not supported by evidence as the basis for strategies to prevent and address research misconduct and detrimental research practices (DRPs) may perpetuate suboptimal responses on the part of the community, causing the negative consequences and damage resulting from misconduct […] to be greater than they need to be*." (NASEM, 2017, p. 93).

In summary, the past decade of research on RI has undeniably produced useful knowledge and improved our understanding of the issues faced by researchers and the research system, and it certainly continues to do so. Our review highlights the areas, methods, and actors that have been most studied, and sheds light on points which have been overlooked. Being aware of unanswered questions in research on RI is a first step toward generating executable knowledge that will allow us to better align the research agenda with the goal of promoting integrity in research.

BEST PRACTICES

As we have thoroughly discussed in the Methods section, studying research on RI is methodologically challenging. The costs from our efforts to keep the review sample manageable must be considered carefully. Importantly, our decision to limit our search to papers from 2005 to 2015 and to automatically categorised 'article' types, our choice of keywords, and our decision to only include articles published in English all had important impacts on our findings. Such decisions were necessary to kept our sample manageable within the resources available, but they also came at a cost. We made efforts maximise the relevance of our results and to ensure transparency throughout the paper, yet it may be important to reiterate that the findings from this study must be used in consideration of the methods and decisions taken. In order to increase transparency and reuse, we encourage authors to examine and complement the dataset shared alongside the paper.

RESEARCH AGENDA

Even though the present work is only a first glance in the broad body of research on research integrity, several points brought up by our analysis may serve to inspire future research agendas.

First, although the predominant involvement of researchers and research students in research on RI is justified given their implication at the core of research practices, involving participants beyond research-producers in future research on RI might help broaden our understanding of the problem. In particular, exploring the perspectives of different research actors and the social context that links these actors might help assess the possibilities, impact, and acceptability of different approaches to foster integrity. In the same way, involving topics and actors who play important roles early in the research process (e.g., research ethics committees, policy makers, funders) may be key to better understand how misconduct can be prevented.

Second, reconnecting the approaches that are proposed and assessed empirically to what is already known from past research on determinants of misconduct may be essential to increase the success of future approaches to foster RI and deter misconduct. In other words, research on research integrity may benefit from developing methods and projects to assess feasibility and success of *systemic* approaches that go beyond researchers' *compliance and awareness*.

EDUCATIONAL IMPLICATIONS

The sense of urgency attached to the topic of misconduct sometimes appears to push scientists to explore new venues for solutions rather than to optimize pre-existing opportunities. We found that past research on RI most often discussed problems with research integrity and reporting standards, and responded by proposing new surveillance, training, and compliance techniques. Nonetheless, in focusing on new approaches, researchers may overlook important insights from past research and useful safeguards which are currently available in the research organisation (e.g., peer-review, whistleblowing, research ethics committees). Building greater cohesiveness in the field of RI to allow comprehensive iterations of past research and approaches might help better optimise existing opportunities for fostering integrity.

REFERENCES

Amsler, M., & Shore, C. (2017). Responsibilisation and leadership in the neoliberal university: a New Zealand perspective. *Discourse: Studies in the Cultural Politics of Education, 38*(1), 123-137. doi:10.1080/01596306.2015.1104857

Antes, A. L., Brown, R. P., Murphy, S. T., Waples, E. P., Mumford, M. D., Connelly, S., & Devenport, L. D. (2007). Personality and Ethical Decision-Making in Research: The Role of Perceptions of Self and Others. *Journal of Empirical Research on Human Research Ethics: An International Journal, 2*(4), 15-34. doi:10.1525/jer.2007.2.4.15

Aubert Bonn, N., Godecharle, S., & Dierickx, K. (2017). European Universities' Guidance on Research Integrity and Misconduct: Accessibility, Approaches, and Content. *Journal of Empirical Research on Human Research Ethics, 12*(1), 33-44. doi:10.1177/1556264616688980

Bailey, C. D. (2015). Psychopathy, academic accountants' attitudes toward unethical research practices, and publication success. *Accounting Review, 90*(4), 1307-1332. doi:10.2308/accr-50970

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in Science Improving the Standard for Basic and Preclinical Research. *Circulation Research, 116*(1), 116-126. doi:10.1161/circresaha.114.303819

Bird, S. J. (2010). Responsible Research: What is Expected? *Science and Engineering Ethics, 16*(4), 693-696. doi:10.1007/s11948-010-9248-9

Butler, D. (2013). Investigating journals: The dark side of publishing. *Nature, 495*, 433–435. doi:10.1038/495433a

Davies, S. R. (2019). An Ethics of the System: Talking to Scientists About Research Integrity. *Science and engineering ethics, 25*(4), 1235-1253. doi:10.1007/s11948-018-0064-y

Davis, M. S., Wester, K. L., & King, B. (2008). Narcissism, entitlement, and questionable research practices in counseling: A pilot study. *Journal of Counseling and Development, 86*(2), 200-210. doi:10.1002/j.1556-6678.2008.tb00498.x

De Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal Misbehavior: Scientists Talk About the Ethics of Research. *Journal of Empirical Research on Human Research Ethics, 1*(1), 43-50. doi:10.1525/jer.2006.1.1.43

Fanelli, D. (2015). We need more research on causes and consequences, as well as on solutions. *Addiction, 110*(1), 11-13. doi:10.1111/add.12772

Fanelli, D., Costas, R., & Larivière, V. (2015). Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. *PLoS ONE, 10*(6), e0127556. doi:10.1371/journal.pone.0127556

Fang, F. C., & Casadevall, A. (2015). Competitive Science: Is Competition Ruining Science? *Infection and Immunity, 83*(4), 1229-1233. doi:10.1128/iai.02939-14

Geller, G., Boyce, A., Ford, D. E., & Sugarman, J. (2010). Beyond "compliance": The role of institutional culture in promoting research integrity. *Academic Medicine, 85*(8), 1296-1302. doi:10.1097/ACM.0b013e3181e5f0e5

Hackett, E. J. (1994). A Social Control Perspective on Scientific Misconduct. *Journal of Higher Education, 65*(3), 242-260. doi:10.1080/00221546.1994.11778499

Hren, D., Vujaklija, A., Ivanišević, R., Knežević, J., Marušić, M., & Marušić, A. (2006). Students' moral reasoning, Machiavellianism and socially desirable responding: Implications for teaching ethics and research integrity. *Medical Education, 40*(3), 269-277. doi:10.1111/j.1365-2929.2006.02391.x

Ioannidis, J. P. A., Boyack, K., & Wouters, P. F. (2016). Citation Metrics: A Primer on How (Not) to Normalize. *PLoS Biology, 14*(9), e1002542. doi:10.1371/journal.pbio.1002542

Koole, S. L., & Lakens, D. (2012). Rewarding Replications: A Sure and Simple Way to Improve Psychological Science. *Perspectives on Psychological Science, 7*(6), 608-614. doi:10.1177/1745691612462586

Kraemer Diaz, A. E., Spears Johnson, C. R., & Arcury, T. A. (2015). Perceptions That Influence the Maintenance of Scientific Integrity in Community-Based Participatory Research. *Health Education and Behavior, 42*(3), 393-401. doi:10.1177/1090198114560016

Larivière, V., Kiermer, V., MacCallum, C. J., McNutt, M., Patterson, M., Pulverer, B., . . . Curry, S. (2016). A simple proposal for the publication of journal citation distributions. *bioRxiv*. doi:10.1101/062109

Martinson, B. C., Anderson, M. S., & De Vries, R. (2005). Scientists behaving badly. *Nature, 435*(7043), 737-738. doi:10.1038/435737a

Marušić, A., Wager, E., Utrobicic, A., Rothstein, H. R., & Sambunjak, D. (2016). Interventions to prevent misconduct and promote integrity in research and publication. *Cochrane Database of Systematic Reviews*(4). doi:10.1002/14651858.MR000038.pub2

Mumford, M. D., Murphy, S. T., Connelly, S., Hill, J. H., Antes, A. L., Brown, R. P., & Devenport, L. D. (2007). Environmental influences on ethical decision making: Climate and environmental predictors of research integrity. *Ethics and Behavior, 17*(4), 337-366. doi:10.1080/10508420701519510

NASEM. (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press.

Okonta, P., & Rossouw, T. (2013). Prevalence of scientific misconduct among a group of researchers in Nigeria. *Developing World Bioethics, 13*(3), 149-157. doi:10.1111/j.1471-8847.2012.00339.x

Phillips, N. (2016, 20 April). US tops global research performance. *Nature Index*. Retrieved from https://www.natureindex.com/news-blog/us-tops-global-research-performance.

Pupovac, V., & Fanelli, D. (2014). Scientists Admitting to Plagiarism: A Meta-analysis of Surveys. *Science and Engineering Ethics, 21*(5), 1331-1352. doi:10.1007/s11948-014-9600-6

Steneck, N. H. (1999). Confronting misconduct in science in the 1980s and 1990s: What has and has not been accomplished? *Science and Engineering Ethics, 5*(2), 161-176. doi:10.1007/s11948-999-0005-x

Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science, 7*(6), 670-688. doi:10.1177/1745691612460687

Tijdink, J. K., Bouter, L. M., Veldkamp, C. L. S., van de Ven, P. M., Wicherts, J. M., & Smulders, Y. M. (2016). Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study. *PLoS ONE, 11*(9), e0163251. doi:10.1371/journal.pone.0163251

Woolley, K. L., Lew, R. A., Stretton, S., Ely, J. A., Bramich, N. J., Keys, J. R., . . . Woolley, M. J. (2011). Lack of involvement of medical writers and the pharmaceutical industry in publications retracted for misconduct: A systematic, controlled, retrospective study. *Current Medical Research and Opinion, 27*(6), 1175-1182. doi:10.1185/03007995.2011.573546

Chapter 2

A short introduction to academic research in Flanders

CONTRIBUTIONS

Conceptualisation: Noémie Aubert Bonn

Investigation: Noémie Aubert Bonn

Writing – original draft: Noémie Aubert Bonn

Writing – review & editing: Noémie Aubert Bonn; Wim Pinxten[2] and Stephanie Ruysschaert[1,] reviewed earlier versions of this chapter; Nele Nivelle[1] reviewed later versions of this chapter and provided constructive feedback on the content of the chapter; Ann Peters[3], who served as external jury to the current thesis, noted mistakes and inconsistencies and provided essential feedback in revising the final version of this chapter.

1. Research Coordination Office, Hasselt University, Hasselt, Belgium
2. Department of Healthcare and Ethics, Hasselt University, Hasselt, Belgium
3. Institute of Tropical Medicine Antwerp, Antwerp, Belgium

ABSTRACT

In the present thesis, we use Flanders — a self-governing region of Belgium — as a sample of study. Focusing our work on Flemish research allowed us to grasp perspectives from interacting and complementary actors within a rich research system while preserving a manageable sample. In this chapter, I briefly introduce Flanders and describe how its Research and Development scene is organized. Amongst other things, I will discuss how Flanders finances and administers academic research and I will describe its relationship with research integrity.

INTRODUCING BELGIUM AND FLANDERS

Belgium is a complex country. To use an expression retold by Renée C. Fox in 1962, Belgium is a "veritable social mosaic" (Fox, 1962).

> *"[W]ithin the 11,779 square miles that comprise Belgium, its more than 9 million inhabitants distribute themselves in countless ways between two cultures, French and Flemish; two languages (each with numerous dialects); two sharply contrasting philosophical-religious attitudes toward life (traditional Catholicism and anticlerical Masonic 'Free Thought'); and four political parties (Social Christian, Liberal, Socialist, and Communist)." (Fox, 1962, p. 436)*

Despite its age, Fox's depiction nicely illustrates the complexity and richness of a country in which philosophical, political, and language divides govern every aspect of public life. Now part of the European Union (and arguably its capital), Belgium's governance is separated through seven different entities: a general federal entity (i.e., the federal government, Chamber of Representatives & Senate), three Communities separated on the basis of language (i.e., the Flemish Community, the French Community, and the German-speaking Community), and three Regions (i.e., the Flemish region, the Walloon Region, and the Brussels-Capital Region) (Bergen & Craps, 2017). Communities and Regions share different powers, but each has its own Parliament and is granted the right to vote laws and regulations applicable for its territories and populations. In the present thesis, we are interested in the Flemish federated entities. These consists of a union of the

Flemish Region and Community, and comprise one Flemish Parliament and one Flemish Government which decide on all matters about Flanders and about the Flemish community in Brussels. Science and Innovations in Flanders depends on these united Flemish entities as well as on the Chamber of Representatives, which has powers that transgress Regions and Communities. For the purpose of the present work, we will target Flemish research from a practical standpoint — how research is organized and managed in Flanders — rather than from a political standpoint — how research is governed in Flanders.

UNIVERSITIES AND RESEARCH INSTITUTES

Flemish research is now recognized as vibrant, cutting edge, and internationally competitive. Flanders is host to a number of research institutes. Among those are five main universities (Katholieke Universiteit Leuven (KULeuven), Universiteit Gent (UGent), Universiteit Antwerpen (UAntwerpen), Vrije Universiteit Brussel (VUB), and Universiteit Hasselt (UHasselt)), four strategic research centres (Interuniversity Microelectronics Centre (IMEC), Vlaams Instituut voor Biotechnologie (VIB), Vlaamse Instelling voor Technologisch Onderzoek (VITO), and Flanders Make), and a few additional institutes which operate on domain-specific topics, such as the Vlaams Instituut voor de Zee (VLIZ; marine sciences), the Institute of Tropical Medicine (ITM; tropical medicine), and the Instituut voor Landbouw, Visserij en Voedingsonderzoek (ILVO; agriculture research).

The higher education and research career tracks in Flanders are similar to those in other countries with some small distinctions. Higher education takes place in universities and in university colleges (hogeschool). University colleges provide professional bachelor’s degrees that allow graduates to be employed as a professional immediately or academic programmes in the fields of Nautical Sciences, Audiovisual and Visual Arts and Music and Performing Arts (Eurydice, 2019). Universities typically offer bachelor's degrees, master's degrees, advanced master's degrees, and PhD degrees (doctoraat). During master’s and PhD theses, students are supervised by a more senior member of the university who is referred to as a ‘mentor’, ‘promotor’, or ‘supervisor’. Teaching assistants, research assistants, and technician positions are available for those with a master's degree, but in absence of a PhD, further academic career is excluded. Those who wish to enter academia will need to complete a PhD and continue through a post-doctoral research assistant position (Doctor-assistent). PhD students and post-doctoral research assistants are most often funded with research grants obtained by the candidate or the promotor, which gives insecurity towards continuity. Alternatively, PhD graduates can become teaching assistant-PhD, a position that is funded by the university and is limited in time to 6 years. After completing one or several post-doctoral research assistant contracts, the candidates can apply for a position as independent academic staff (ZAP, zelfstandig academisch personeel), however ZAP positions are scarce compared to the number of PhD graduates and post-doctoral positions. In Flanders, all ZAP immediately receive the title of professor and are fully independent (i.e., no supervision). There are four levels within the ZAP: assistant professor (docent), associate professor (hoofddocent), and two levels as full professor (hoogleraar and gewoon hoogleraar). Assistant professor positions (docent) are temporary positions offered as tenured tracks, generally allowing three to five years to fulfil the requirements of the contract with interim evaluations. If successful, the candidate obtains a permanent contract and is promoted to associate professor (hoofddocent), if not the contract comes to an end. Further promotion to the two degrees of full professor depend on the track record of professors with a minimum number of years in between each promotion (European University Institute, 2018).

According to figures from VLIR fact sheets, Flemish universities employed over 25 000 research staff (including 3303 Tenured academic staff, 3404 post-doctoral Researchers, and 9660 PhD Researchers who are also considered staff in Flanders) and involved almost 146 000 students in 2016 (Vlaamse interuniversitaire Raad [Flemish Interuniversity Council]). In a recent public funding observatory from the European Universities Association, Flanders was shown as the European country with the second highest 10-year increase in its number of students; an increase by 52% of its students since 2008 (Pruvot, Estermann, & Lisi, 2018). This increase may be even more marked for PhD students, as earlier assessments estimated the increase in PhD students to be around 70% from 2005 to 2014 (Furniere, 2016). VLIR reports that nearly two thousand PhD degrees were granted in 2015. Nonetheless, the increase in professorships did not follow the growing number of PhD students, leading to a stable or even decreasing percentage of PhD students securing academic positions after completing their studies. Even though almost 60% of PhD students aspire to pursue a professorship after their degree, only one in five will obtain an appointment in a Flemish institution, a number which includes part-time positions as low as 10% (Debacker & Vandevelde, 2016). This high discrepancy between the number of PhD student and the number of academic positions available will become important to contextualise some of the findings from the following chapters.

HOW RESEARCH IS SUPPORTED

Flemish Research and Development is supported from public and private sources. Public sources of funding, which include the Flemish Government, federal budgets, European Union Horizon 2020 programs, and initiatives on research and innovation from the European Union Regional Policy 2014-2020 (Geerts, Langenhove, Viaene, & Verdoodt, 2017) are added to private sources to provide Flanders with nearly 7.5 billion Euros for Research and Development (figures from 2017; Debackere, Hoskens, Joosten, Verheyden, & Viaene, 2017). In an optic to maximise Research and Development, the Flemish Government aims to provide 3% of its total GDP to Research and Development by 2020. Falling short from this objective by only 0.11% as of 2017 (Debackere et al., 2017), and with an increase of over 50% of the total expenditures for Research and Development in the past ten years, Flemish research is rapidly growing. Such big figures compose the general funding and functioning of Flemish Research and Developments, including structural funding and investments.

As in most states, a significant portion of funding for Flemish universities comes from the Flemish Government. In 2018, the government budget for core institutional funding of universities in Flanders was over nine hundred million euros, while the budget for research and innovation activities accounted for an additional two hundred million euros (Peters, 2019). The total core funding, which also includes universities' functional and procedural expenses, consists of up to 44% of all funding sources dedicated to Education and Training (VLIR, 2017), and is distributed between universities using specific distribution rules . The core funding is divided between four sub-budgets: two fixed lump-sums (with an upper and lower limit), one for education and one for research, and two variable lump-sums one for education and one for research (Peters, 2019). The fixed lump-sum for education is based on course credits at the Bachelor's and Master's level, while the variable lump-sum considers additional elements such as acquired study points and diplomas. The fixed lump-sum for research considers the number of doctoral degrees and the number of publications, while the variable one considers the number of academic oriented Bachelor's and Master's degrees, the number of doctoral degrees conferred, the number of publications and citations, and the number of first appointment of external female researchers as autonomous academic staff.

Each university also receives a budget for research and innovation activities from the special research fund (BOF; Bijzondere Onderzoeksfonds) and the industrial research fund (IOF; Industrieel onderzoeksfonds). The budget from the BOF is distributed between universities using a specific key that contains a structural portion (Part A, which comprises of Bachelor's and master's degrees awarded, defended PhDs, and gender diversity), and a bibliometric portion (Part B, which comprises of publications in the Web of Science, publications in the Flemish Academic Bibliometric Database for the Social and Human Sciences (VABB-SHW), and citations in the Web of Science). Each publication is weighed depending on the Journal Impact Factor percentile of the subfield to which it corresponds in the Web of Science (Engels & Guns, 2018). Until 2019, and thus in validity for all the interviews and focus groups undertaken in the present project, part B of the BOF-key two accounted for 40% of the final score for distribution (Zacharewicz, 2016). In the summer of 2019, a new version of the BOF key has been developed and released, but the new calculation still determines a large proportion of the resource allocation on bibliometrics.

Project funding, scholarships, and contractual salaries are then funded either directly from the institutions or from external funders. PhD students and post-doctoral researchers generally obtain funding through mandates and scholarships, or by being paid directly by a tenured researcher as part of a research project (Debacker & Vandevelde, 2016). Most scholarship and mandates are obtained either through BOF funding — thereby provided directly from the university's structural funding — or through FWO (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen [Research Foundation – Flanders]) and VLAIO (Vlaanderen Agentschap Innoveren & Ondernemen [Flemish Agency for Innovation & Entrepreneurship]) mandates — provided from national public foundations which are external to universities. FWO focuses on fundamental research while VLAIO focuses on applied entrepreneurship and innovations. Scholarships and project funding are generally based on competitive funding. The success rates of FWO and VLAIO tend to be lower than BOF scholarships, hovering between 20% and 30% success. Alternative funding channels are available for scholarships, but the majority comes from the three channels described above. As a general rule, scholarships and mandates evaluate both the proposed project and the profiles of applicants. Consequently, past successes — in particular past publications and associated impact metrics —of the junior applicant and the associated supervisor are relevant.

Beyond scholarships and mandates, Flemish research is also broadly financed from project funding. In such cases, the BOF, FWO, and VLAIO still play a major part, but other key programs such as those from the Belgian Science Policy Office (BELSPO) and the European Commission (Horizon 2020, ERC) also provide a big portion of the funds. In those cases, project funds are attributed to a specific research project elaborated by a specific research team. Although these grants are largely dependent on the assessed innovation, feasibility, and prospective outcomes of the proposed research project, they also consider individual characteristics from the researchers' profiles and past success — especially past publications and associated impact metrics.

Project funding can also come from the private sector, either through industry funding — which is very important in Flanders ("Vlaamse universiteiten wereldkampioen," 2013) — or through private funding agencies. Nonetheless, the share of private funding is much smaller than public funding in Flemish research.

Considering these different funding pathways, it is clear that the distribution of research funding in Flanders — both core funding to universities and grant funding to individual researchers — is highly based on performance and outputs (Peters, 2019).

FLANDERS AND INTEGRITY

The above figures simply serve to illustrate the growth and general functioning of Research and Innovations in Flanders, and to contextualise some of the findings from upcoming chapters. Given the core interest of the present thesis however, it is also relevant to consider Flanders' relationship with research integrity.

In terms of guidance and policies, Flemish institutions now jointly endorse the European Code of Conduct for Research Integrity from the All European Academies (All European Academies (ALLEA), 2017), but historically relied on diverse institutional and national guidance. The most notable national guidance is the 'Codes of ethics for scientific research in Belgium', issued in 2009 by the Royal science academies of Belgium and still broadly endorsed today. The 'Codes of ethics for scientific research in Belgium' are largely value based (S. Godecharle, Nemery, & Dierickx, 2014) and propose six key principles to guide researchers: rigour, caution, reliability, verifiability, independence, and impartiality (Royal science academies of Belgium, 2009). Although they provide broad descriptions of principles of good research, the Codes of the Royal science academies of Belgium provide limited guidance on what needs to be done if such principles are infringed (Aubert Bonn, Godecharle, & Dierickx, 2017).

In the years that followed the establishment of the Royal science academies' Codes of ethics, awareness of the importance of research integrity rocketed throughout Europe. In March 2011, the first version of The European Code of Conduct for Research was issued (European Science Foundation and ALLEA, 2011). The code was rapidly adopted in Europe and its revised version (All European Academies (ALLEA), 2017) is now endorsed explicitly throughout Flemish institutions. At the same time, misconduct scandals started to surface, adding to the concern and the awareness on the topic. In 2011, the scandal of Diederick Stapel, a Dutch psychologist from Tilburg university who blatantly fabricated data for at least 58 publications (Palus, 2015) began to propagate in worldwide media (Callaway, 2011). Only a few months later, another case happened in the Dutch social psychology arena, this time from Dirk Smeesters, a researcher who undertook his studies in Flanders but worked at the Rotterdam School of Management in the Netherlands when he falsified his data (Oransky, 2012b). Despite erupting in the neighbouring country, these scandals increased the alert for integrity and misconduct in Flanders and awoke the need to control and monitor research integrity locally. Later in the same year, a complex and controversial Belgian case added to the growing discussion of research integrity in Belgium. This case involved Professor Donnez, a clinical researcher at the Université Catholique de Louvain whose reports of an ovarian transplant raised concerns. The case was very complex and included issues with the ethics approval of the study, possible manipulation of the data, unverifiable claims, editorial expression of concerns, authorship disputes, as well as a fire which burnt the documentation to prove the case (Dardenne, 2012; Oransky, 2012a, 2012c). The investigation of the institution recommended to terminate the employment of the professor, but the university decided otherwise, further adding to the confusion. Likely alarmed by these defaming cases, Flemish institutions began to equip themselves with the means to detect and address research misconduct. By 2013, each Flemish university had established a local research integrity office (referred as Commissie voor Wetenschappelijke Integriteit (CWI) in Flanders), and the Royal Flemish Academy of Belgium for Sciences and the Arts (KVAB) had put together the Flemish Committee for Scientific Integrity (Vlaamse Commissie voor Wetenschappelijke Integriteit; VCWI), a Flanders-wide platform for scientific integrity which provides second opinions on allegations conducted by Flemish institutions (Furniere, 2013). Little by little, university CWIs gained credence and obtained the power to enforce policy plans on research integrity (Furniere, 2015b). Clear allegation procedures, guidance, and integrity charters emerged in Flemish institutions in the following years. In 2013, the FWO also set out a taskforce on research integrity. The taskforce culminated in 2015 and led to the introduction of a clause on research integrity that applicants are now obliged to sign before applying for funding.

Amid the growing awareness in research integrity, an article published in the scientific magazine Eos in 2013 revealed that research integrity in Flanders was far from laudable. According to the article, around 8% of Flemish researchers admitted to having fabricated or falsified data (Verbeke, 2013). Comparing this figure to the estimated 2% of scientists who admit to falsification and fabrication worldwide (Fanelli, 2009), the disproportionately high Flemish figure raised the alarm. Only one day after the Eos article was released, a case of Flemish research fraud hit the news (Vervaeke, 2013). Although the case was initially kept largely confidential, rumours and assumptions quickly proposed that the case happened at the Vrije Universiteit Brussels (VUB). In the media, the case was compared with the case of Diederick Stapel, but newspapers argued that, since detection happened before the data was published, the damage of the Flemish case did not compare to the Stapel case. Only a couple of months after this first Flemish scandal, another one hit the news, this time leading to the retraction of a PhD degree ("Nieuw geval van wetenschapsfraude," 2013). This case tackled a student who defended his thesis in 2005, also at VUB, but had since left Belgium. Beyond the personal aspect of the highly visible cases, concerns about the stressful academic system also started to emerge. Newspapers explained that bad research practices may result from a 'deadly cocktail' between high expectations of research outcomes and limited research resources (Hoet, 2013). Such depictions reflected the reactions of movements and groups of scientists who, in the same period, argued that the pressures and the fierce competition of the current system encouraged bad practices and discouraged scientific progress. For instance, the Slow Science Movements were gaining grounds to support civic engagement of scientists and to denounce the pressures induced by the commercialisation of science ("Actie tijdens opening KULeuven," 2011). A Slow Science Belgium group was even developed, putting forth seminars and encouraging scientists to support the *Slow Science manifesto* (Slow Science Belgium). Along the same lines, the *Charte de la désexcellence* was advanced by researchers from French-speaking universities in Belgium to denounce and combat publication pressures (L'Atelier des Chercheurs, 2014).

In public communications, rectors from Flemish institutions acknowledged the growing concerns of the 'publish or perish' climates, but they initially focused their attention elsewhere. The rector of KU Leuven — Mark Waer at the time — argued for the importance of whistleblowing and invited researchers to come forward and denounce misconduct and inappropriate practices (Vanherle, 2013), a view which will later be supported by Rik Torfs, his successor. At VUB, on the other hand, representatives advocated that universities must take firm and decisive action on trespassers, and that students must be educated and sensitized to integrity in research ("In de schijnwerper," 2013). The need for better integrity training was later reiterated by KU Leuven (Meyvis, 2014), but the focus on the importance of a safe and confidential channel to raise concerns remained predominant. Throughout the years, the diversity of cases and issues accumulated and inspired new guidance and training. Cases of plagiarism and image manipulation (Eckert & Cools, 2018; Han, 2017) led to clear institution policies and the development of an interuniversity workshop on the topic in 2019. Issues on predatory journals also surfaced and raised awareness on the topic which had been overlooked in the past ("Honderden Belgische onderzoekers," 2018; "Ignaas Devisch," 2018). Some Flemish researchers also faced retractions from honest errors (Oransky, 2013b), giving the media an opportunity to differentiate honest error from misconduct.

For each of these cases however, the approach of the media revealed a vivid fear of reputational damage between institutions. Institutions in which misconduct was mediatized sometimes pointed at the neighbour to justify their issues ("In de schijnwerper," 2013), jeopardizing inter-institutional trust and cooperation. Others made hasty public announcements towards the innocence of their researchers to avoid damaging the reputation of their institution. For instance, in 2017, a conflicting plagiarism case hit the media ("Marc Hooghe onder ", 2017), was refuted immediately ("Plagiaat? Niets van aan," 2017), and was re-confirmed two years later (Debusschere & Gordts, 2019). The former Rector of KU Leuven who faced another particularly notorious case — the case of Stefaan Van Gool (Stern, 2017b) — defended this desire for secrecy as a mean to protect the individual who committed misconduct from public humiliation (Stern, 2017a). Doubting this perspective however, other media rather claimed that institutions covered up stories to minimize reputational damage and exploit research funds (Chini, 2019; Eckert, 2017). Whatever the reason behind such secrecy, keeping misconduct cases behind closed doors is thought to have allowed misbehaving researchers to remain employed years after concerns had been expressed (Chini, 2019), or to be dismissed discretely and to migrate to a neighbouring institution without adequate alert (see for example Oransky, 2013a).

Eager to understand how to protect the integrity of science, funding agencies such as the European Commission, FWO, the King Baudouin Foundation, and BOF institutional funding then began allocating funds to scientific projects investigating research integrity. As a result, Flanders joined international efforts and started to investigate research misconduct and integrity scientifically. Amongst others, Simon Godecharle — one of the first Flemish PhD students to focus his thesis on research integrity — allowed us to reassess the prevalence of misconduct in Flemish research. Focusing on both university and industry researchers, Godecharle showed that, unlike the reports from Eos and Tijdink (Tijdink, Verbeke, & Smulders, 2014), the prevalence of misconduct in Flanders was similar to international reports (Simon Godecharle, Fieuws, Nemery, & Dierickx, 2018). Far from lowering the interest on the topic, Godecharle’s findings rather highlighted areas where Flemish institutions should concentrate their efforts, such as authorship recognition. Later in the same year, Luc Sels was elected rector of KU Leuven. At the opening of the academic year, he dedicated the quasi entirety of his speech to research integrity. His key message mostly supported the need for early training and safe whistleblowing channels, the need to keep trust and faith in science, but also the need to continue research on research integrity ("Rector Sels on scientific integrity," 2018). Adding to Godecharle's results, a thorough report from a project funded by the King Baudouin Foundation added rich insights in the perceptions of Flemish researchers on the culture and integrity of research (Mergaert & Raeymaekers, 2017). The broad report, which included 1720 respondents from French- and Dutch-speaking universities in Belgium, showed that almost 90% believed that pressures on researchers may lead to ‘compromises on research integrity and standards’, and 85% believed that quality should be the main criterion for evaluating research. The report also showed that respondents from Flemish universities experienced lower satisfaction with the balance between their research, teaching, and services workload than respondents from French-speaking universities; that they were especially unsatisfied with the amount of time they had available for research; and that they perceived funding and career assessments as more problematic then their French-speaking colleagues.

As institutions focused on strengthening their CWIs, providing safe disclosure channels, and mandating integrity training, researchers continued to argue that the problem was in the system rather than in a lack of awareness and oversight (see some examples in Flemish news and scientific literature such as: "Academics protest against pressure," 2013; "Flemish academics rebel," 2013; Furniere, 2014; Tijdink et al., 2014). Exaggerated pressures and maladapted research assessments were among the most frequent issues described in the media. These issues, added to the stiff pyramidal and highly competitive employment climate, were later shown to generate a mental health vulnerability among younger scientists in Flanders (Levecque, Anseel, De Beuckelaer, Van der Heyden, & Gisle, 2017; Van de Velde, Levecque, Mortier, & De Beuckelaer, 2019).

The growing concerns about pressures and inappropriate research assessments slowly paved their way into Flemish institutions and encouraged the reconsideration of researcher assessment and promotion. The Brussels Declaration on Open Access of 2012 played an early part in the changes by opening the door to green access repositories in which new metrics were displayed. Some funders in Wallonia (e.g., Fund for Scientific Research F.R.S.-FNRS) then decided to allow researchers to use these indicators in their funding application if they preferred them over traditional indicators (Miedema, Mayer, Holmberg, & Leonelli, 2018). In Flanders however, most funders preferred to stick to the traditional impact factors which are still the gold standard today. As we mentioned above, the BOF key which is used to distribute special research funds between Flemish institution, divides a large portion of its funding based on bibliometric data, and despite continued concerns and controversies (Stroobants, Godecharle, & Brouwers, 2013), traditional impact metrics remained essential also in the 2019 update of the BOF key. At the level of universities on the other hand, important steps to tackle the ‘publish or perish’ culture were taken. One of the early changes happened in the differentiation of senior academic positions. In UHasselt for instance, since 2014, senior professors can choose to enter in a differentiation track as an associate professor (hoofddocent). The differentiated track allows professors to decide on a pillar of specialisation. In this regard, instead of having to focus on all three traditional pillars (i.e., research, teaching, and services) and to fulfill the usual deliverables (i.e., publications, grant applications, etc.), associate professors in the differentiated track can choose a domain of specialisation in which they focus on providing important services to the university or the wider community (e.g., innovation / valorisation, internal functioning / community, etc.) without needing to fulfil the full delivery outputs of the traditional tenure track. This differentiation constitutes an important step for allowing the development of broader coordinated teams in which personal skills are valorized (see Chapter 5 for a discussion on the lack of team valorisation in more traditional systems). Another initiative came a couple of years later, when KU Leuven adopted a new evaluation system in which they would assess researchers based on biosketchs (Furniere, 2015a). This new system aimed to de-focalise assessments from imposed performance indicators and to allow for a diversity of profiles to strive in universities. In the end of 2018, UGent introduced a new career model to reconsider recommendations from international movements such as the San Francisco Declaration on Research Assessments (American Society for Cell Biology, 2013). In this new career model, evaluations are not based on metrics, the evaluative burden is reduced, and researchers set their own objectives in the format that suits them (*Ghent University is changing course*, 7 December 2018). This drastic change, in momentum with the current international debate, made the news internationally. Although we will need to wait some years before understanding whether these evaluative changes had the desired effects, they indicate that Flemish universities are willing to accept disruption to help protect the integrity of science. With recent changes in the Netherlands where institutions and funders agreed to sign DORA and to stop looking at impact factors ("VSNU, NWO, NFU and ZonMw," 2018), I am eager to see the changes that Flanders will consent to in the coming years.

REFERENCES

Academics protest against pressure to publish. (2013, 28 August). *Flanders Today*. Retrieved from http://www.flanderstoday.eu/current-affairs/academics-protest-against-pressure-publish.

Actie tijdens opening KULeuven-academiejaar tegen ontslag Barbara Van Dyck [Action during the opening of the KULeuven academic year against the dismissal of Barbara Van Dyck]. (2011, 26 September). *De Standaard*. Retrieved from https://www.standaard.be/cnt/dmf20110926_041.

All European Academies (ALLEA). (2017). *A European Code of Conduct for Research Integrity: Revised Edition*. [ISBN: 978-2-918428-37-4]. Retrieved from https://allea.org/portfolio-item/the-european-code-of-conduct-for-research-integrity-2/

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Aubert Bonn, N., Godecharle, S., & Dierickx, K. (2017). European Universities' Guidance on Research Integrity and Misconduct: Accessibility, Approaches, and Content. *Journal of Empirical Research on Human Research Ethics, 12*(1), 33-44. doi:10.1177/1556264616688980

Bergen, D., & Craps, G. (2017). *The Parliaments of Belgium and their International Powers*. [D/2017/3933/3]. Retrieved from http://docs.vlaamsparlement.be/docs/varia/brochures//the-parliaments-of-belgium-international-uk.pdf

Callaway, E. (2011). Report finds massive fraud at Dutch universities. *Nature, 479*. doi:doi:10.1038/479015a

Chini, M. (2019, 1 July). University covered up for scholar accused of fraud, harassment: reports. *The Brussels Times*. Retrieved from https://www.brusselstimes.com/brussels-2/59763/university-covered-up-for-scholar-accused-of-fraud-harassment-reports/.

Dardenne, L. (2012, 20 July). UCL: Et dans le rôle principal, le Dr Donnez [UCL: And in the leading role, Dr. Donnez]. *La Libre*. Retrieved from https://www.lalibre.be/belgique/ucl-et-dans-le-role-principal-le-dr-donnez-51b8eea4e4b0de6db9c76647.

Debacker, N., & Vandevelde, K. (2016). From PhD to professor in Flanders. *ECOOM Brief (no. 11)*. Retrieved from https://biblio.ugent.be/publication/8043010.

Debackere, K., Hoskens, M., Joosten, W., Verheyden, L., & Viaene, P. (2017). *Totale O&O-intensiteit in Vlaanderen 2007-2017 "3% nota" [Total R&D intensity in Flanders 2007-2017 "3% note".]*. ECOOM. Retrieved from https://www.ecoom.be/assets/194

Debusschere, B., & Gordts, P. (2019, 23 April). Leuvense sterrenprof pleegt verregaande vorm van auteursfraude [Leuven star Professor commits far-reaching form of authorship fraud]. *De Morgen*. Retrieved from https://www.demorgen.be/nieuws/leuvense-sterrenprof-pleegt-verregaande-vorm-van-auteursfraude~b860d2e2/.

Eckert, M. (2017, 28 March). Torfs gunde 'topdokter' zachte exit [Torfs gived the 'top doctor' a soft exit]. *De Standaard*. Retrieved from https://www.standaard.be/cnt/dmf20170327_02802645.

Eckert, M., & Cools, S. (2018, 10 March). Wetenschapsfraude met Photoshop [Scientific misconduct with Photoshop]. *De Standaard*. Retrieved from https://www.standaard.be/cnt/dmf20180309_03400010.

Engels, T. C. E., & Guns, R. (2018). The Flemish Performance-based Research Funding System: A Unique Variant of the Norwegian Model. *Journal of Data and Information Science, 3*(4), 45-60. doi:https://doi.org/10.2478/jdis-2018-0020

European Science Foundation and ALLEA. (2011). *The European Code of Conduct for Research Integrity*. [ISBN 978-3-00-055767-5]. Retrieved from https://allea.org/portfolio-item/the-european-code-of-conduct-for-research-integrity/

European University Institute. (2018). *Belgium Flanders, Academic Career Structure*. Retrieved from https://www.eui.eu/ProgrammesAndFellowships/AcademicCareersObservatory/AcademicCareersbyCountry/BelgiumFlanders#RequirementsPositions

Eurydice. (2019). *Conditions of Service for Academic Staff Working in Higher Education*. Retrieved from https://eacea.ec.europa.eu/national-policies/eurydice/content/conditions-service-academic-staff-working-higher-education-3_en

Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE, 4*(5), e5738 (5731–5711). doi:10.1371/journal.pone.0005738

Flemish academics rebel against pressure to publish. (2013, 20 August). *The Bulletin*. Retrieved from https://www.thebulletin.be/flemish-academics-rebel-against-pressure-publish.

Fox, R. C. (1962). Medical scientists in a Chateau. *Science, 136*(3515), 476-483. doi:10.1126/science.136.3515.476

Furniere, A. (2013, 10 October). New commission battles scientific research fraud. *Flanders Today*. Retrieved from http://www.flanderstoday.eu/education/new-commission-battles-scientific-research-fraud.

Furniere, A. (2014, 15 May). Q&A: Geert Molenberghs on the battle against scientific fraud. *Flanders Today*. Retrieved from http://www.flanderstoday.eu/innovation/qa-geert-molenberghs-battle-against-scientific-fraud.

Furniere, A. (2015a, 22 September). KU Leuven to battle pressure to publish with biographical profiles. *Flanders Today*. Retrieved from http://www.flanderstoday.eu/education/ku-leuven-battle-pressure-publish-biographical-profiles.

Furniere, A. (2015b, 27 May). UGent first Flemish university with official research ethics policy. *Flanders Today*. Retrieved from http://www.flanderstoday.eu/education/ugent-first-flemish-university-official-research-ethics-policy.

Furniere, A. (2016, 23 November). 70% more PhD graduates in Flanders. *Flanders Today*. Retrieved from http://www.flanderstoday.eu/education/70-more-phd-graduates-flanders.

Geerts, N., Langenhove, M. V., Viaene, P., & Verdoodt, P. (2017). *STI in Flanders*. The Flemish Government Department Economy, Science and Innovation. [D/2017/3241/344]. Retrieved from https://www.ewi-vlaanderen.be/sites/default/files/bestanden/sti_in_flanders_2017.pdf

Ghent University is changing course with a new career model for professorial staff. (7 December 2018). Retrieved from https://www.ugent.be/en/news-events/new-career-model-professorial-staff.htm

Godecharle, S., Fieuws, S., Nemery, B., & Dierickx, K. (2018). Scientists Still Behaving Badly? A Survey Within Industry and Universities. *Science and Engineering Ethics, 24*(6), 1697-1717. doi:10.1007/s11948-017-9957-4

Godecharle, S., Nemery, B., & Dierickx, K. (2014). Heterogeneity in european research integrity guidance: Relying on values or norms. *Journal of Empirical Research on Human Research Ethics, 9*(3), 79-90. doi:10.1177/1556264614540594

Han, A. P. (2017, 31 August). Author cops to “randomly” choosing data for figures in paper, colleagues say. *Retraction Watch*. Retrieved from https://retractionwatch.com/2017/08/31/author-cops-randomly-choosing-data-figures-paper-colleagues-say/.

Hoet, C. (2013, 21 March). Wetenschapsfraude: het failliet van een perfide systeem [Science fraud: the failure of a perfidious system]. *De Wereld Morgen*. Retrieved from https://www.dewereldmorgen.be/artikel/2013/03/21/wetenschapsfraude-het-failliet-van-een-perfide-systeem/.

Honderden Belgische onderzoekers publiceerden in 'neptijdschriften' [Hundreds of Belgian researchers published in 'fake journals’]. (2018, 10 August). *De Morgen*. Retrieved from https://www.demorgen.be/wetenschap/honderden-belgische-onderzoekers-publiceerden-in-neptijdschriften-bc348747/.

Ignaas Devisch: 'Ik ben in de val getrapt' [Ignaas Devisch: "I fell into a trap”]. (2018, 10 August). *Knack*. Retrieved from https://www.knack.be/nieuws/wetenschap/ignaas-devisch-ik-ben-in-de-val-getrapt/article-normal-1183481.html.

In de schijnwerper: wetenschapsfraude [In the spotlight: science fraud]. (2013, 20 August). *VUB Niews*. Retrieved from https://www.vub.be/nieuws/2013/08/20/schijnwerper-wetenschapsfraude.

L'Atelier des Chercheurs. (2014). *Charte de la désexcellence [Charter of desexcellence]*. Retrieved from http://lac.ulb.ac.be/LAC/charte_files/Charte_Desexcellence_1-1.pdf

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868-879. doi:10.1016/j.respol.2017.02.008

Marc Hooghe onder vuur voor plagiaat [Marc Hooghe under fire for plagiarism]. (2017, 21 March). *De Standaard*. Retrieved from https://www.standaard.be/cnt/dmf20170321_02791797.

Mergaert, L., & Raeymaekers, P. (2017). *Researchers at Belgian Universities: What drives them? Which obstacles do they encounter?* Kind Baudouin Foundation. [ISBN: D/2893/2017/16]. Retrieved from https://www.kbs-frb.be/en/Virtual-Library/2017/20171113PP

Meyvis, L. (2014, 25 February). Wetenschappelijke integriteit: verder bouwen aan een huis van vertrouwen [Scientific integrity: continuing to build a house of trust]. *KU Leuven Campuskrant*. Retrieved from https://nieuws.kuleuven.be/nl/campuskrant/1314/06/wetenschappelijke-integriteit--verder-bouwen-aan-een-huis-van-vertrouwen.

Miedema, F., Mayer, K., Holmberg, K., & Leonelli, S. (2018). *Mutual Learning Exercise - Open Science: Altmetrics and Rewards*. European Commission. Retrieved from http://publications.europa.eu/resource/cellar/449cc187-693f-11e8-ab9c-01aa75ed71a1.0001.03/DOC_2

Nieuw geval van wetenschapsfraude aan VUB [New case of science fraud at VUB]. (2013, 30 May). *De Morgen*. Retrieved from https://www.demorgen.be/nieuws/nieuw-geval-van-wetenschapsfraude-aan-vub~bc11c6a8/.

Oransky, I. (2012a, 25 July). Fireworks: Belgian dispute over ovarian transplant findings includes claims of theft, arson. *Retraction Watch*. Retrieved from http://retractionwatch.com/2012/07/25/fireworks-belgian-dispute-over-ovarian-transplant-findings-includes-claims-of-theft-arson/.

Oransky, I. (2012b, 25 June). Following investigation, Erasmus social psychology professor retracts two studies, resigns. *Retraction Watch*. Retrieved from https://retractionwatch.com/2012/06/25/following-investigation-erasmus-social-psychology-professor-retracts-two-studies-resigns/.

Oransky, I. (2012c, 26 July). Ovarian transplant update: Authors of 2004 live-birth follow-up letter ask Lancet to retract it. *Retraction Watch*. Retrieved from https://retractionwatch.com/2012/07/26/ovarian-transplant-update-authors-of-2004-live-birth-follow-up-letter-ask-lancet-to-retract-it/.

Oransky, I. (2013a, 26 November). Cell update: Co-corresponding author let go from Belgian university; retraction notice language changed. *Retraction Watch*. Retrieved from https://retractionwatch.com/2013/11/26/cell-update-co-corresponding-author-let-go-from-belgian-university-retraction-notice-language-changed/.

Oransky, I. (2013b, 22 May). Doing the right thing: Psychology researchers retract after realizing data "were not analyzed properly". *Retraction Watch*. Retrieved from https://retractionwatch.com/2013/05/22/doing-the-right-thing-psychology-researchers-retract-after-realizing-data-were-not-analyzed-properly/.

Palus, S. (2015, 8 December). Diederik Stapel now has 58 retractions. *Retraction Watch*. Retrieved from https://retractionwatch.com/2015/12/08/diederik-stapel-now-has-58-retractions/#more-34952.

Peters, A. (2019). *Public funding of universities, and indicators used in the funding systems : a Euregional perspective.* Université de Liège, Liège, Belgium. Retrieved from https://matheo.uliege.be/handle/2268.2/8014

'Plagiaat? Niets van aan' ['Plagiarism? None of it']. (2017, 22 March). *De Standaard*. Retrieved from https://www.standaard.be/cnt/dmf20170321_02793120.

Pruvot, E. B., Estermann, T., & Lisi, V. (2018). *EUA Public Funding Observatory Report 2018*. European Universities Association. Retrieved from https://eua.eu/101-projects/586-public-funding-observatory.html

Rector Sels on scientific integrity: "No pillory, no witch-hunt". (2018, 25 September). *KU Leuven News*. Retrieved from https://nieuws.kuleuven.be/en/content/2018/sels-scientific-integrity.

Royal science academies of Belgium. (2009). *Codes of ethics for scientific research in Belgium*. Belgian Federal Science Policy Office. [D/2009/1191/6]. Retrieved from http://www.belspo.be/belspo/organisation/publ/pub_ostc/Eth_code/ethcode_en.pdf

Slow Science Belgium. *The Slow Science Manifesto*. Retrieved from https://slowscience.be/the-slow-science-manifesto-2/

Stern, V. (2017a, 26 June). Soon-to-be-ex-rector of top Belgium university blames coverage of misconduct case for ouster. *Retraction Watch*. Retrieved from https://retractionwatch.com/2017/06/26/soon-ex-rector-top-belgium-university-blames-coverage-misconduct-case-ouster/.

Stern, V. (2017b, 5 April). Star pediatric oncologist committed misconduct, ethical violations: reports. *Retraction Watch*. Retrieved from https://retractionwatch.com/2017/04/05/star-pediatric-oncologist-committed-misconduct-ethical-violations-reports/.

Stroobants, K., Godecharle, S., & Brouwers, S. (2013). Flanders overrates impact factors. *Nature, 500*(7460), 29-29. doi:10.1038/500029b

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication Pressure and Scientific Misconduct in Medical Scientists. *Journal of Empirical Research on Human Research Ethics, 9*(5), 64-71. doi:10.1177/1556264614552421

Van de Velde, J., Levecque, K., Mortier, A., & De Beuckelaer, A. (2019, September). Waarom doctorandi in Vlaanderen denken aan stoppen met doctoreren [Why PhD students in Flanders think about stopping their PhDs]. *ECOOM Brief (no. 20)*. Retrieved from http://hdl.handle.net/1854/LU-8630419.

Vanherle, M. (2013, 21 March). Zwaar geval van fraude aan Vlaamse universiteiten [Serious case of fraud at Flemish universities]. *VRT NWS*. Retrieved from https://www.vrt.be/vrtnws/nl/2013/03/21/zwaar_geval_van_fraudeaan vlaamseuniversiteiten-1-1580672/.

Verbeke, R. (2013, 20 March). Fraude bij een op de twaalf medische wetenschappers [Fraud from up to one in twelve medical scientist]. *EOS Wetenschap*. Retrieved from https://www.eoswetenschap.eu/natuurwetenschappen/fraude-bij-een-op-de-twaalf-medische-wetenschappers.

Vervaeke, L. (2013, 22 March). Belgische wetenschapper ontslagen wegens fraude [Belgian scientist fired for fraud]. *Volkskrant*. Retrieved from https://www.volkskrant.nl/nieuws-achtergrond/belgische-wetenschapper-ontslagen-wegens-fraude~b4593e4c/.

Vlaamse universiteiten wereldkampioen sponsoring [Flemish universities are world champions in industry funding]. (2013, 24 June). *De Standaard*. Retrieved from https://www.standaard.be/cnt/dmf20130624_00633829.

VLIR. (2017). *Mission of excellence: facts and figures*. Retrieved from http://www.vlir.be/media/docs/VLIR_fact%20sheets_feb%202018.pdf

VSNU, NWO, NFU and ZonMw drive change in recognition and reward of academics. (2018, 26 November). Retrieved from https://vsnu.nl/en_GB/news-items/nieuwsbericht/489-vsnu-nwo-nfu-and-zonmw-drive-change-in-recognition-and-reward-of-academics.html.

Zacharewicz, K. J. T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment*. European Commission. [ISBN 978-92-79-68715-0 - ISSN 1831-9424]. Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC101043/kj1a27837enn.pdf

Chapter 3

Rethinking success, integrity, and culture in research: A multi-actor take on success in science

An adaptation of this chapter has been submitted to the journal *Research Integrity and Peer Review* and is currently available as a preprint on biorXiv at https://doi.org/10.1101/2020.02.12.945733.

CONTRIBUTIONS

Conceptualization: Noémie Aubert Bonn, Wim Pinxten[1], Raymond De Vries[2]

Funding acquisition: Wim Pinxten. Funding granted by the Bijzonder Onderzoeksfonds (BOF) 15NI05

Project administration: Noémie Aubert Bonn, Wim Pinxten

Methodology: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries

Resources: Noémie Aubert Bonn, Wim Pinxten (general); Raymond de Vries, Melissa S. Anderson[3], and Brian C. Martinson[4] (focus group guide); Ines Steffens[5], Inge Thijs[6], and Igna Rutten[6] (focus group organisation and help in recruiting participants)

Investigation: Noémie Aubert Bonn

Data curation: Noémie Aubert Bonn

Formal analysis: Noémie Aubert Bonn

Visualization: Noémie Aubert Bonn

Validation: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries

Supervision: Wim Pinxten

Writing – original draft: Noémie Aubert Bonn

Writing – review & editing: Noémie Aubert Bonn, Wim Pinxten (intermediate and final versions)

1. Department of Healthcare and Ethics, Hasselt University, Hasselt, Belgium
2. Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor (MI), USA
3. Department of Organizational Leadership, Policy, and Development, University of Minnesota, Minneapolis (MN), USA
4. HealthPartners Institute, Minneapolis VA Medical Center, Center for Care Delivery and Outcomes Research, and University of Minnesota, Department of Medicine, Minneapolis (MN), USA.
5. European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden
6. Faculty of Biomedical and Life Sciences, Hasselt University, Hasselt, Belgium

ABSTRACT

Background: Success shapes the lives and careers of scientists. But success in science is difficult to define, let alone to translate in indicators that can be used for assessment. In the past few years, several groups expressed their dissatisfaction with the indicators currently used for assessing researchers. But given a lack of agreement on what should constitute success in science, most propositions remain unanswered. This paper aims to complement our understanding of success in science and to document areas of tension and conflict in research assessments.

Methods: We conducted semi-structured interviews and focus groups with policy makers, funders, institution leaders, editors or publishers, research integrity office members, research on research integrity community members, laboratory technicians, researchers, research students, and former researchers who changed career to inquire on the topics of success, integrity, and responsibilities in science. We used the Flemish biomedical landscape as a sample to be able to grasp the views of interacting and complementary actors in a system setting.

Results: Given the breadth of our results, we divided our findings in two chapters, with the current chapter focusing on what defines and determines success in science. Respondents depicted success as a multi-factorial, context-dependent, and mutable variable. Success appeared to be an interaction between characteristics from the researcher (Who), research outputs (What), processes (How), and luck. Interviewees noted that current research assessments overvalued outputs but largely ignored the processes deemed essential for research quality and integrity. Interviewees suggested that science needs a diversity of indicators that are transparent, robust, and valid, and that also allow a balanced and diverse view of success; that assessment of scientists should not blindly depend on metrics but also value human input; and that quality should be valued over quantity.

Conclusions: The objective of research assessments may be to encourage good researchers, to benefit society, or simply to advance science. Yet we show that current assessments fall short on each of these objectives. Open and transparent inter-actor dialogue is needed to understand what research assessments aim for and how they can best achieve their objective.

INTRODUCTION

Excellence is a prominent theme in any funding scheme, university mission, and research policy. The concept of excellence, however, is not self-explanatory. Apart from the fact that excellence is hard to define, it is complicated to translate it into concrete criteria for evaluating whether researchers are successful or not in their pursuit of scientific excellence. Nonetheless, in today's highly competitive setting where talent is plenty and money is tight, determining evaluation and assessment criteria is a necessity.

When researchers are being assessed for success, it is important that the criteria used for determining success are compatible with our concepts of scientific excellence. However, with poorly defined concepts of excellence (e.g., Moore, Neylon, Paul Eve, Paul O'Donnell, & Pattinson, 2017) and assessment criteria that raise considerable controversy, there is no guarantee that this is actually the case.

The issue has increasingly attracted the attention of influential voices and fora, which resulted in a growing number of statements and documents on the topic, including the Declaration on Research Assessment (DORA; American Society for Cell Biology, 2013), the Leiden Manifesto (Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015), The Metric Tide (Wilsdon et al., 2015), and more recently the Hong Kong Principles for Assessing Researchers (Moher et al., 2019). In a review of 22 of these documents, Moher and colleagues pointed out that current research incentive systems are open for improvement, particularly in further addressing the societal value of research, in developing reliable and responsible indicators, in valuing complete, transparent, and accessible reporting of research results as well as reproducibility, and in providing room for intellectual risk taking (Moher et al., 2018). As many of the documents mention, however, changing scientific assessment is not straightforward and is likely to face resistance from diverse parties. One of the reasons for this resistance may be the complex inter-actor exchange that governs research and academia. As the European Universities Association (EUA) made clear in a recent report on research assessments, research institutions, funders, and policy makers must "work together to develop and implement more accurate, transparent and responsible approaches to research evaluations" (Saenen & Borell-Damián, 2019, p. 13). But although certain actors such as researchers and scientific editors have been highly involved in the debate on research assessments, other actors have been largely missing from the discussion.

The present research contributes to this discussion by expanding the understanding of success in science and by exploring the connections between success and research integrity. We use the Flemish biomedical research landscape as a lens to study what success means in science, how it is pursued, and how it is assessed. Noticing that most research on research integrity captures the perspective of researchers and research student (Chapter 1), we decided to extend our understanding of success and integrity by including the perspectives from a broad range of research actors. Not only did we involve researchers and research students in our interviews and focus groups, but we also obtained input from policy makers, funders, institution leaders, editors or publishers, research integrity office members, research integrity network members, laboratory technicians, and former researchers who changed career. Our findings, divided in a two-chapter series (see Chapter 4 for our associate findings describing the problems that currently affect academia), resonate with past efforts by suggesting that, in their current state, research assessments may fuel detrimental research practice and damage the integrity of science. In this first chapter, we discuss the way in which different research actors perceive success in science.

METHODS

PARTICIPANTS

The present chapter reports findings from a series of qualitative interviews and focus groups we conducted with different research actors. This qualitative work was part of the broader project Re-SInC (Rethinking success, integrity, and culture in science), the initial workplan is available at our preregistration at https://osf.io/ap4kn/.

In Re-SInC, we captured the views of different research actors on scientific success, problems in science, and responsibilities for integrity. Being aware that the term 'research actor' may be ambiguous, we defined research actors as any person having a role in the setup, funding, execution, organisation, evaluation, and/or publication of research. In other words, we included actors connected to the policing, the funding, the evaluation, the regulation, the publishing, the production (i.e., undertaking the research itself), and the practical work of research, but we did not include consumers of science or end users of new technologies.

We used Flanders as a setting, including participants who either participate in, influence, or reflect (directly or indirectly) upon the Flemish research scene. In most cases, participants did not know the interviewer before the interviews and focus groups. In selecting participants, we aimed to capture the breadth of the Flemish research scene. Using Flanders as a research setting had the advantage of allowing us to capture perspectives from an entire research system in a feasible setting. The Flemish research scene comprises five main universities and a few external research institutes, major funding agencies, a federal research policy department, and one advising integrity office external to research institutions (see Chapter 2 for a greater description of research in Flanders). We chose to concentrate our research on three of the five universities, and to include partnering European funding and policy organisations as well as international journals and publisher to build a realistic but manageable system sample. When participants were affiliated with a university, we focused on the faculty of biomedical sciences. Given the exploratory and qualitative nature of this project, we did not aim for an exhaustive nor a fully representative sample. Our objective was to shift the focus from the narrow view targeting mainly researchers to a broader view that includes a broad range of research actors. Accordingly, we maximized the diversity of participants in each actor group to ensure that each group encompassed a wide range of potentially different perspectives.

Our main actor categories are PhD students, post-doctoral researchers (PostDoc), faculty researchers (Researchers), laboratory technicians (LT), policy makers and influencer (PMI), funding agencies (FA), research institution leaders (RIL), research integrity office members (RIO), editors and publishers (EP), research integrity network members (RIN), and researchers who changed career (RCC). The composition of each actor group is detailed in Table 1.

Table 1. Demographics of participants

Actor group	Abbrev.	Sample description	N, setting, and gender*
Researchers	Researchers	Faculty researchers from the faculty of medical and life sciences of the host institution.	[■ ■ ▲ ▲]
Post Doctoral Researchers	PostDoc	Post-Doctoral researchers enrolled in the faculty of medical and life sciences of the host institution.	[■ ■ ■ ▲ ▲]
PhD Students	PhD	PhD students enrolled in the faculty of medical and life sciences of the host institution.	[■ ■ ■ ■ ■ ■]
Lab Technicians	LT	Laboratory technicians from the faculty of medical and life sciences of the host institution.	[■ ■ ■ ■ ■]
Past Researchers who changed career	RCC	Although this group was not part of our pre-registration, one RCC asked us whether she could take part in our study after seeing the invitation email. After having a chat with her, we realized that hearing the narrative and perspectives of individuals who did research work but decided to leave academia would deeply enrich our results and inform us on problems which are big enough to drive researchers away from research. Therefore, we invited a few researchers who changed careers (i.e., researchers or research students who decided to leave academia) to participate in interviews. In this group, we selected individuals from each of the three universities included in our project, and ensured to have individuals who left academia during their PhD, after their PhD, after their post-doc, and during a tenure track. Recruitment of those participants was helped by recommendations from colleagues who were aware of the profiles we were looking for.	■ ■ ■ ■ ▲
Research Institution Leaders	RIL	We included three Flemish universities in our study. In each institution, we involved several members from the board of directors. These included directors of research, deans, or directors of doctoral schools from the faculties of medicine and life sciences or equivalent.	▲ ▲ ▲ ▲ ▲ ▲ ▲
Research Integrity Office Members	RIO	We included different members from offices in charge of investigating allegations of research integrity and misconduct in three Flemish research institutions and outside research institutions in Flanders (e.g., research integrity officer, policy officers, etc.)	■ ■ ■ ▲
Editors and Publishers	EP	We invited both big and small editors and publishers, and were fortunate to be able to involve journals and publishers with a broad range of editorial practices (i.e., open access and subscription based; published in local language and published in English; focusing on reviews and focusing on groundbreaking empirical findings). To select the interviewees, we first invited a selection of journals from the top twenty highest Impact Factor for 2017 under the category of 'Medicine, general and internal' in the Journal Citation Reports (Clarivate Analytics), purposively picking different publishing models. In addition, we invited select publishers to take part in our research. After conducting individual interviews with a few agreeing participants from this sub-selection, we organized a small focus group with editors of smaller or differing journals, allowing us to involve a great diversity of editors and publishers.	[■ ■ ▲ ▲] ■ ■ ■ ▲
Funding Agencies	FA	We selected national, as well as European funding agencies, making sure to target different funding styles. We made sure to include perspectives from regional public funders, regional private funders, international funders, as well as funders focusing on applied research and funders focusing on fundamental research.	[▲ ▲] ■ ▲ ▲
Policy Makers or Influencers	PMI	In this group, we included both organisations responsible for setting science policy, and organizations which influenced such policies by serving as informers. Consequently, PMIs do not necessarily write nor decide science policies, but may also be asked to provide data which later influences policy decisions.	■ ▲ ▲ ●
Research Integrity Network Members	RIN	We selected a few actors from the research integrity core experts. These included researchers involved with important European research projects on research integrity as well as one actor involved in writing the European Code of Conduct for Researchers.	■ ▲ ▲
			TOTAL = 56 participants

*Square bullets (■) represent female participants; triangle bullets (▲) represent male participants, and round bullets (●) represent participants with undefined gender ('prefer not to answer'). Bullets displayed in brackets represent participants with whom we conducted as focus groups or joint interviews.

It is important to keep in mind that the research world is complex and not organized in distinct actor groups. Consequently, participants could often fit in more than one category, and sometimes felt the need to justify circumstances that would make them fit in the category we selected. Before the interview, we asked participants whether they agreed with the category we assigned them in, and we refined and exemplified the definitions of our actor groups to reflect the participants' distinctions (i.e., further explaining the slight differences between the groups planned in the registration and those used here).

Recruitment

We used several recruitment strategies. For the focus groups with PhD students and researchers, we circulated an email to everyone in the faculty of biomedical and life sciences of the host university and invited them to register on an interest list. We later scheduled a convenient time with those who registered. We used a similar strategy for a focus group of editors and publishers, but circulated the invitation in a conference of scientific editors. For focus groups with lab technicians and post-doctoral researchers, key players helped us recruit and organize the focus group.

For interviews, we invited participants directly via email. We sent up to three reminder emails, but did not pursue further if no response was obtained at the third reminder email. All participation was on a voluntary basis.

Design and setting

We conducted semi-structured interviews and focus groups, meaning that we asked broad questions in an open manner to target main themes rather than specific answers. All interviews and focus groups were audio recorded and transcribed verbatim. Details about the tools used to guide the interviews and focus groups are available in the tool description below.

To maximise transparency, we provide a copy of the COnsolidated criteria for REporting Qualitative research checklist (COREQ) in Appendix 2 and extended descriptions of the interviewer and the setting of the interviews in Appendix 3.

ETHICS AND CONFIDENTIALITY

The project was approved by the Medical Ethics Committee of the Faculty of Medicine and Life Science of Hasselt University (protocol number CME2016/679), and all participants provided written consent for participation, for use and publication of anonymized direct quotes, and for dissemination of the findings from this project. A copy of the consent forms is available in the registration of this project (Aubert Bonn & Pinxten, 2016). We protected the confidentiality of participants by removing identifiers from quotes included in the text. Nonetheless, Flanders is a small research system and given our actor-specific sample, personal identification within quotes remain a risk despite our efforts. To further protect participants' confidentiality and avoid that identification of individual quotes lead to identification of all quotes from the same participant, we decided not to specify respondent numbers in individual quotes, but to refer only to actor groups. If quotes contained personal experiences, we contacted the participants and agreed on the quote with them before publishing.

Following this reasoning, we are unable to share full transcripts, but attempted to be as transparent as possible by providing numerous quotes in the text, in tables, and in appendices.

TOOL

To build our focus group guide, we inspired our style and questions from the focus group guide developed by Raymond De Vries, Melissa S. Anderson, and Brian C. Martinson and used in a study funded by the NIH (Anderson, Ronning, De Vries, & Martinson, 2007). We obtained a copy of the guide after approval from the original authors, and revised the guide to tailor questions to the topics we wished to target, namely 'success in science' and 'responsibilities for research integrity'. We revised our focus group guide several times before data collection and discussed it with Raymond De Vries — expert in qualitative inquiries and part of the team that built the original guide upon which we inspired ours. We built interview guides based on our revised focus group guide. We adapted specific questions (e.g., responsibilities, evaluation) to each actor group, but preserved the general structure and themes for all interviewees. A general version of the interview and focus group guides are available in Appendix 4 and 5. More specific group guides can be provided upon request. All guides were constructed around the following four topics:

i) **Success in science**: What makes a researcher successful? Are these characteristics captured in current assessments? What are indicators for success?
ii) **Current problems** (including misconduct and questionable research practices): Do you have experience with research that crossed the lines of good science? How can we draw the line, what are red flags? Why do bad practices happen? Can they happen to anyone?
iii) **Responsibilities towards integrity**: What is your responsibility towards integrity? Where does it end? Who else is responsible? In what ways are other actors responsible?
iv) If you were granted a fairy wish and could **change one thing in how science works**, what would you pick?

It is important to consider that the interview guide was not used mechanically like a fixed questionnaire, but sometimes shortened, expended, or reordered to capture responses, interest, and to respect time constraints.

Analysis

Recordings were first transcribed verbatim and, where necessary, personal or highly identifiable information was anonymized. We analyzed the transcripts using an inductive thematic analysis with the help of the NVivo 12 Software to manage the data. The analysis proceeded in the following order:

v) **Initial inductive coding**: NAB first analyzed two focus groups (i.e., researchers and PhD student) and five interviews (i.e., RIL, RIO, PMI, RCC, and RIN) to have an initial structure of the themes targeted. In this step, she used an inductive thematic analysis (Elo & Kyngas, 2008) while keeping the three main categories — i.e., success, integrity, and responsibilities — as a baseline. Using the inductive method avoided that we limit our analysis to the order and specific questions included in our guide, and allowed us to identify and note themes that were raised spontaneously or beyond our initial focus.

vi) **Axial coding**: With this first structure, NAB and WP met and took a joint outlook at these initial themes to reorganize them in broader categories and identify relationships between categories. For this step, NAB built figures representing the connections between the main themes, and refined the figures and the codes after the meeting.

vii) **Continued semi-inductive coding**: NAB continued the coding for the remaining transcripts, sometimes coding deductively from the themes already defined in steps 1 and 2, and sometimes inductively adding or refining themes that were missing or imprecise.

viii) **Constant comparison process**: NAB and WP repeated the axial coding and refining steps several times throughout this process, constantly revisiting nodes (i.e., individually coded themes) by re-reading quotes. The nodes and structure were then discussed with RDV to reconsider the general organisations of the nodes. This constant comparison process is common in qualitative analyses and is commonly used, for example, in the Qualitative Analysis Guide of Leuven (QUAGOL; Dierckx de Casterlé, Gastmans, Bryon, & Denier, 2012). This repeated comparison led to a substantially solid set of nodes which later guided further coding in a more deductive manner, though we made efforts to remain open to possible new themes in respect of our inductive analysis.

ix) **Lexical optimization**: Finally, after having coded all transcripts, NAB and WP further discussed the choice of words for each node and reorganized the themes to ensure that they were an ideal fit with the data they were describing. NAB and RDV met to have a final outlook of the general structure and to reorganise the nodes in clean and natural categories.

LIMITATIONS TO CONSIDER

A few points are important to consider when interpreting our findings. First, given the exploratory and qualitative nature of this project, our sample is neither exhaustive nor fully representative. We chose to ask for personal perspectives rather than official institution or organisation views since we believed it would allow us to capture genuine beliefs and opinions and would avoid rote answers.

We thus encouraged participants to share their personal thoughts rather than the thoughts that could be attributed to their entire actor groups, institution, or organisation. We consider that these personal beliefs and opinions are crucial in shaping the more general views of organisations, yet we urge our readers to remain careful when making group comparison and generalisations.

Adding to this first concern, and most relevant to Chapter 4, it is important to consider that when discussing a topic such as research integrity, participants may feel that they have to defend or conceal the practices in place at the organisation where they work, resulting in a possible lack of transparency. Ensuring confidentiality is essential to obtain transparent answers. To minimize risks of identification, we grouped responses by general actor group rather than by individual participants, and decided that any potentially damaging information revealed during our interviews or focus groups would remain confidential, even if it revealed possible misconduct. After the focus group discussions, researchers, research students, and laboratory technician were given a list of contacts where they could safely declare or discuss possible misbehaviours, but the research team preserved full confidentiality on possible misconduct revealed within this project and did not intervene further. Consequently, although we cannot guaranty the accuracy and transparency of participants' response, we ensured that participants felt confident that they could be honest without risk.

As started earlier, it is also important to keep in mind that the research world is complex and not organized in distinct actor groups. Participants could often fit in more than one category by endorsing several research roles. As we mention above, we asked all participants whether they agreed with the category we assigned them in, and we refined and exemplified the definitions of our actor groups to reflect the participants' distinctions. Yet, we must consider each actor category not as a closed box with characteristic opinions, but as a continuum which may or may not hold divergent views from other actor groups. Our findings help capture views that may have been overlooked in past research which focused on researchers, but should not be used to discriminate or represent the opinions of entire actor groups.

As pointed out by one of the jury members, it may also be important to reflect on possible 'survivor biases' — biases caused by interviewing participants who already survived the research system. This point is very relevant and should be taken into account in any research using researchers as participants. In our results, we consider that the inclusion of researchers who left academia as well as early career researchers may counterbalance this bias. Yet, it is important to understand that differences between groups that we tend to attribute to seniority may in fact be the result of these survivor biases.

Finally, it is important to consider that given the richness of the information gathered, certain findings may be displayed with greater importance than others simply based on the authors' personal interests. We were careful to include also the views we disagreed with or found to be of limited interest, yet it is inevitable that some of the selection and interpretation of our findings was influenced by our own perspectives. To maximise transparency on the genuine views of our informers, we supplement our interpretation of the findings with quotes whenever possible.

RESULTS

The purpose of this chapter is to retell, connect, and extend on the issues that the different actors raised in our study. Aiming to maximise transparency and to minimise selective reporting, we provide numerous quotes and personal stories to illustrate our claims. The result, however, is a lengthy chapter in which we explain the breadth of the concerns raised by our participants. Given the length of the resulting chapter, a short summary of results is available at the end of the results section, and select findings are re-examined and extended in the discussion.

RESEARCHERS' PERSONAL SUCCESSES

Before reporting on the views of all interviewees on research success, we believed it would be important to look at the answers of researchers and research students. Focus groups with researchers and research students comprised an additional question in which we asked participants to describe their personal satisfactions and successes. Given the limited number of researchers and research students involved in our research, it would be naive to infer that our findings represent the breadth of researchers' view on success. Nevertheless, we believed that capturing what researchers and research students describe as 'satisfying' was important to understand and contextualise the general perspectives of success in science.

In their answers, interviewees described several factors which made them feel satisfied or which they interpreted as personal success. First, PhD students and post-doctoral researchers strongly supported that *making a change* in practice was something that was central to them.

> *"I agree with the fact that that feeling that something is done with what you found is crucial for your own feeling. […] I think that's crucial. Even more than the publications…" (PostDoc)*

> *"Yeah it was part of my motivation to give something back to the clinical field by doing research." (PhD student)*

For PhD students, realising that their results would remain theoretical or would be too small to make a difference was raised as one of the disappointment they faced in research.

> "Participant A: *If I can help people by doing this project, that gives me a lot of satisfaction I think.*
> Participant B: *That's true but that was also my first idea when I started, but I have to be honest, my project is so fundamental that I'm almost finishing up, and I don't see anything that will be going to the clinic for years or something. So at that point for me it was a bit disappointing, because... Ok, I wanted to, but I'm so fundamental, basically really molecular stuff, that I don't see it to get really... […]*
> Participant C: *Yeah I think for me it's the same. Because I'm working on a project that's like this very tiny subset of a subset of [specialised group of] cells. And then at the beginning you think 'I'm going to change the field with this research', but yeah I don't know." (PhD students)*

Although some researchers also supported that translating their findings in practice was satisfying, they acknowledge that *theoretical knowledge* or simply following their *curiosity* became their "*main drive*", or at least provided its share of satisfaction.

> *"For me it's good if it goes this direction* [i.e., is translated in practice] *but also just creating new knowledge which doesn't really directly impact people, I think is also very very interesting, or I'm also very passionate about that. So it shouldn't always have an implication." (Researcher)*

For researchers, external satisfactions, such as *peer appreciation* or *fulfilling institutional requirements* were "*also very important*" to personal satisfaction, but as secondary aspects which were not enough for feeling completely satisfied.

> *"I also have some… still some criteria which I have to do that I also think about those things. But I don't feel bad about it that it's my only drive for some things that it's just publication. On the other hand, I also feel that I cannot be satisfied alone by those things."* (Researcher)

Finally, post-doctoral researchers added two intriguing dimensions to the concept of success. First, they stated that successes are personal, and that each researcher will likely be successful in different ways. In this sense, personal success was seen to reflect aptitudes and skills in which individuals excel, rather than a universally shared idea.

> "[In my group, we don't have strict requirements]*, and I think it's very beautiful because we have* [dozens of] *PhD students and they're all — or 99% of them are — successful, but they are so different in being successful. Some are really being successful in the number of publications, some of them are really successful in the network they have with other companies, with other research institutes, some of them are really successful in the perseverance to do something really new and to make it happen, only if there's a small study on 20 patients, but it's so new and they will really make it happen in the hospital. So, they're so successful on so many different levels and I really like the fact that we don't judge them all in the same way because they can be themselves and be successful in the way that they want to be successful." (PostDoc)*

The need for diversity of successes was thus valued, even though it was acknowledged to be a rare feature in research assessments. A second intriguing dimension, also raised by a post-doctoral researcher was that even within individual researchers, personal conceptions of success may be mutable, likely influenced by career stages, work environments, and expectations of others.

> "Participant A: *For me I think my idea on what success is is* ***changing a lot*** *of course. When you're a PhD student you just want a breakthrough in your project, that's success, and then by the time you're finishing your PhD you're looking at what… Is there a possibility for PostDoc then you realise 'OK they're counting publications, they're doing this' and then you're looking around and then you sometimes get this mixed feeling of someone who you feel was not very creative or did not have to do a lot of work themselves, it was very guided and clear steps, and they have a lot of publications and so they get a PostDoc position. And then that's sometimes difficult, and you think like 'How does this work here?' […] then I went into more research coordination and then I was in a [different group] and then it was all the time about metrics. Because the money was divided by metrics,*

> *and it was like publications and project funding and… And then I felt like everything revolved around that. It wasn't important anymore like what projects we're doing as long as it was a project on that funding channel because that counted higher on the metrics and… So ok, and* ***then you're really like that****. And now being here in this setting I'm really seeing the impact of research.* ***Now it's changed again****. Now it's really like that kind of research where you can make a difference for an organisation, for patients… That's the thing that's success. And I think that maybe like you say that in the long run that's what you have to do. But it's kind of the short-term mechanisms, and not always…*
> Participant B: *Yeah, I think that* ***the definition of success is highly dependent of the institute and the environment you're in*** *like you're mentioning. And if you're constantly told 'This is how we measure success' then…*
> Participant A: *Yeah, so* ***then you're really guiding yourself to get those key indicators****." (PostDoc, bold added for emphasis)*

In other words, interviewees revealed that personal success was a mutable variable which could change depending on contexts, demands, and career stages.

INTER-ACTOR VIEWS ON SUCCESS

Now that we have glanced at the perspectives of researchers and research students, let's look at the views of all research actors on the more general idea of success. In order to avoid rote answers, we asked about success indirectly. Rather than asking 'What is success in science?', we asked interviewees about 'What makes researchers successful?'.

In their answers, interviewees mentioned several factors which they believe are essential or useful in becoming a successful researcher. We classified these factors in four main categories: factors visible in the *researchers* themselves (Who), factors from the research *process* (How), factors from the research outputs (What) and, unexpectedly, factors related to *luck*, which was thought to play an important role in success. Figure 1 illustrates the different categories we captured.

Who

Researcher. Several features related to the *researchers* were considered important in determining and yielding success. While all these individual factors were said to play a role in producing success, they were also described as indicators to look for when selecting researchers for a position, thereby influencing careers and promotions. Among those, participants highlighted *personal traits*, such as ambition, passion, rigorousness, and intelligence, as well as acquired *skills and expertise*, such as business potential, management skills, writing skills, and scientific expertise. Certain respondents also believed that success could be influenced by specific *situation* in which individuals find themselves. In this regard, gender and ethnicity were mentioned as possible obstacles — through pregnancy leaves, family obligations, prejudice, or language inequalities — or advantage for success — through employment quotas. Along the same line, childlessness and celibacy were mentioned as advantages for yielding success since they allowed researchers to devote more time to their work.

Beyond the advantage that extra time and flexibility could provide, they were sometimes considered as conditions to a successful research career. Indeed, some interviewees believed that researchers and research students *should* be able to devote themselves to their career by being mobile and by working beyond regular schedules and conditions.

> *"I think people have to realize when you do a PhD, it's a stressful thing, you really are going to get the highest degree there is at a university, it doesn't fit between 9 and 5." (RIL)*

> *"…being passionate about science is almost like being an artist. You live in poverty because you want to pursue your art." (PMI).*

> *"That's also what we ask for, excellence for people when they come here. […] Usually those people need to have been abroad for at least six months. But if it is two years it's better. So these are important factors to create excellence." (RIL)*

Many of the researchers who changed career mentioned that the expectation that they should sacrifice family life and private comfort for science played a role in their decision to leave academia. We will explore this idea further when discussing unrealistic expectations in the associate chapter (Chapter 4).

Finally, the network and status that researchers bring along with them was also seen as determinant to success. Having an established network and personal recognition from peers was thought to be key to success.

Figure 1. Main themes captured as determinants of success in science.

WHO **researcher**

Situation
Skills and expertise
Personality traits
Status and network

WHAT **outputs**

Publications
Past successes
Impact metrics
Applicability and social value
Teaching and services

HOW **processes**

Collaborations / multidisciplinarity
Disinterestedness
Appropriate methodology
Adherence to ethical standards
Good research idea
Innovation and creativity
Feasibility
Focus
Openness and transparency
Reproducibility
Science communication

OTHER

Luck

What

Outputs. Indicators which provide information about what researchers have accomplished were univocally considered crucial in determining success. Among those, high academic grades, past success in obtaining funding, publications, and publication metrics (e.g., impact factor, citations, H index) were mentioned as currently being used for determining success, although not all interviewees agreed on the individual value of these determinants. In addition, less traditional products of research were also mentioned, such as the applicability and societal value of the research findings and the researcher's involvement in teaching and services (i.e., mostly referred to as serving on institutional boards, committees, and scientific societies).

How

Processes. But features which indicate 'how researchers work' (i.e., *processes*) were also deemed integral to success, regardless of the output they generate. On the one hand, some processes were thought to play a part in the success of individual research projects. Collaborations, multidisciplinarity, appropriate methodology, adherence to ethical requirements, good and innovative research ideas, feasibility, and focus were all viewed as pathways to achieve good outputs and related successes. On the other hand, respondents also identified a number of processes which they considered impacted beyond individual projects and were essential to the success of science at large. Openness and transparency, for example, were repeatedly viewed as important aspects of the collegiality which promotes the success of science as a common goal. One interviewee explained that openness was "*very important to help the research enterprise because it's really about facilitating the fact that other people can build upon a research"* (EP). Along the same lines, reproducibility was qualified as the "*most important thing*" (RIL) and as "*a very important element in science*" (PMI). Yet, interviewees noted that reproducibility is "*often lacking*" from research (PMI), and that replication studies are under-appreciated in current success assessments (Researcher) or even possibly wasting research money (RIL). Finally, public engagement, mainly in the form of communicating scientific findings to the public, was also mentioned as part of the broader scientific success by building trust in science and by potentially contributing to the quality of research.

What truly differentiated *outputs* from *processes* was the perspective that the latter contributed to success regardless of the final result.

Other

Luck. Interviewees also attributed success to *luck*, a feature which transcended outputs, processes, and individuals. In our analysis, we discerned three different definitions of what it meant to be 'lucky' in science. First, researchers could be considered lucky if they worked with distinguished colleagues or in established labs, given that such settings maximized the opportunities for obtaining high end material, publications, and grants. This first meaning brings back the idea of the network that researchers bring with them, and adds an element of arbitrariness to the control that researchers have in building their network. Second, luck was also employed to refer to unexpected evolutions and trends, such as working on a topic which suddenly boomed in visibility and media attention or being "*somewhere at the right moment at the right time*" (FA). In this second signification, luck was perceived as something that one could partially create, or at least grasp and maximize. Finally, luck was sometimes attributed to the output of research results, with positive findings being lucky, and negative findings being unlucky. In this last sense, luck was a factor that was out of researchers' control and independent of their skills. In all three senses, luck was both described as something that had helped mediocre researchers move ahead in their career and as something that had wrecked the success of otherwise talented researchers.

CURRENT ASSESSMENTS OF SUCCESS

In the above section, we describe the different features which were used to describe success in science. Although this broad array of features describe the overall picture of success our interviewees revealed, current research assessments do not necessarily value these elements equally. In answering our question about 'What makes researchers successful?', several interviewees spontaneously identified the tension between what currently determines success

— through formal rewards and recognition — and what they believed *should* determine success in science.

> *"Hm… What the current situation is, or what I think success should be? (laughs)" (EP)*

> *"I think that you have different views on looking on it. You have the measurable parts, and you have the non-measurable part. And I think that these two are sometimes in contradiction." (RIO)*

> *"...I started this PhD project because I wanted to have results useful to clinical practice, and I said "I want to do this". And [my supervisors] were already saying for a year "No, no, it's not interesting, no we shouldn't do that." and I said "I want to do this, or my project failed for me." [...] Ok, I know it's not going to be so big that it's so interesting for journals, but I think for our clinical field, for Flanders, it's important that we do a study like that. And it was… that was the chapter that people from clinical practice were most interested in too. So... I think when you ask us 'What is success in research', we've got our own points of success, and what we know that's expected from us by the system. So those are two different lists. (laughs)" (PhD)*

More precisely, assessments of success were described as currently focusing on research outputs — generally measured through rigid metrics and quantity indicators — while largely ignoring other important features from the process through which science is performed. In this respect, several interviewees mentioned that although output-related successes helped researchers advance their career and made them feel satisfied in some way, they also felt a lack of reward for processes which provided an indication about the quality of the work, its usability for the scientific community, and the quality of science as such.

Despite a general agreement on the need to reintroduce processes in research assessments, respondents sometimes contradicted each other when we asked them to give precise examples of outputs that bothered them, or to describe processes they thought should be valued more in research assessments.

Disagreements on outputs

Publications. The emphasis on publication in research assessments raised such a disagreement. Some respondents considered that relying on publications to measure success was problematic and even damaging for science, while others saw publications as a necessary and representative measure of scientific success. We distinguished three arguments against, and three arguments in favor of using publications as the main indicator for assessing research success (Figure 2). Illustrative quotes are available in Appendix 6.

Figure 2. Summarising scheme of arguments for and against using publications as the main indicator of success

The first argument against using publications as a main indicator of success was based on the idea that publications, constitute a *reductionistic* measure of success. In other words, using publications as the main measure for success ignored "*other very important contributions to the scientific enterprise*" (EP). Additionally, the reductionistic scope of publications was said to sometimes unfairly disadvantage researchers who "*have the qualifications to be good researcher*" (PhD student) but are simply unsuccessful in publishing their results.

The second argument against focusing on publication for evaluating success resided on the belief that publications are an *arbitrary* measure which does not represent merit, efforts, and quality. Researchers and research students in particular worried that publications often resulted from arranged connections rather than from high scientific value or efforts, an argument we will discuss further in the associated chapter (Chapter 4). Researchers and research students also supported that highly cited publications in recognized journals were not necessarily of high quality, making the link between papers and quality arbitrary. Adding to this, several interviewees supported that publications could be a mere matter of luck. While such reflections may support the need for journals to focus on quality and to reduce publication biases as much as possible, one editor or publisher rather explained that research assessments are the real issue which needs to be resolved. According to this participant, basing research assessments on journals' editorial decisions is opening them to biases which are embedded in the publication system and are unlikely to change (e.g., publishing controversial articles or impressive results to increase readership and visibility). To be realistic, efforts should focus on changing research assessments and their exaggerated reliance on select scientific publications, rather than expecting that journals change themselves.

> *"…the problems that I was describing with the ecosystem, it's not necessarily… It's not happening at the journal level necessarily. Journals do what journals do. They select papers based on subjective criteria that are specific to each journal, and that's their editorial mission. That's what they do. The fact that only some of these journals and publications count as real measures of success is… This is a problem of research assessment." (EP)*

As a third argument against focusing on publications to evaluate researchers, interviewees worried that the increasing dependency on publication output (i.e., the publish or perish culture) may introduce *perverse* incentives which might threaten the integrity of research. On the one hand, publication pressures may tempt researchers to engage in questionable practices to maximise their publication output. On the other hand, the emphasis on publications may shift the main objective of research projects towards 'sexy' and 'publishable' topics rather than topics that are 'interesting' or 'relevant' to advance science.

Despite these three arguments against the focus on publications for evaluating success, several respondents also identified arguments in favour of this focus, sometimes directly opposing the arguments introduced above. First, publications were described as a *necessary* aspect of scientific advancement, and the emphasis that evaluations give to publications was seen as a way to ensure that researchers keep publications in the forefront of their priorities. Second, some respondents described publications as *representative* indicators of good research and merit. In fact, considering that publications are the endpoint of an extensive and difficult process which could not happen without hard work, some argued that publication outputs helped identify good researchers. Finally, publications were also described by many as the only tangible way to value and *measure* science, which added to the credibility of research assessments.

Impact factor. Similar conflicts were observed when looking at the impact factor. First, the impact factor was acknowledged by many as a being useful for its *measurability*, its *simplicity*, its *acceptability*, and even — as some mentioned — for its perceived correlation with the *quality* of the review process.

> *"So if you hire a PhD student, but even more if you hire a PostDoc or a young professor, then they evaluate it of course. And then, the bibliometric parameters are much more important so you look at the number and the quality of publications.* ***How do you measure the quality, of course the impact factor****. So if somebody with a Nature paper comes of course this person is considered to be more 'valuable', in quotation marks, and gets a higher score in the end and probably the job, compared to a person with a publication record which has lower impact factors. So impact factors are still very important, and grants..." (RIL, bold added for emphasis)*

> *"Of course what you always want to have is one of the two champions that are really picky in the graph, but I think for us it's also important to really see that the whole group is evolving to improved* ***quality as measured by the impact factor*** *and of course I know the discussion that this is only one way to look at quality, but* ***it's still the most accepted way to look at quality*** *I think, in our field." (RIL, bold added for emphasis)*

> *"I have to say that generally* ***there is a big correlation between the impact factor and the quality*** *of the content…" (EP, bold added for emphasis)*

> *"OK, when we select something, somebody for an academic position, we will look at publications, at the numbers, and below 10 you will never get something in an academic position, below 10 papers. Of course, suppose somebody comes with two Nature, one Lancet, and one NEJM, then we have to re-think. So... In a way, today it's still a balance between numbers and impact factors, it's still playing a role. But the whole issue is that there is something which goes together.* ***A journal with a high impact factor has to improve its review process. Because you cannot keep your high impact... I think that when you send your paper to Lancet or NEJM, you will have tough review. While when you send it to a low impact factor journal, [...] you can send a completely fake paper to reviewers who will judge it perfect and let it publish.****" (RIL, bold added for emphasis)*

Nonetheless, using the impact factor as a measure of success yielded overwhelmingly negative responses, even among participants who believed it served as an indicator of quality. Most participants mentioned that the impact factor was not adapted to their disciplines, that it was not representative of the impact of individual papers, that it was open to biases and manipulation, and even that it could disrupt science by discouraging research in fields with traditionally lower impact factor.

> *"I think [current metrics are]* ***far too simple****. You know like* ***impact factor is useless I think in evaluating the importance of an individual paper****, because impact factor relates to a journal. So it's not an article level measure of any kind." (EP, bold added for emphasis)*

> *"Publishing is important but* ***I hate the impact factor thing****. I would more look into the quartile thing, if you are in a* ***field*** *that has low impact factors but you are in the top ten of your field, that's just fine. I mean it doesn't have to be Nature, it can also be [a small specific journal], if that's your top, in your field. So I think there is a tendency towards going that way but I like that a lot more than the impact factor shizzle, yuck!" (RCC, bold added for emphasis)*

> "Interviewer: *Which [indicators] do you think are the most toxic and less representative of quality?*
> Participant: *The urge to publish in Q1. […] I understand that there needs to be an impact factor, but the whole issue of the weight of an impact factor in the personal career of a researcher… because then I would advise anybody who wants to go in research "Please go in cancer research". Try to get to Lancet cancer or whatever other journal, of NEJM and then you're safe. Don't do anything like plastic surgery or [smaller topics]... So* ***that's one of the most toxic factors I think****. The pressure of... Because the* ***impact factor is not reflecting really the importance of the research****. You could say that cancer is of course important, and then you see that for instance [the biggest journals in other discipline] which has an impact factor of 16, they only publish on cancer [...] and they* ***manipulate the impact factor****. Of course, because when you, as an author, you don't have enough references referring to their own journal you get from the reviewer report that you need to put those in... (RIL, bold added for emphasis)*

> *"Well the problem with the impact factor as a standard, most appreciated metrics, even though we don't want to do that [laughs], is that it is* ***not essentially an indicator of quality neither of the article, neither of the journal****, but why? Because there could be less articles of lesser quality, published by renowned scientists in higher impact factor journals, and you can have a good research from scientists coming from some small country and who is not so famous internationally, and he will not, or she* ***will not be able to publish in the higher impact factor journals because they are usually biased****, and I know because I come from a country, when you read someone's last name you usually… they can know that you are from that country [laughs]." (EP, bold added for emphasis)*

A few interviewees proposed that direct citations would be more relevant in personal impact assessments, but they also acknowledged that determining the impact of individual articles using direct citations could take years, if not decades in some disciplines. Furthermore, researchers added that their most cited paper was not necessarily the one they considered most important, and that citation counts tended to refer to novelty and timing rather than to quality. Consequently, despite an overwhelming aversion towards using impact factors for scientific evaluation, concrete alternatives were more difficult to nail down.

Disagreements on processes

Science communication. The importance of science communication also raised conflicting opinions among interviewees. Many supported that sharing science through popular channels such as Twitter, YouTube, Facebook, or Wikipedia should be considered in career evaluations. For example, one respondent in the PMI group noticed that "*Researchers who do a lot of work on Wikipedia are not rewarded for it, but they're doing a lot of good work!*". The same interviewee however, later warned that appearing in the media was different than actively making the effort to communicate science, "*because then the media decides who is successful*" and "*a lot of researchers will also be successful and you will never hear of it*". A few researchers mentioned that science communication was essential to maximize the interdisciplinary impact of one's work, and that presenting findings in broad conferences and participating on Twitter could foster this interdisciplinarity. Other respondents even regarded science communication and the ability to simplify and share one's findings with different stakeholders as a core requisite for the quality of research.

> *"I mean we work a lot on, or we try to promote everything which has to do with public engagement and science communication, all these things, but if you're not able to explain to lay people what your research is about… [shakes head meaning it's not a good sign]. I think it's a sort of, how do you call it, a litmus test in a certain way […] Sometimes sort of public engagement… the arguments are sort of normative. You have to do this because you're working with public money and you have to be accountable or… which is ok, but, I really believe it's better than that. It's more important than that. I really believe it's good to discuss with philosophers, with ethicists, with citizens, with patients… For the quality of your research. To be stretched. It's another type of checks and balances than the ones which are done in peer review. It doesn't replace peer review, it's just another level. To look at the relevance, to, yeah… to be confronted with questions you probably haven't ever asked… To be better in communicating which… Better communicating will help you better thinking. I mean I think there is a lot of quality gains." (FA)*

But although some perceived science communication to be an essential component of quality work, others saw it as a component which did not indicate the quality nor the efforts invested in the research. Some researchers even thuoght that the quality of science might be threatened by the lack of quality control of social media.

> *"I feel there is also a kind of danger in those things, because for example I follow some researchers on Twitter, which have a very... I feel that they're on Twitter all day long I'd say, and everybody follows them... But it's not... the research is not always that good, but because of the fact that everybody is following, this is going to be the new reality, and I start to... yeah... These things, worldwide, have <an> impact on the research impurities, and it's shifting towards.... yeah, it's not controlled, the quality of those things." (Researcher)*

This perspective was echoed by a participant from the RIO group who admitted having faced substantial resistance from researchers when presenting an action plan meant to promote and value science communication in her institution. This RIO received responses such as "*yeah... that's the one who's always with his head in the newspapers, but is he writing A1s*[5]*?*", and concluded that researchers might not be in favour of such a shift. In sum, even though science communication is an important aspect currently put forward in new evaluation processes and policies, researchers do not all agree on its value and its impact on the quality of science.

Openness. Openness also raised diverse thoughts from our interviewees. Although most agreed that open science and transparency were important or even "*necessary for the community of researchers"* (PMI), some doubted that open science would help foster integrity, proposing that it might simply bring a "*different level of cheating*" (RIL). We also understood from researchers and especially from research students that the fear of 'being scooped' was still too vivid for openness to fully happen, at least before publication. PhD students

---

[5] A1's are a category of publications in research assessments in Flanders. They relate to articles included in Web of Science's Science Citation Index Expanded, Social Science Citation Index and/or Arts and Humanities Citation Index, whose document type is labelled as "Article", "Review", "Letter", "Note" and/or "Proceedings Paper"; or in journals included in the Journal Citation Reports of Web of Science.

expressed frustration but also helplessness towards their will to be more open, admitting that the risks of losing their data tended to overcome their will to be more open and that opening their work was often discouraged by their supervisors.

> "Participant A: *Yeah we are now trying, or in our group someone is trying to put up a database for all of the data on [our topic]. But then researchers would need to hand over their data to make it accessible. And there is a lot of discussion about it, if people would be willing to do that, to hand out your unpublished data... I think it will help the research, and it will help patients, but I don't know if everyone is willing, I don't know if I would be willing to, just put it in...*
> Interviewer: Would you all… would you be willing to put your data in a server?
> Participant B: *I had the question once, but by a supervisor, and we're PhD students so you asked, he said 'No, no we're just going to publish first and then when we did that then we can say here's the data'. […]*
> Participant C: *The problem with research is also it's really a competition in research. I also have it now that I can't present on a congress because there are only three articles published on the subject I'm studying, so the supervisors are scared if I make a poster or I present that other researchers will get interested in the same topic, and then, if they publish first all I'm doing is a waste of time... not exactly waste, but... yeah... so I think in research you really have a lot of competition because some people are focusing on the same subject and the data is not published so they will be first, and they want to be first, and... that's a problem with research. And I think it's also a problem that no one wants to share their unpublished data because they are scared that someone else will go and take the data and will publish first and then, you don't have it anymore.*
> Participant D: *Yeah I completely understand the feeling because what we are doing it's also new so it's never been done and my promotor is always so reluctant to let me go and show the data to other people. […] he is always so scared that other people are going to steal his ideas... Sometimes I do understand, but sometimes I'm also like, I don't really like this kind of environment, it struggles with my personality a lot, I think." (PhD students)*

Beyond open data, issues surrounding open access were also brought up in our interviews. We noticed that PhD students, who are directly affected by the inability to access research articles, strongly supported open access. Some university leaders also encouraged open access and criticized the monopole of big publishers, noting that we faced a growing problem where subscriptions may become "*unpayable in the long run*" (RIL). In this first perspective, the subscription costs that university libraries (and hence universities) need to invest in closed-access publishers was pointed as the issue. Conversely, other university leaders stated that, for financial reasons, they would not advise their researchers to publish in open access journals. In this second perspective, the resources needed to finance the article processing charges of open access journals were considered problematic, possibly because these charges are often deducted directly from the budgets of departments and are, therefore, more visible to the faculty leaders. This second institution leader further perceived the model of open access as biased towards accepting papers regardless of their quality.

> *"That's another big issue. That's the open access eh? The model of the open access is unfair because the journal makes the profit by publishing because the author has to pay. So I think the review process is probably more biased. […] I believe. I think that... OK, there is, in my very small field, there is some open access journals, and I feel like whatever review you do they all get published because they get the money. So, is that a good solution? No I don't believe it's a good solution.*
> Interviewer: *Yeah. So would you not advise to your researchers to...*
> *Participant: I don't, no, because we don't have the money (laughs)" (RIL)*

One editor or publisher explained that bad publicity surrounding open access journals may come from the unfortunate reality that the open access model "*opened the door for a number of the so called predatory journals*". Nonetheless, this interviewee also declared that "*at the end of the day, whether the article is robust and well evaluated is not a function of the business model of the journal, it's a function of the editorial process of the journal*".

Other indicators raised polarized views, such as the need for societal benefit, and the need for focused areas of expertise. In sum, respondents agreed that current research evaluations were sub-optimal, that they valued quantity over quality, and that they relied disproportionally on outputs while ignoring the processes that provide essential information on the quality of the research. Nonetheless, disagreements on the specific indicators which should and should not be used in attributing success suggest that solutions are far from simple.

OPTIMIZING RESEARCH ASSESSMENTS

Despite persisting disagreement on the content of good research assessments, several respondents proposed concrete recommendations on the form research assessments should take. Four main characteristics were put forward as essential for fair and representative evaluations (See Table 2 for sample quotes representing the four criteria).

Diversity of indicators. First, many interviewees mentioned that it is essential to use a diversity of indicators to be able to measure different aspects of research. Many respondents worried about the current overreliance on outputs (publications and impact factors especially). Interviewees believed that relying on one or few metrics generated important biases, opened the door to manipulation, and ignored important processes which relied both on different metrics and on other types of evaluations, such as openness, societal impact, or science communication.

Human input. Second, respondents also believed that it was necessary to have *human input* — in the form of peer review — in the evaluation process to capture what some called a holistic view of success. Peer review was, however, said to also share important weaknesses which must be taken into account. Among those, (i) the *potential for conflicting interest* (especially worrisome to researchers and students who perceived that funding depended more on status and network than on the quality of the project proposed), (ii) *conservatism* (an issue we will explore further in Chapter 4), (iii) *subjectivity*, and (iv) *costs*[6] were mentioned. One research funder proposed that repeating evaluations in different contexts, institutions, and with boards of mixed affiliations could help balance these problems. Another respondent proposed that, to reduce the costs and increase the availability of peer-review, peer-review itself should be rewarded in research assessment.

> *"Why shouldn't people be given credit for doing this kind of work? It's really important work, it keeps the whole academic system alive. So I think it's crazy that it's not included as a, you know, a metric, a possible metric or an indicator of being a successful scientist!" (RIN)*

[6] This aspect was raised when discussing external expert panels for integrity issues, in which one RIO mentioned that "*They're paid for leading the report and making the preparations, you have to bring them to your university, you have to put them in a nice hotel, obviously, you have to dine, I mean, the amount of money is just enormous. And then you have... ok what is coming out of this? You have some remarks... […] Yeah I'm a bit critical towards that system*".

Table 2. Characteristics for good evaluation

Characteristic	Sample quote	Actor
Diversity of indicators	*With metrics I think there is an important rule to keep in mind is that if you're going to use metrics, you need to use many of them. And you need to really understand what they mean and whether they answer what you're looking for.*	EP
	And maybe also, and I think that the idea of taking other impacts into account can be helpful. I have no one solution, but I think this can be helpful.	FA
	I think that you have to have different parameters. I think that's important. Not to focus on just one or a few, but have different parameters that focus on different aspects and put these together together with alternatives.	RIO
	…you need to use [metrics] in combination in terms of other indicators, you need to use what we say 'a basket of metrics', you cannot evaluate people just based on one single metric. I would argue that you need several, and then of course you have different metrics evaluating different things. One thing is valuating excellence. Impact factor is going to be among that one. But then when you evaluate the education part, the capacity of someone to be a good professor, you need different educators also. So that's the first one. You need… It's never in isolation.	EP
Human input	*I don't think you can rely on one or several indicators without human input. You can't make sense of a number on its own.*	EP
	That's also why there is not one penny of research money allocated in the university that is not based on peer-review. Everything is based on peer review. So every proposal submitted is based on peer review	PMI
	So you can have indicators, publications of course is a good indicator, but it's an indicator. You should also do, let's say an holistic, what we call an holistic approach, have an holistic approach.	FA
Quality over quantiy	*I think they should evaluate again true quality in terms of… And that isn't done easily via metric, one metric, you know. It's actually… You know I would suggest, it's just up the top of my head… That for any appointment, the people say to a researcher 'OK, please choose your best two or three papers in the past five years, and then two additional ones if you want, where there's no time limit. You know, something that you might have published 20 years ago, but is really important. And you submit that with your application. And then people… And you say to us why it's important. And then people need to of course evaluate that. So that would get away to, that would take into account the more longer term strategy of someone, but also… But it needs a qualitative.*	EP
Transparent (robust and valid) indicators	*I think that it needs to be transparent, robust, validated, etc. […] Everybody can see the methodology, they can reproduce it if they want. […] As long as you describe really well what the rankings take into account, and why you are first or hundredth… What would be bad is if you rank and you don't tell people on what basis you rank. There you go. You might disagree with the indicator, you might disagree with the ranking, but as long as it's transparent, well validated, robust, etc. There you go.*	EP

Quality over quantity. Third, the importance of evaluating the quality over the quantity was raised many times by different research actors. Many proposed that presenting only a subset of the most relevant work (e.g., three papers most important to the researchers, and why) could help by permitting in depth evaluation rather than reliance on quantity and metrics. Nevertheless, funders and policy makers mentioned that despite criticism from researchers about the over reliance on quantity, peer reviewers — generally researchers themselves — often asked for the full list of publications, the H index, or other quantifiable indicators when evaluating proposals, even when the proposal was purposively adapted to contain only a subset of relevant work. Overcoming this quantifiable culture thus seems to be a must for initiating a change.

Transparent, robust, and valid indicators. Finally, the transparency, robustness (consistency between evaluations) and validity (measuring what is intended) of indicators were also mentioned as a requirement for good evaluation. These last criteria are basic criteria for any reliable metric, yet they are not always met by newly proposed indicators, and the way current indicators are used sometimes compromises the validity of the intended measure (e.g., assessing *quality* of *single* publications using the impact factors, which qualifies *journal* average *citations*). Added to these four essential characteristics, the importance of being consistent in how evaluations are conducted while considering differences in fields and disciplines were often raised by interviewees.

A WISH FOR CHANGE

At the end of our interviews, we ask participants what they would do if they had a 'fairy wish' to changes anything in science. In other words, we ask them to describe one aspect of science they believe need priority for change. Although not all answers targeted research assessments, the majority of respondents discussed changes relating to research assessments or research funding as their 'fairy wish'. In changing research assessments, the need to value quality over quantity, to reduce output pressure and competition, and to broaden and adapt indicators of success to reflect not only the output, but other aspects of science were mentioned. In changing research funding, the need for fairer evaluations and distribution, including the suggestion that resources are not distributed based on assessment but rather equally distributed among scientists, and the wish for long-term funding schemes and baseline research allowances were 'wished' from our participants. Appendix 7 illustrates these ideas with a selection of quotes from diverse participants.

Short summary of findings

Our investigation of the perspectives of success in science reveals that the way in which we currently define science and the way in which we assess scientific success generates conflicting perspectives within and between actors.

First, we realised that the way in which researchers define their personal successes was not necessarily standard, and that definitions of successes seem to change with different contexts, demands, and career stages. For instance, the desire to make a change in society was particularly strong in early career researchers, while more established researchers also valued simple curiosity, and relational successes.

When involving all different research actors, we were able to build a representation of success which was nuanced and multifactorial. Success appeared to be an interaction between characteristics from the researcher (Who), research outputs (What), processes (How), and luck. Interviewees noted that current research assessments tended to value outputs but to largely ignore processes, even though these were deemed essential not only for the quality of science, but for the collegiality and the sense of community that unites scientists. Luck was thought to play a crucial role in success and was often used to explain cases where evaluations of success were considered unfair: bad luck explained the lack of reward for excellent researchers, while good luck explained that regular researchers moved ahead without deserving it more than others.

Interviewees generally agreed that current research assessments did not capture the whole picture of success, and there were a number of disagreement on the specific indicators used to attribute success. The relevance of publications, impact factors, science communication, and openness in research assessments raised such disagreements.

Interviewees provided insights on the characteristics they considered essential to any fair and representative assessments. Among those, interviewees suggested that science needs a diversity of indicators that are transparent, robust, and valid, and that allow a balanced and diverse view of success; that assessment of scientists should not blindly depend on metrics but also value human input; and that quality should be valued over quantity.

Finally, when asked what they would change in science, many respondents targeted the way in which research is being assessed and rewarded, reiterating that there is an urgent need for fairer distribution of resources and rewards in science.

DISCUSSION

To advance or even maintain their career, researchers need to be successful. But meanings of success in science are not univocal. Different research actors shared their perspective with us, depicting success as a multi-factorial, context-dependent, and mutable construct which is difficult to define. Unsurprisingly, translating the complex idea of success into concrete assessments is challenging.

Many of our respondents worried that current research assessments conflicted with their personal views on success, and they are not the first to express such concerns. Research assessments have been under the radar for quite some time and the discussion on how to improve them continues to grow. The current chapter adds to this discussion by showing that, even when considering the perspectives of different research actors on the way success is defined in science, research assessments generate a lot of criticism. One recurrent criticism was the fact that current assessments over-rely on research *outputs*, thereby ignoring, if not discouraging, important *processes* that contribute to the quality of research. This issue is central to current discussions on the topic. For instance, just a few months ago, Jeremy Farrar, director of Wellcome UK, stated that the "relentless drive for research excellence has created a culture in modern science that cares exclusively about what is achieved and not about how it is achieved" (Farrar, 2019). Resonating this perspective, The Hong Kong Principles for Assessing Researchers state that researchers should be assessed on the process of science, including on responsible practices (principle 1), transparency (principle 2), and openness (principle 3), and that a diversity of research activities, such as scholarship and outreach should be taken into account (principles 4 and 5).

Part of the criticism when assessing outputs also comes from the dominance of inflexible and reductionistic metrics, an issue that was also significant in our findings. By definition, metrics transform complex concepts in simple numerical estimates, thereby inevitably making decisions on what matters and what should be ignored. Impact factors, for example, ignore not only research processes and research contexts, but also slow citation (i.e., citations two years after publication), thereby potentially ignoring innovative research (Schmidt, 2020). Direct citations raise similar issues, considering only recognition and visibility among scientists, but ignoring the impact that the research has beyond the academy (Lebel & McLean, 2018). These limited metrics shape perspectives and research practices, giving a particular agenda to what determines success in science. Echoing such concerns, the Declaration on research assessments (DORA; American Society for Cell Biology, 2013) directly advocates against using the impact factor for individual evaluations, while the Leiden Manifesto and the Metrics Tide reports pledge for the development and adoption of better, fairer, and more responsible metrics (Hicks et al., 2015; Wilsdon et al., 2015). In this regard, the issues raised by our interviewees are at the heart of current discussions on research assessments. But even though the overarching criticisms towards research assessments appear to be aligned, our findings also reveal that perspectives on specific assessments remain multi-sided, and that priorities and desired changes are far from univocal. In connecting these different perspectives, we realised that a key question remains unanswered in the debate on research assessments, namely, 'What do we want from research assessments?'

Considering both the views on success and the perspective of the problems that were raised in our project (Chapter 4), we identified three main objectives of research assessments.

First, one of the objective of research assessments appears to *promote and value good researchers*. Our respondents suggest that success in science has an important individual facet. Assessments were often described as a way to reflect personal merits and to provide recognition for skills, competencies, and efforts. Research assessments are thus expected to aim, at least in part, for the fair recognition of researchers' accomplishments. And indeed, fairness was central to our discussions on success. Interviewees expressed their concern for fairness by blaming luck (and bad luck) for inexplicable successes (or lack thereof) and by worrying that connections, seniority, and renown could yield unfair advantages which are not related to genuine merit (Chapter 4). Valuing researchers also means building capacities and nurturing autonomy in order to create strong and sustainable research units. Accordingly, if the goal for research assessments is to promote excellent researchers, they should also facilitate, support, and sustain strong research teams. This perspective reinforces the importance of rewarding not only personal merit, but also teamwork, diversity, inclusion, and collegiality. Yet, our respondents identified important problems in current research climates which may inhibit these essential features by fostering competition, mutual blame, and mistrust (Chapter 4). Many of these problems have been echoed in past research, such as the precariousness of research careers (European Science Foundation (ESF) & Science Connect, 2017), the vulnerability of researchers' well-being (e.g., Evans, Bira, Gastelum, Weiss, & Vanderford, 2018; Levecque, Anseel, De Beuckelaer, Van der Heyden, & Gisle, 2017), and the perceived lack of institutional support for researchers (Heffernan & Heffernan, 2019). Beyond disrupting the collegiality between researchers, these issues also appear to influence the perceptions that researchers hold of their institutions (Chapter 4). Knowing that researchers' perceptions of research climates can directly influence research practices (Baker, 2015; Martinson, Crain, De Vries, & Anderson, 2010), it seems urgent to address issues embedded in research climates before assessments can truly value good researchers and fulfil the advancement of strong, sustainable, and flourishing research teams.

Other approaches rather focus on the *benefits that science can bring to society*. A common argument for the need to benefit society is the fact that science is primarily financed through public money and should thus profit back (tangibly and intellectually) to society. Following this perspective, research assessments should aim to ensure that scientists involve, communicate, and implement their findings within society. Applicability of research findings, public engagement, science communication, open access, and feasibility would be at the heart of this objective. But in practice, research assessments often neglect public dimensions of research (Alperin et al., 2019). Our interviewees offered polarised views on this topic, with some valuing and others downgrading science communication. We have also explained that, at least within our modest sample, the values of open science and the desire for implementation seemed to diminish as career advanced. While this finding may be anecdotal, it could also suggest that the broad neglect for societal benefit in current assessments shapes researchers' perspectives of success, encouraging them to prioritize competition and metrics over openness and societal value. Consequently, if research assessments aim to promote and value societal benefit, they might need to reconsider the impact that assessments have had on research cultures.

Finally, we should not overlook research's primary and inherent *goal of advancing science and knowledge*. Knowledge is often described as the common objective and the end in itself of science. Two aspects are then essential to consider here. First, to advance science, we need to ensure that research is conducted with integrity. Assessments should thus encourage the processes which maximise the integrity and the quality of research. Yet, openness, reproducibility, rigorousness, and transparency were recurrently mentioned as missing from current research assessments. Certain aspects of current assessments were even thought to discourage integrity and research quality. Many interviewees supported that the lack of consideration for negative results caused tremendous research waste, that competitiveness of assessments compromised collaborative efforts and transparency, and that the current focus on 'extraordinary findings' discouraged openness and transparency (Chapter 4). Evidently, if the reason for assessing research is to promote the advancement of science, processes which foster integrity must be given due recognition. But even when integrity and quality of research are ensured, advancing science requires continued innovation, creativity, and productivity. According to our findings, this is where most research assessment currently focus. Publications, impact metrics, and past successes all ensure that 'new' knowledge is created. Yet, the overemphasis on quantity and outputs and the negligence of quality and processes was highly criticized by our interviewees. Current assessments, for instance, were said to shift researchers' focus from 'what is needed to advance the field' to 'what is sexy to publish', or 'what will attract funding'. Current assessment systems were further criticized for their conservatism and for the difficulty to pass disrupting and truly innovative ideas through peer-review (see Chapter 4). Short-term funding schemes and the high pace of research evaluations were also criticized, with interviewees noting that innovative research requires long-term investments and sufficient freedom for failure. In sum, both the overlook of research processes and the expectation of quick, positive research results suggest that current research assessments are not optimized to advance knowledge.

Our findings do not provide an answer to what research assessments *should* aim for, but rather illustrate that resolving this first — and often overlooked — question is already challenging. The variety of answers we collected suggests that the perceived objectives of research assessments may differ from person to person, and that specific actor roles may come into play. In theory, it seems reasonable that universities aim to create sustainable and empowered research teams, that funders and policy makers aim to maximize the societal value of science, and that publishers, editors and researchers aim to contribute solid advances to the existing pool of knowledge. But such simple perspectives do not reflected reality, where research actors are themselves individuals with personal perspectives, experiences, and convictions. The complex association of perspective indubitably provides richness to the scientific system, and it would be absurd, if not damaging, to aim for a single unified perspective. Yet, the 'end' goal of research assessments is often obscured from discussions on the topic. Most discussions aim to find the 'means' (i.e., metrics, indicators) to fit the 'end', but fail to define the end of research assessments. Ensuring that the discussion on research assessments listens to the perspectives of all research actors — including the forgotten voices such as early career researchers — and that all parties are transparent and explicit about what they wish to achieve by assessing researchers may be a first step for an open dialogue to enable concrete changes to take place.

CONCLUSIONS

The present chapter describes the perspectives of different research actors on what defines and determines success in research. In their answers, interviewees raised a number of shortcomings about the approaches currently used for assessing success in science, and these shortcomings lead to important problems in the functioning of science (see Chapter 4). Most notably, participants noted that current research assessments place too much emphasis on research outputs and on quantity, while they largely overlook research processes and indicators of quality.

Issues with research assessments have been on the priority agenda for some years already. But although reflections and ideas for change are on the rise, concrete changes are still moderate and sporadic. In this chapter, we bring the debate one step back to ask 'What do we really want from research assessments?'. Are assessments meant to value and encourage good researchers, to benefit society, or to advance science? We argue that current research assessments fall short on each of these core objectives and need to be addressed.

Assessing researchers is an issue that has high stakes, not only for individual researchers who wish to continue their career and seek recognition, but also for the future of science. Our findings reiterate that current research assessments need to be revisited, that all research actors must be involved in the discussion, and that the dialogue must be open, inclusive, transparent, and explicit. Acceptability, trust, and joint efforts can only be increased if all actors are involved, understand the other's perspective, and work together to build a solution.

REFERENCES

Alperin, J. P., Muñoz Nieves, C., Schimanski, L. A., Fischman, G. E., Niles, M. T., & McKiernan, E. C. (2019). How significant are the public dimensions of faculty work in review, promotion and tenure documents? *eLife, 8*, e42254. doi:10.7554/eLife.42254

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics, 13*(4), 437-461. doi:10.1007/s11948-007-9042-5

Aubert Bonn, N., & Pinxten, W. (2016). Rethinking success, integrity, and culture in science (Re-SInC). *Open Science Framework*. doi:10.17605/osf.io/ap4kn

Baker, M. (2015). Metrics for Ethics. *Nature, 520*, 713. doi:10.1038/nj7549-713a

Dierckx de Casterlé, B., Gastmans, C., Bryon, E., & Denier, Y. (2012). QUAGOL: A guide for qualitative data analysis. *International Journal of Nursing Studies, 49*(3), 360-371. doi:https://doi.org/10.1016/j.ijnurstu.2011.09.012

Elo, S., & Kyngas, H. (2008). The qualitative content analysis process. *J Adv Nurs, 62*(1), 107-115. doi:10.1111/j.1365-2648.2007.04569.x

European Science Foundation (ESF), & Science Connect. (2017). *2017 Career Tracking Survey of Doctorate Holders*. Retrieved from http://www.esf.org/fileadmin/user_upload/esf/F-FINAL-Career_Tracking_Survey_2017__Project_Report.pdf

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*, 282. doi:10.1038/nbt.4089

Farrar, J. (2019). Why we need to reimagine how we do research. *Wellcome - News and views*. Retrieved from https://wellcome.ac.uk/news/why-we-need-reimagine-how-we-do-research

Heffernan, T. A., & Heffernan, A. (2019). The academic exodus: the role of institutional support in academics leaving universities and the academy. *Professional Development in Education, 45*(1), 102-113. doi:10.1080/19415257.2018.1474491

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature News, 520*, 429-431. doi:10.1038/520429a

Lebel, J., & McLean, R. (2018). A better measure of research from the global south. *Nature, 559*, 23-26. doi:10.1038/d41586-018-05581-4

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868-879. doi:10.1016/j.respol.2017.02.008

Martinson, B. C., Crain, A. L., De Vries, R., & Anderson, M. S. (2010). The Importance of Organizational Justice in Ensuring Research Integrity. *Journal of Empirical Research on Human Research Ethics, 5*(3), 67-83. doi:10.1525/jer.2010.5.3.67

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2019). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *OSF Preprints*. doi:10.31219/osf.io/m9abx

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology, 16*(3), e2004089. doi:10.1371/journal.pbio.2004089

Moore, S., Neylon, C., Paul Eve, M., Paul O'Donnell, D., & Pattinson, D. (2017). "Excellence R Us": university research and the fetishisation of excellence. *Palgrave Communications, 3*, 16105. doi:10.1057/palcomms.2016.105

Saenen, B., & Borell-Damián, L. (2019). *EUA Briefing - Reflections on University Research Assessment: Key concepts, issues and actors*. European Universities Association. Retrieved from https://eua.eu/resources/publications/825:reflections-on-university-research-assessment-key-concepts,-issues-and-actors.html

Schmidt, R. (2020, 24 August). The Benefits of Statistical Noise. *Behavioral Scientist*. Retrieved from https://behavioralscientist.org/the-benefits-of-statistical-noise/

Wilsdon, J., Liz Allen, Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from https://re.ukri.org/documents/hefce-documents/metric-tide-2015-pdf/

Chapter 4

Rethinking success, integrity, and culture in research: A multi-actor take on problems of science

An adaptation of this chapter has been submitted to the journal *Research Integrity and Peer Review* and is currently available as a preprint on biorXiv at https://doi.org/10.1101/2020.02.12.945899.

CONTRIBUTIONS

Conceptualization: Noémie Aubert Bonn, Wim Pinxten[1], Raymond De Vries[2]

Funding acquisition: Wim Pinxten. Funding granted by the Bijzonder Onderzoeksfonds (BOF) 15NI05

Project administration: Noémie Aubert Bonn, Wim Pinxten

Methodology: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries

Resources: Noémie Aubert Bonn, Wim Pinxten (general); Raymond de Vries, Melissa S. Anderson[3], and Brian C. Martinson[4] (focus group guide); Ines Steffens[5], Inge Thijs[6], and Igna Rutten[6] (focus group organisation and help in recruiting participants)

Investigation: Noémie Aubert Bonn

Data curation: Noémie Aubert Bonn

Formal analysis: Noémie Aubert Bonn

Visualization: Noémie Aubert Bonn

Validation: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries

Supervision: Wim Pinxten

Writing – original draft: Noémie Aubert Bonn

Writing – review & editing: Noémie Aubert Bonn, Wim Pinxten (intermediate and final versions)

1. Department of Healthcare and Ethics, Hasselt University, Hasselt, Belgium
2. Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor (MI), USA
3. Department of Organizational Leadership, Policy, and Development, University of Minnesota, Minneapolis (MN), USA
4. HealthPartners Institute, Minneapolis VA Medical Center, Center for Care Delivery and Outcomes Research, and University of Minnesota, Department of Medicine, Minneapolis (MN), USA.
5. European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden
6. Faculty of Biomedical and Life Sciences, Hasselt University, Hasselt, Belgium

ABSTRACT

Background: Research misconduct and questionable research practices have been the subject of increasing attention in the past few years. But despite the rich body of research available, few empirical works provide the perspectives of non-researcher stakeholders.

Methods: To capture some of the forgotten voices, we conducted semi-structured interviews and focus groups with policy makers, funders, institution leaders, editors or publishers, research integrity office members, research integrity network members, laboratory technicians, researchers, research students, and former-researchers who changed career to inquire on the topics of success, integrity, and responsibilities in science. We used the Flemish biomedical landscape as a baseline to be able to grasp the views of interacting and complementary actors in a system setting.

Results: Given the breadth of our results, we divided our findings in two chapters with the current chapter focusing on the problems that affect the quality and integrity of science. We first discovered that perspectives on misconduct, including the core reasons for condemning misconduct, differed between individuals and actor groups. Beyond misconduct, interviewees also identified numerous problems which affect the integrity of research. Issues related to personalities and attitudes, lack of knowledge of good practices, and research climate were mentioned. Elements that were described as essential for success (Chapter 3) were often thought to accentuate the problems of research climates by disrupting research cultures and research environments. Even though everyone agreed that current research climates need to be addressed, no one felt responsible nor capable of initiating change. Instead, respondents revealed a circle of blame and mistrust between actor groups.

Conclusions: Our findings resonate with recent debates and extrapolate a few action points which might help advance the discussion. First, we must tackle how we assess researchers. Second, approaches to promote better science should be revisited: not only should they directly address the impact of climates on research practices, but they should also redefine their objective to empower and support researchers rather than to capitalize on compliance. Finally, inter-actor dialogues and shared decision making are crucial to building joint objectives for change.

INTRODUCTION

When performing scientific research, researchers agree to abide by principles and standards of practice. We know, however, that best practices are not always upheld (Fanelli, 2009; Martinson, Anderson, & De Vries, 2005; Pupovac & Fanelli, 2014). Obvious deviations from accepted practices are generally known as misconduct. But misconduct is difficult to define. At the moment, one of the most widely accepted definition of misconduct comes from the US Department of Health and Human Services 42 CFR Part 93. This definition is endorsed by the US National Institute of Health and Research Integrity Office, and defines misconduct as "fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results." Nonetheless, the definition also specifies that "research misconduct does NOT include honest error or differences of opinion" (National Institute of Health). In other words, even in its simplest definition, misconduct remains contextual and nuanced, further complicating what constitutes research integrity. Adding to this complexity, several behaviours which cannot be characterised as manifest misconduct are also thought to deviate from research integrity. These behaviours, referred to as questionable — or detrimental — research practices, are so common in the scientific community (Fanelli, 2009; Martinson et al., 2005) that their cumulative damage is believed to surpass the damage of manifest misconduct (Bouter, Tijdink, Axelsen, Martinson, & ter Riet, 2016). Nonetheless, questionable research practices are not univocally condemned, adding to the challenge of distinguishing acceptable from inacceptable practices.

Beyond the complexity of identifying which behaviours transgress research integrity, the causes that may lead to integrity deviations also bring confusion and disagreement. A vast body of research on the topic suggests that both personal and environmental factors are at play. Some studies condemn personal factors such as ego and personality (e.g., Antes et al., 2007; Bailey, 2015; Brown et al., 2011; Davis, Riske-Morris, & Diaz, 2007; Davis, Wester, & King, 2008; Hren et al., 2006; Miller, Shoptaugh, & Wooldridge, 2011; Okonta & Rossouw, 2013), gender (e.g., Fang, Bennett, & Casadevall, 2013; Ghias, Lakho, Asim, Azam, & Saeed, 2014), and career stage (e.g., Fanelli, Costas, & Larivière, 2015;

Martinson et al., 2005). A few others instead believe that researchers' lack of awareness of good practices (e.g., Adeleye & Adebamowo, 2012; Babalola, 2012; Kraemer Diaz, Spears Johnson, & Arcury, 2015), inadequate leadership modelling and mentoring (Buljan, Barać, & Marušić, 2018; Wright, Titus, & Cornelison, 2008), and inefficient oversight (Mumford et al., 2006) are to blame. But some studies also suggest that issues embedded in the research system are at play (Davies, 2019). Among those, the pressure to publish (e.g., Anderson, Ronning, De Vries, & Martinson, 2007; Fanelli, 2010; Singh & Guram, 2014; Tijdink, Verbeke, & Smulders, 2014; Wester, Willse, & Davis, 2010), perverse incentives and conflicting interests (e.g., DuBois et al., 2013; Kaiser et al., 2012; Lundh, Krogsbøll, & Gøtzsche, 2012; Shrader-Frechette, 2011), and competition (Anderson et al., 2007) are the most frequent suspects. In light of these works, integrity seems to depend on a complex interaction between individual and social factors, climates, and awareness.

Despite the rich body of research available to explain what threatens research integrity, few empirical works target the perspectives of the stakeholders beyond researchers (Chapter 1). Given the diversity of actors involved in research systems, focalising the integrity discourse on researchers inevitably overlooks essential voices.

To add some of the forgotten voices to the discourse and understand how non-researchers perceive research climates, we captured the perspectives of policy makers, funders, institution leaders, editors or publishers, research integrity office members, research integrity network members, lab technicians, researchers, research students, and former-researchers who changed career on the topics of success, integrity, and responsibilities in science. We used the Flemish biomedical landscape as a baseline to be able to grasp the views of interacting and complementary actors. Given the breadth of our results, we divided our findings in a two-chapter series, with the current chapter focusing on the problems that affect the integrity of science (see the associated findings on success in Chapter 3).

METHODS

In the current chapter, we retell the perspectives of different research actors on misconduct and on the problems which affect the integrity of science. Our data comes from interviews and focus groups with PhD students (PhD; n=6, focus group), post-doctoral researchers (PostDoc; n=5, focus group), faculty researchers (researchers; n=4, focus group), laboratory technicians (LT; n=5, focus group), researcher who changed career (RCC; n=5), members from research integrity offices (RIO; n=4, interviews), research institution leaders (n=7), policy maker or influencers (PMI; n=4, interviews), members of the network of research on research integrity (RIN; n=3, interviews), research funders (FA; n=5, interviews), and editors or publishers (EP; n=8, focus group and interviews), see Table 1 in Chapter 3 for more details on participants. The project was conducted in Flanders, Belgium, and most participants came from, or were connected with, the Flemish research system. This chapter complements the findings detailed in Chapter 3. The full methods, materials, and participants are detailed in the associated Chapter 3.

RESULTS

The purpose of this chapter is to retell, connect, and extend on the issues that the different actors raised in our study. Aiming to maximise transparency and to minimise selective reporting, we provide numerous quotes and personal stories to illustrate our claims. The result, however, is a lengthy chapter in which we explain the breadth of the concerns raised by our participants. Given its length, a short summary of results is available at the end of the results section, and select findings are re-examined and extended in the discussion.

MISCONDUCT

Why misconduct matters

As we explain in the introduction, defining misconduct is challenging and very likely dependent on the context and the research culture in place. Probing directly for these complex definitions risked generating rote answers from our interviewees. Consequently, instead of asking our respondents to define misconduct, we asked them about the 'red flags' that indicate when researchers may be involved in inacceptable practices. By explaining these red flags, interviewees went beyond a finite list of research behaviours that they believed lacked integrity, and hinted at the reasons and personal perceptions of integrity in science.

Many interviewees started by explaining that misconduct was very difficult to detect. Some explained that the continuum of questionable research practices blurred the distinction between what may be considered misconduct, what may be punishable, and what may be acceptable despite deviating from best practices. Others explained that misconduct had a "shifting" definition which challenged accusations of past misbehaviours. But most interviewees mentioned that the biggest challenge in detecting misconduct was the difficulty to prove intention. Interviewees who had to deal with cases of misconduct mentioned being able to 'feel' when a case was intentional, but often missed the elements to prove it. The following quote illustrate this thought.

> *"But I know that there is a problem with integrity in that person. I can feel it. We have no proof." (RIL)*
>
> *"We had this case once of a guy… and I, up until now, I'm still convinced that he completely fabricated his research. I know for sure that he did. But we weren't able to prove it because it's very difficult to prove that something is not there.* […goes on to describe the case in which the researcher deleted all possible evidence] *So there was no proof anywhere. And it was the adding up of all these coincidental things that made us believe – and in fact of course his attitude and the entire person, him as a person being was very unreliable with a lot of lies, with a lot of contradictions, stories that didn't add up, very negatively threatening… so it was a very nasty one. And at that point you sense that there is something off." (RIO)*

Asking about red flags also allowed us to grasp what made specific practices inacceptable. We found that the reason for condemning practices ranged from general worries about the impact on science to worries about the morality and motives of the researchers. We illustrated these main positions in Figure 1, and illustrated each position with quotes in Appendix 8.

The answers were mixed and diverse, but some group-specific characteristics could be observed. Among those who worried most about the impact on science, some interviewees emphasized that the potential to alter conclusions or change the course of science was what made misconduct troublesome. Editors and publishers were particularly strong on this view. Although they acknowledged the importance of intention to condemn misconduct, editors and publishers emphasized that, given their late entry in the research process, their main concern was on the effect that misconduct may have on the scientific record. This view was not only the view of editors. Some institution leaders also highlighted that not all bad intentions shape equal forms of misconduct. For example, while intentions to save efforts and be lazy can clearly harm the quality of results, they might be of a different order than the intention to change conclusions for your personal benefit.

> *"The only misconduct I've picked up was just stupidity. PhD students who scanned a little too short and had to go back to the scanner and thought "I could just copy-paste the bottom bit because there's nothing on it anyway". That's real misconduct, but at the same time, that's not scientific fraud. Well it was, it is scientific fraud, but he was not changing a conclusion, he was just too lazy to scan a really nice experiment [...] What I consider cheating is that you leave out the data that don't suit your model. Or you make up data to get your model correctly. That is what I call cheating." (RIL)*

Although both cases are unquestionably intentional, in the first case science is harmed as a side effect of pursuing a goal extrinsic to science (i.e. laziness), while in the second case science is harmed by explicitly by going against its intrinsic goals (i.e. producing inaccurate results).

The same interviewee later mentioned that small deviations, including misconduct in early stages of research, might not be so problematic since they were likely to be corrected early on and had little risk of changing the course of science (See quotes in Appendix 8). Nevertheless, research institution leaders varied greatly in their answers, and some rather supported that the intentions and truthfulness of researchers were most important since they affected them as employers. Somewhere between these views, other interviewees argued that misuse of research money immediately constituted misconduct. In this regard, one policy maker or influencer believed that any misrepresentation or duplication in an application done purposively '*to win money'* should be considered fraud. Another added that producing weak or low quality results which could not be generalized or used for further research was also "*sloppy or bad practice*" since the results will not represent reality. One research funder supported this perspective by adding that poor quality and delays in delivery were crucial to them since their goal was to "*guarantee the most efficient use of public money*". Finally, other interviewees focused more on the individual than on the impact on research and resources. Research integrity office members in particular tended to talk about intention or morals as the aspect that was most important to flag and determine misconduct. One interviewee explicitly mentioned that even if the conclusions were unchanged and the results were simply slightly embellished, the intention and moral mismatch was what made these practices inacceptable.

Integrity jargon

Although research integrity office members, research integrity network members, and editors or publishers used the key terms from the integrity literature (e.g., falsification, fabrication, plagiarism (FFP), misconduct, and questionable research practices (QRP)), many other interviewees, including funders, policy makers or influencers, researchers, and some institutions leaders, appeared less familiar with this jargon. They would use descriptions such as 'changing your data', 'faking data', 'cheating', rather than the more familiar FFP and QRP terms. Even the term 'misconduct' was rarely used, most often replaced by 'fraud'. This unfamiliarity with integrity jargon may be due to Dutch-speaking nuances, or to our sampling strategy (i.e., we intentionally include interviewees who were not integrity experts in order to obtain a perspective that was unbiased by the integrity literature). Nonetheless, this finding also means that researchers working on research integrity should be aware that common terms such as FFP, misconduct, QRP, and other key terms may still be jargon to actors who are, ultimately, the intended audience.

Figure 1. What makes questionable practices unacceptable?

What causes misconduct?

We asked our respondents *why* they think misconduct and questionable research practices happen, and whether they think it *can happen to anyone*. The main themes mentioned are illustrated in Figure 2, and illustrative quotes for each theme may be seen in Appendix 9. Since some of these themes were also mentioned as general problems of academia, we will repeat and expand on the themes later on in the 'Problems beyond misconduct' section.

Figure 2. Main themes mentioned as causes for misconduct

Theme	Number of interviewees mentioning the theme
Pressure	>20
Ego and personal morals Normalisation of smaller misbehaviours	≥10
Perverse incentives Lack of awareness Lack of controls	≥5
Unrealistic demands Lack of openness to failure and negative findings Overspecialisation Cultural background	<5

Pressure was among the most mentioned potential causes for misconduct and questionable research practices. Despite the frequent reference to pressure as being excessive and problematic, at least ten interviewees (including LT, PMI, FA, RIL, and RIO) supported that pressure ultimately does not discharge researchers from their personal responsibilities to act with integrity. Key arguments for this position included the fact that research is not the only profession in which pressures are high, the view that pressures cannot justify moral deviance, and the perspective that even though pressures are high, pursuing research careers is a choice of which researchers are ultimately responsible. Select quotes expressing these ideas are available in Table 2. **Egos and morals**, or the 'bad apple' idea, was also recurrently mentioned as a possible explanation for misconduct. The high prevalence of interviewees who mentioned egos and morals might have been primed by our question 'can misconduct happen

Table 2. Select quotes supporting that pressure may not be an excuse for misconduct

Research is not the only career with high pressure
"I would say that pressure is definitely a part, but I find it a bit of a cheap explanation if I may say so. Because you have pressure in a lot of professions eh, it's not just in science. OK there is some insecurity about the future and projects and things like that, but also that is not all that uncommon." (RIO)
"Because again there's discussions in newspapers, to say well you should understand people with that high pressure of publications and... I don't... I don't understand this, because if you have a truck driver tomorrow he also has pressure to drive 150 kilometres on the highway, not to care about crossing pedestrians, he wouldn't do that, or we wouldn't accept it. So we will not accept it from a professor in a lab neither, I don't think so, I think everybody would do it." (FA)
Pressures do not justify moral deviance
"An answer that is often given is exactly what we already discussed: Publication pressure. I'm not sure. I <have not> seen the proof of it. I can imagine that for a whole set of reasons, publication pressure that is too big, that is too strong, can have negative results. And that is not a rational thing to do to put so much pressure on people just to produce for the publication's sake. The sake of publishing in itself. Is that also the reason why people start to fabricate or manipulate research? I'm not sure. I'm really not sure." (FA)
"I can imagine some people being forced to deliver something they're not 100% happy about, but I find it a bit harsh to say that just because you have to publish three papers and you're under a lot of stress it's an excuse to make <up> some publications." (RIO)
Researchers ultimately choose to do research
"I mean, what we always forget when we discuss this is: people choose to be in this context. Nobody puts a gun to your head and says 'you have to work in this lab and produce a Nature paper'. It's a choice. But of course you can say 'yeah, family pressure, peer pressure, and you get into this dynamic and we know group dynamics are very heavy, and they have caused, there are people who think 'I have no choice I have to do this otherwise I lose my work... Then you ask 'and what happens if you lose your job? Do you have to kill yourself? No of course not. Even if you have to drive a taxi or so. But it's a choice and this is not clear. And in the entire discussion, this I find completely left out, that people are not forced with a gun to their head to work in a high-performance laboratory. It's their own ambition." (RIL)

to anyone?' (as we will see later), yet many interviewees spontaneously mentioned the influence of personalities and morals on misconduct. Respondents especially linked egos to 'big misconduct cases' such as the cases that appear in the news. **Growing tolerance of misbehaviours**, either by seeing colleagues perform bad science or by getting away with small suboptimal practices, was also often discussed as a catalyst for detrimental practices and misconduct.

Beyond these three major culprits, additional determinants of misconduct were raised. Among those, the perverse effects that unadapted incentives may incur, added to a lack of control for compliance, were thought to make misconduct a low-risk high-gain prospect. The overspecialization of research fields was also thought to challenge monitoring and reproducibility. Finally, research cultures were also thought to threaten integrity, for instance through the cultural background of researchers or students, the lack of openness for failure and negative findings, and the lack of realism in expectations and demands. We will revisit these three themes later on since they were also mentioned as more general problems of science.

Is everyone subject to misconduct?

When asking respondents whether misconduct could happen to anyone, the answers were varied and contradicting. Although respondents identified pressures and problems from the research culture and environment as major causes for misconduct, most respondents also supported that certain types of personalities were more prone to misconduct than others. The extent of this perception, however, varied from interviewee to interviewee, and appeared to be linked to personal experiences with misconduct cases rather than to actor groups.

First, some interviewees perceived researchers as inherently good by default. Statements along the lines of "*I believe in the goodness of researchers*" (RIO), "*She was a real scientist… I could not believe that she would ever, yeah, <commit> misconduct on purpose.*" (RCC), or "*I find it very hard to believe that somebody who would go into science, go into research to intend really to go and do wrong things.*" (RIO) illustrate this perspective. Nonetheless, the same interviewees later explained that despite researchers' inherent goodness, academia sometimes placed so much pressure on researchers that it may push them to deviate from integrity. Corroborating the view that research culture and environments may drive virtually anyone to commit misconduct, other interviewees were more explicit in linking propensity for misconduct to individuals. These interviewees admitted that pressures played a key role in misconduct, but believed that certain researchers were more prone to misconduct than others.

> *"I definitely think there is a pathological end of the spectrum. […] But I also think that there is so much pressure, especially on people at the beginning of their careers, that I don't think anyone is completely immune to actually committing something" (EP)*

> *"The truly white and the truly blacks are rare. […] many people will be willing to cut a small corner somewhere in an experiment. But really cutting a corner meaning 'I come up with an answer that I don't have yet, but I assume it will be this and I'll give myself the data for free'... I think requires a mentality". (RIL)*

Finally, a minority of interviewees believed that individual characteristics were the biggest (or sole) determinant of integrity. Although this perspective was only supported by a few interviewees, it suggests that integrity is sometimes perceived as independent from training and climates. Supporters of this view questioned the benefit of training and support in promoting research integrity, rather asserting that, to build good researchers, institutions must choose the right individuals.

> *"Sloppy science, first and foremost is the product of sloppy scientists. It's not the product of a system, it's the product of a person. […] there are persons who are striving for high levels of integrity, and there are people who are not doing so" (PMI).*

> *"Integrity is in the person. […] integrity is something that is in you. You have it or you don't have it. I mean you have it, it's there. And when you don't have it, you don't have it. So we cannot create integrity, it's something that's in the people. Working together and being involved, that's something [universities] can create by offering a structure. But I'm a strong believer that the integrity is inborn, it's in you". (RIL)*

Our interviews revealed that perspectives, knowledge, and convictions about misconduct can vary greatly between individuals and actor groups. Not only are the terms used to talk about misconduct still jargon to many research actors, but the views on why misconduct matters, what causes it, and who is susceptible to it also varied greatly between interviewees.

THE PROBLEMS BEYOND MISCONDUCT

In performing our interviews, we noticed that what respondents were most concerned with was not 'misconduct' per-se, but rather a number of more general problems that affect the integrity of research.

> *"The experience that I have in research is that really [misconduct] is exceptional. It makes… It's breaking news, because it's something that we, in the community of research, we consider inacceptable, but it's rare. It's rare." (FA)*

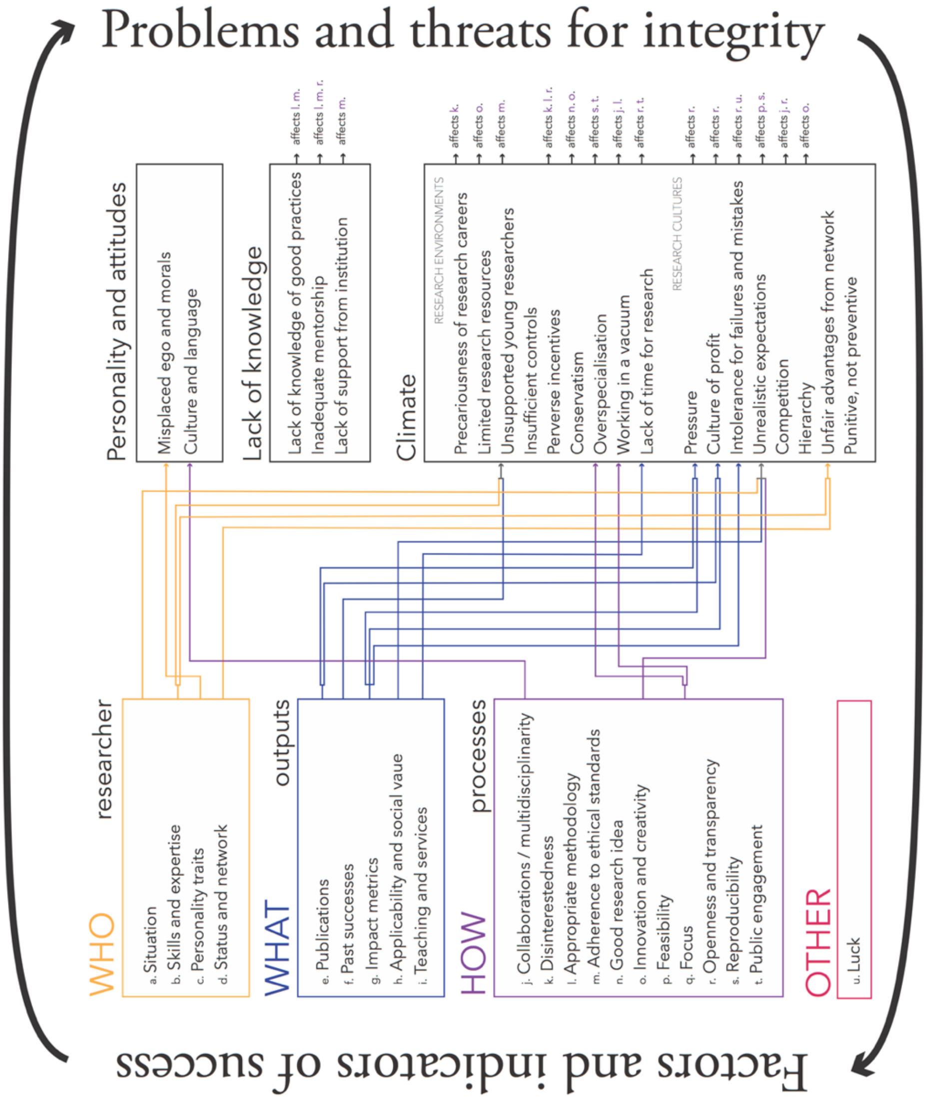
> *"The very serious misconduct is not such a big problem. It's… it's more the grey area that is a problem because of, yeah, the amount of… of, yeah, the bad practices." (RIO)*

Indeed, respondents discussed what they found problematic in research and what frustrated them much more spontaneously than genuine misconduct. They expressed these problems and frustrations throughout the interview, even during discussions on success and responsibilities, often forcing the interviewer to re-focus the discussions on the 'positive side' of science. Results from the following sections are thus based on spontaneous reflections expressed throughout the interview rather than limited to specific interview questions.

A tight connection between success indicators and threats to integrity

The first thing we noticed when analyzing the problems and frustrations raised by our respondents is that many of them are intimately connected to the way success is attributed in science. In Figure 3, we linked the different themes of success that we reported in Chapter 3 to the different problems or frustrations mentioned by interviewees. Despite the oversimplification of this illustration, we can see at first glance that the two topics are highly interconnected. In fact, not only are determinants of success seen as aggravating some of the problems mentioned by interviewees, but some of the problems mentioned are also seen as blocking or damaging success in science. In addition, problems appeared to influence one another, escalating into bigger issues until some of them became big enough to generate misconduct. In this regard, some of the problems described as causes for misconduct were first described as general problems of science.

Figure 3. Simple depiction of the complex interaction between the factors and indicators that currently define success and the problems which currently threaten research integrity.



Note: Although this figure oversimplifies the complex interaction between success and integrity, it shows how diverse and circular the connection is, with both success generating problems, and problems influencing and blocking the processes needed for success.

Three categories of threats to integrity in science

Interviewees raised multiple problems which could potentially threaten the integrity of science. We organized the different problems mentioned in three big categories: problems related the *personality and attitudes*, problems related to *a lack of knowledge from researchers*, and problems related to *research climate*s, which include research environments and research cultures.

ISSUES RELATED TO THE PERSONALITY AND ATTITUDES

Interviewees mentioned several individual characteristics that could be problematic and might impede on the integrity of research. We have already mentioned a few of these issues — such as misplaced ego and morals — above when discussing individual propensity to misconduct. But a few points raised conflicting dualities with what was believed essential to success. For instance, interviewees supported that ambition, passion, and tenacity were key elements of success (Chapter 3). Nonetheless, they also supported that hyper-ambition or excessive desire to be successful could bias conclusions and encourage researchers to loosen their integrity.

> *"That's very important, because we're always talking about misconduct as if it's deliberate, as if you're cheating, I think maybe the most dangerous thing in research is in your wishful thinking, of self-fulling prophecies, you want it so badly that you will see it, you will see it the results, if you're out in a complex…" (FA)*

> *"And the good researchers, like I said, they're really passionate, they only think about their own research, they want to get things done, they want to get their results, so… What we usually see is that people then don't really follow the rules as they should, so they don't see why these rules are important. […] It's always that this researcher is really heart to… motivated to get the results done, and then bypasses procedures and rules, and that they don't see why these rules are in place and why they're so important to have them, why it's also protecting other people in the field… So this is mainly I think one of the reasons…" (RIL)*

Several respondents also associated problematic attitudes with cultural backgrounds, proposing, for instance, that the perceived seriousness of detrimental research practices may differ between cultures.

> *"…Sometimes I find that it's a matter of some cultural differences, in some cultures it seems that every meaning is justifiable to achieve the goal, and so they are trying to do anything they can do just to get*

> *their research published. So they will falsify data without a problem [laughs], they will not hesitate…" (EP)*

> *"I think [certain cultures] have this mentality that it's almost, you honor somebody by plagiarising them. And they just want to get their diploma so they can do a PostDoc in America." (RIL)*

> *"I dare to say that [different cultures] have a slightly different opinion about rules." (RIL)*

Cultural and language differences were also mentioned as challenging the ability to communicate and increasing the risk of loneliness, misunderstandings, and mistrust. For instance, laboratory technicians mentioned that cultural and language differences could decrease students' willingness to ask questions and disclose mistakes, thereby increasing the cumulative severity of mistakes and the temptation to conceal them. But despite raising these personal issues as a potential risk for integrity, interviewees failed to propose concrete solutions or improvements to minimise the risks resulting from personalities and attitudes.

ISSUES RELATED TO A LACK OF KNOWLEDGE

Lack of knowledge of good practices

Several respondents mentioned that researchers were sometimes unaware of good practices. Lack of knowledge of good practices was not only perceived as a problem of individual researchers who lack insight in their own behaviour, but also as a systemic issue caused by insufficient training and inadequate mentoring within the larger scientific community.

Insufficient support, mentorship, and guidance

Most concerns related to the lack of knowledge of good practices pointed to a lack of mentorship and guidance for early career researchers. This issue was discussed on different levels. One the one hand, students mentioned that they lacked guidance, support, and time from their supervisors. PhD students and researchers who changed career were especially vocal on this point.

> *"Well it was generally just my supervisor messing up. That was just the worst. Always the worst. Always. (laughs) And I'm not telling you… you know. So, not responding to emails, you know, for a very long time. Not being present. Not giving any useful feedback, if they give*

> *feedback, giving feedback that just makes your work worse instead of better… Not knowing how to supervise basically." (RCC)*

> *"I think everything I learnt, I learnt because of doing myself. I expected when I started my PhD project, that I would learn a lot from my supervisors, but now at the end of my PhD I think I didn't really learn a lot from them, so I'm a bit disappointed about that.". (PhD)*

Although the lack of mentoring in such cases is not necessarily causing an unfamiliarity with good practices, young researchers often felt lonely, stressed, and frustrated about the lack of support they receive. Loneliness, in turn, was described as a possible red flag for integrity failures.

> *"I get suspicious when PhD students are complaining, for instance, or feel alone or feel pressured to do things. Of course, in a certain sense this always happens. If you ask my PhD students, they will also say there are moments in which they were alone or pressured, so you cannot really prevent all of that but if that becomes too big, then I think there is something wrong" (RIN)*

On the other hand, researchers themselves mentioned that they lacked support and guidance on how they should meet integrity and ethical standards. For example, one researcher mentioned that funders tended to increase the number of 'tick boxes' without increasing training and capacity.

> *"Just having you as a researcher filling all these tick boxes, and not being responsible... [Funders] really should work on that. Also the same goes now for the data protection. They will make an extra box, and we should think that everything is arranged for data protection while no university in Flanders is ready for that by May 28th." (Researcher)*

Along the same line, one participant in the RIO group observed that integrity training generally comes in the form of specialized, intensive courses, when in fact it should be integrally embedded throughout the research training in order to become "*part of the research process*" for every researcher.

Issues related to the research climate: Research environment

The third broad category of problems raised by our interviewees were linked to the research climate, targeting both the working environment and the research culture.

Precariousness of research careers and limited research resources

The precariousness of research careers and the constant insecurity linked to short-term contracts and scarce opportunities for advancement was a recurrent issue mentioned by our interviewees. Policy makers and research institution leaders were particularly concerned about this issue. One policy maker explained that, in Flanders, the number of students completing a PhD highly exceeds the number of academic positions available and that, despite this imbalance, the current number of PhD students in Flanders is still below the target set by the Organisation for Economic Co-operation and Development (OECD). This interviewee further explained that most PhD students will have to find a career outside academia, often finding careers at a Masters level rather than at a PhD level. Another policy maker or influencer added that the ambition to continue in academia was the default option for PhD students. He further explained that this is problematic since there are "*phantom pains attached to it. People think that it's kind of a lost battle — a defeat — when they leave university and go to work in a company, or go to work for another agency.*" Trying to find solutions to the problem, some interviewees supported that the lack of opportunities in research may result from limited research resources, and that investing more money in research would help solve the problem. Nonetheless, other respondents concluded that because of the way the system is organized, capital investment in research would not necessarily solve this problem.

> *"Because research is a human activity, more money into research means also more people getting involved in research. I'm always amazed to see that people think that this should lead to more academic careers in research. That's a kind of incongruence because of the fact that the public sector cannot bloat universities, in the sense that we cannot multiply by an order of magnitude the amount of positions available." (PMI)*

Unsupported young researchers

Adding to the lack of stability and security embedded in research careers, the struggles of early careers was another important theme in our interviews. Young researchers and former-researchers mentioned that they felt unsupported while juggling with too many tasks to be able to focus on the outputs required for advancing their career. Early career researchers, former-researchers, and post-doctoral researchers also believed that their modest output records disadvantaged them in the fierce competition needed to secure grants and careers in academia. As a result, on top of the duties of early adult life (building a family, buying a house, caring for aging parents, etc.), young researchers struggled with an insecure future, excessive pressures for output, insufficient resources, and the inability to compete with established researchers.

> *"There's a certain starters package that I got, but it's not enough, you have to find your own money which is very difficult because you don't have the publication list. […] Because the first thing [funders] do is, they look at who is asking, and then at your resumé and then they say 'oh no too junior' or 'not well enough established in the filed' or, you know, stuff like that. […] you need money to publish, you need to publish to get money, you know, it's a circle." (RCC)*

> *"Yeah so for me it's because I'm in this end stage, the insecurity of the future is really something that I'm struggling with. Not every day, or not all day every day, but every day at least 5 minutes (laughs). […] The fact that you don't have a permanent position is also really ambiguous about it. I would like to have a little bit more future, and also not to have to find my own money all the time because I have the feeling that I'm not actually doing something myself. I'm constantly finding and looking for more money, so to hire people who are actually doing something." (PostDoc)*

Most former-researchers said that the desire for a stable career with a sane work–life balance influenced their decision to leave academia.

> *"Why should I stay in the academic world, why should I go? […] If I go for the academic world, I'm going to have to tell my wife, that was pregnant [at the time…], I have to tell my wife "well we're going to a financially uncertain situation for at least 10 to 15 years. And maybe when I'm 30 or 35 and I have said no to you an enormous amount of times, I'm going to be so successful that I can say 'It's ok now, we can pay the bills.' But I'm still going to say no to you because I have to compete with the other people. Whereas if I choose another life or career, you get, for example a contract that lasts for your entire life, and you can build your life. You can start building your life. You can settle in a way, you can… You can make plans. Whereas in the academic world you can only make plans for 2 or 3 years. And that was the kind of life that I didn't want to <live>." (RCC)*

Two of the interviewed former-researchers admitted having been mentally and emotionally affected with symptoms of burnout, and all recalled a certain distress from their time in academia (we will get back to the emotional distress when discussing unrealistic expectations below). One interviewee proposed that this distress may be accentuated by "*the enormous discrepancy we have today between the job security of professors and the job non-security […] of PhD students <and> PostDocs*". This interviewee recalled stories of supervisors who continuously reminded their students that they could be replaced anytime if they didn't meet expectations. Such security discrepancy was also thought to create an environment in which young researchers may not feel safe enough to be open or transparent about issues and mistakes — a problem we will target later when discussing the research culture. Finally, there appeared to be strong emotional implications for researchers who decided to quit academia. Even though all former-researchers interviewed expressed a great sense of relief from leaving academia, most admitted that the decision to leave had been difficult to make since it would be perceived as a failure in their career. The emotional involvement was often linked to a sense of personal disappointment or shame, rather than to a frustration against an unrealistic and unsupportive system.

> *"I am the idiot that gave up [a professorship]. That's what it is, I worked my entire career to get at that point, I was in it for [a few] years and I gave it up. And so many people in the academy want to be in that position, and I gave it up. What kind of an idiot am I?* […later in the discussion…] *In the end I was like […] What's the chance that I'll ever help any patients, because that's basically why we all start doing it, to make a difference. But that's for the happy few, and those happy few have big names behind them and get money. They are not struggling to be at home, to put children to bed or whatever. The daily things that were too hard for me, and now that I don't have to do it anymore I'm a happier person. So maybe I'm not a real, real scientist." (RCC)*

In this last quote, the perception of not being a "*real, real scientist"* clearly expresses that researchers who leave academia risk blaming themselves rather than the system's unreasonable demands, a perspective which further deepens the wound and pain from leaving. In sum, the strength of the aspirations that young researchers hold to continue in academia may increase their vulnerability by imposing escalating expectations upon themselves. Knowing that less than 10% of PhD students will be able to pursue academic careers, the current dynamics clearly generate disappointments, self-doubt, and emotional distress among early career researchers for whom the future is uncertain.

Inefficient controls and perverse incentives

Issues around inefficient controls were also raised by a few stakeholders who feared that misconduct and detrimental practices often go unnoticed or unsanctioned. Research integrity officers complemented this idea by mentioning there are also insufficient incentives for integrity.

In fact, **current incentives** were often thought to discourage integrity. One interviewee in the editor and publisher group mentioned that researchers are incentivized to find "*big bold claims*" and to publish in "*very selective journals*", which led to low quality research practices such as performing research on smaller populations or choosing inappropriate statistical controls and analyses to inflate significance. We will get back on this point later on when discussing issues of unrealistic expectations and the culture of publish-or-perish that results from such expectations. Along the same lines, an institution leader mentioned that "*short term financing situations*" which expect high publication outcomes may be the "*worst perverse incentive you can give a scientist*".

Conservatism

Adding to the above concerns, funding distribution was also criticized for being conservative and for discouraging high risk research (i.e., research with important possible outcomes but with high potential for yielding negative results). Interviewees considered that high risk research was important for scientific advancement and innovation, but they worried the reliance on experts for reviews decreased the chances of obtaining funding for high risk or simply unusual research.

> *"…peer review has this tendency to be a little bit conservative. Because since you have experts in your panels, people who already have proven themselves […] and also mostly are senior people, they can also sometimes, not all of them because you shouldn't generalize, but sometimes they can get a little bit conservative. Because they think that they have found the holy grail." (FA)*

One policy maker proposed that the problem also came from within institutions, supporting that "*more and more, institutions, universities don't want to fund high risk research. So they only want to fund research that gives good results that can be used for society and so on.*" As a coping strategy, both PhD students and researchers admitted having heard of situations where applicants "*get funding for something that's already proven, and they just explain it and they turn it… they describe it in such a way that it's new, and then they get funding*" (PhD student), or where researchers "*write a project where <they>*

already have the data and... so <they’re> asking money for something that <they have> already done.” (Researcher). Researchers considered that such coping strategies were problematic because they limited innovation and prevented new research groups from obtaining funding in topics that were investigated by other, more established groups[7].

> *“You need to show that you have every technology in hand to do this new idea. And this is really a problem for me. I think that many researchers are now playing at the safe side. Because they already have shown that they work in this field, they will continue on this field, and they will not go broader, because probably they will not get funding because it's a new idea and they don't have any evidence at work.” (Researcher)*

In response to these concerns, one research funder stated that there also existed private, smaller funders which “*could, and therefore also probably should <be> somewhat more risk taking than public funders.*”

Overspecialisation, working in a vacuum, and lack of time for research

Interviewees also shed light on a problematic interplay between overspecialisation, isolation, and lack of time for research. As introduced in the potential causes for misconduct, overspecialization was criticized for potentially deterring the replicability of research, thereby undermining the detection of mistakes and misconduct. But overspecialisation was also criticized for increasingly isolating researchers from one another and discouraging collaborations. Interviewees often felt that researchers work in a vacuum rather than within a shared community. Evidencing this idea, PhD students supported that research “*is sometimes a bit lonely*” and that they were often unaware of the research that was happening around them. An interviewee from the RIO conceded that “*very often […] researchers don't even know what is happening within their*

[7] It has been pointed out by a reviewer of the present thesis that applying for funding with a project which is already ongoing may, in turn, be a way to stimulate innovation. Indeed, it is possible that, when a grant is given to a project which is already initiated, the full research funds are unlikely to be used up, and the remaining resources may then be used for high risk and innovative projects without needing a new grant application. This idea was hinted in our focus group, but it was also referred as problematic since it infringed the funding agreement, with researchers stating: "*But then it seems it's fraud, because you got funding and then you use maybe for something else. That's what happens to be creative.*" The need to trick applications to allow for creative research further indicates that true creativity and high risk research is not easily accepted in current research funding schemes.

own buildings" and that this isolation probably lead to unnecessary duplication and waste of research resource. Working alone also means that researchers are expected to have highly versatile abilities to be able to coordinate and respond to the expectations of their position.

> *"The advantage of academics is that you have many tasks, but this is also a disadvantage. Sometimes you have to do everything, you have to be good <in> English and <grammar>, in statistics, in everything, and... which is not always our expertise and also neither our interest." (Researcher)*

Another researcher added that the lack of collaboration also reduced the possibility of blinding experimenters and increased the risks of bias. But beyond research multitasking, the three pillars required to be employed in a university — i.e., 'Teaching', 'Research', and 'Services' — also sparked the debate.

> *"The problem is that today we cannot deliberate between those three pillars. And I believe that if you're excellent in education and you spend 80% of your time in education and you do only 20% in research, and you don't get your criteria for research but you overdo your criteria in education, why not make a balance?". (RIL)*

This lack of flexibility played an important role in the decision to quit research of one of our interviewee. Asked to teach approximately 80% of the time, this interviewee recalled that there was too little time left to fulfill the research requirements expected by the institution. According to this interviewee, the lack of flexibility from the three pillars of research careers (i.e., everyone is expected to perform research, teach, and contribute to services) neglected personal skills and preferences.

> *"I believe there are very good researchers that have to teach and that suck at it, and I believe that there are good teachers that have to do research and suck at it. Or at least are not top notch at it. But no we all have to be equally good at both and we all have to divide our time exactly the same. […] I struggled at doing everything the way I want to do it. I want to do everything in a good way. And when you have to do teaching and research I didn't manage. I didn't manage to do both in a good way." (RCC)*

Finally, with such diverse tasks and pillars of expertise, and with the bureaucratic demands of research work, researchers felt that they lacked time to actually do research.

> *"I think the research part you're so passionate about it and then, you know, you feel, you always have to fight to get your time to do it. And there is many, so many things that always come unexpectedly, or expectedly in between, that disable you from writing that article, or from doing your field work yourself... […] because we have to do also education and we have to do managing tasks, and then we have a curriculum reform, then we have to think also about the new education, and then we have, we are responsible for clinical placement and things go wrong on the clinical placement, and then... I mean because I'm juggling many balls, it always... seems like I always, for one reason or the other, have to be juggling those balls instead of being able to do, to spend more time on my research. And we have a tremendous amount of meetings... The amount or time that I'm sitting in the meeting room is … (sighs)" (Researcher)*

Ultimately, this lack of time played back and aggravated a number of issues we just mentioned, such as inadequate mentorship and the difficulty to build one's status as an early career researcher.

ISSUES RELATED TO THE RESEARCH CLIMATE: RESEARCH CULTURE

Deeper into the habits and customs of researchers, several issues embedded in the culture of research were also seen as problematic. Once again, Figure 3 showcases a few of the interactions between issues embedded in scientific culture and current indicators of success.

Pressure

Pressure to perform — and especially to deliver — and the culture of publish or perish were the issues that were mentioned by the biggest number of interviewees. Interviewees described such pressure as potentially causing misconduct, as threatening the quality of science, and as impeding on researchers' health and happiness. By listening to multiple research actors, we also discovered that pressures are multilevel and that they affect more than researchers alone. In fact, the publish-or-perish culture was said to fuel a cascade of expectations and demands which increase pressures on a broad range of research actors. Starting with students and researchers, we first found that pressures did not only come from the institutions and superiors, but sometimes came from the researchers themselves, in the form of personal aspirations and ambitions. On a second level, students and researchers expressed feeling substantial pressure coming from the supervisor and the institution. But institution leaders also expressed that they felt, as an institute, a pressure to deliver more and faster in order to promote the excellence of their research and their attractiveness to the international research community. One research funder explained that in Flanders, where institutional funding depends largely on research outputs[8], institutions must continuously increase their outputs in order to keep their share of structural funding.

> *"…to even conserve [their] share, make sure that [they] will not get less than the previous year, [universities] have to work always harder, [they] have to produce more publications. Because if [their] competitors — other universities — produces more than [them], then the share of that same amount of money will decrease" (FA)*

One level further, funders expressed that they also felt pressures. The increasing number of applications for funding increased workload and generated internal pressure and struggle to find adequate peer-reviewers. Journals expressed a similar concern, stating that the pressure to publish and the current focus on quantity often led them to receive more manuscripts than they could review, and forced them to use greater scrutiny to ensure the quality of their publications, but also to charge higher article processing charges and subscriptions. This whole circle of pressures then appears to link back to policy makers. Specific to Flanders, policy makers were especially criticized for the performance based funding model that was used to distribute research funding between research institutions. Such distribution keys (in particular the BOF-key[9]) were mentioned several times by interviewees and were said to be *"the reason that publications are so paramount in the assessments"* (PMI). Nevertheless, policy makers and influencers clearly expressed that although they were aware of the criticism generated by the BOF-key (and were currently working on a

[8] In Flanders, an important portion of the federal funding for research institutions is distributed using a performance based calculation (Chapter 2). Specifically, a portion of both the core structural funding and the special research funds (BOF: bijzondere onderzoeksfonds) are distributed based on a number of performance indicators. These may include input such as number of students and staff, or outputs such as awarded degrees, defended doctorates, publications, and citations. The Flemish funding model is highly dependent on output indicators (Peters, 2019). In the distribution for the BOF funding, publication metrics account for over 40% of the final score for distribution (Decreet betreffende de organisatie en financiering van het wetenschaps en innovatiebeleid, 2009)(Zacharewicz, 2016). Performance-based funding models are not unique to Flanders and similar systems exist all around the world (Franzoni, Scellato, & Stephan, 2011; Zacharewicz, 2016).

[9] Although the BOF-key was most often mentioned by interviewees, one of the jury members explained that the core structural funding of universities is also partly based on outputs and performance indicators (Peters, 2019).

revision[10]), it was its inadequate transfer within universities (i.e., using its parameters at the individual level) that was at the source of the problem.

> *"The BOF-key was actually only created to divide the money under the universities. And what we see is that the same parameters are being used within the universities themselves to fund the individual researchers. That was never our intention… So that's the negative effect of this key that we never wanted." (PMI)*

> *"The BOF key is just one thing. It's a distribution rule that has to divide a pot of money among five universities. In one way or another, you will always need some distribution mechanism. The BOF Key — That's also why we never report on individual researchers — but what the BOF key does is just aggregated at the level of a university: count PhD output, count publication output (certain type of publication output I'll come back to that). […] I know from hearing and feedback that I get that certain institutions try to, what I would call extrapolate, or interpolate the BOF Key into individual level research output. I think that's wrong, that's even stupid. But it may happen. But the BOF key it's not there to do this." (Another PMI)*

Consequently, even though the pressures and the culture of publish-or-perish were raised by nearly all interviewees, the root of the problem appears to be transferred from one actor to the next. This circle of blame further seemed to create an unappreciation of individual responsibilities and actionable solutions, leaving most actors feeling frustrated and helpless. We will discuss this problem later on.

Culture of profit

In tight connection with the pressures and the culture of publish or perish, the whole culture of profit that characterizes current academia was also questioned by a few interviewees. More specifically, the emphasis on profit and outcomes was seen as potentially undermining the care and consideration that should be given to researchers themselves.

> *"In research you're not making research results, you're making good researchers. And you have to develop and support the people, and not just the research. And I think that that entire culture of care is missing too much. We see them too much as producers of research results,*

---

[10] Since the interviews have been conducted, a new version of the BOF key has been developed and released (mid 2019). Nonetheless, a large proportion of the resource allocation distributed through the BOF-Key still depends on output and publications metrics.

instead of 'we are making a good researcher that will, hopefully go on a lifetime making good research results'." (RIO)

This forgotten need for care easily links back to the lack of support faced by young researchers, the precariousness of research careers, and the lack of support for meeting integrity requirements, while it also feeds into the dominant intolerance for failure and mistakes.

Intolerance for failures and mistakes

Interviewees from all actor groups spontaneously explained that failure, negative findings, and mistakes were almost invisible in science. Yet interviewees also believed that failures were *very "important*", "*valuable*", and "*interesting*", and that they could act as "*a motor to drive you to success*". Intolerance for failure and mistakes was even thought to be a potential incentive for falsification of results (see Appendix 9). One researcher told of a personal experience when discovering a mistake in a team project. From the story, the different reaction and the overall worry that mistakes can generate in science is obvious.

"I think what is also a problem is the fact that it's still a 'taboo' I would say, just to come open with the fact within research "I made a mistake" in the past. We had something in the past in our group that there was… suddenly there was… everybody thought that a measurement was wrong. Something in the system and all of the data that were captured were therefore wrong, and <these> were already data which were published. And then it should be decided what to do. Should we do a correction to the journal or not? And there was a lot of pressure from the professors, because it came higher and higher in the university, and some people were afraid, and some were like 'Whatever!', and everybody had another opinion, the PhD students just had to follow... But I feel the big difference between some people who were very ethically committed, like we have to correct it and we have to send it to the journal, and others were like 'nobody will see it, it's in the past', and... Yeah, I saw a lot of things which should not have been happening." (Researcher)

Although the fear of mistakes seemed deeply engrained in the research culture, many supported that efforts must be made to normalise mistakes in science. One former-researcher eloquently summarised this idea by stating that "*If one place in our world should be a place where people are free to make mistakes, even though we pay them a lot, and we hope they don't make mistakes, then it's [academia].*" (RCC), further advancing that intolerance for failure was "*not justifiable*" in academia. The under-appreciation for negative results was also mentioned very frequently. Interviewees worried that unpublished negative results wasted research resources and could potentially endangered research participants. Yet, on the research floor, the apprehension for negative results was still palpable. Researchers, research students, and lab technicians described negative results as highly frustrating or as '*unlucky*' (see the discussion on luck in Chapter 3) and admitted that projects with negative results were often abandoned early. The quest for positive results also influenced research designs. Students admitted with unease that many experiments seemed designed to ensure publications rather than scientific relevance. Researchers and lab technicians added that data fishing and selective publication were common practice, even sometimes part of the strategies required for success. But when asked about responsibility, interviewees once again seemed to pass the ball to one another. Researchers claimed that they were pushed to look for positive findings since journals would not accept negative results and funders expected their projects to yield positive findings. Nonetheless, both journals and funders refuted this perspective, supporting that their true concern was the value and the quality of the work, not the outcome. Editors and publishers added that they rarely, if ever, received manuscripts with only negative results. One interviewee even told the story of a new journal dedicated to negative results which had to be shut down because it received "*no submission whatsoever*" (EP). The issue thus appears to be deeply embedded within the research culture, possibly even budding at the micro level within the research teams themselves.

> *"If you're really interested in the success of research environment, it's an environment that says 'you don't have to be successful'. 'You may fail. And it's OK. As long as your research methodology is accurate'. […] Now you see that the rector for example is also saying this, so I think change is coming in a certain way… In a certain way. But I'm not that sure if it's really coming because the culture is defined by your promoter. […] You may get trainings every single day. If your promoter or the head of the lab doesn't agree, then it won't happen." (RCC)*

Unrealistic expectations

Intolerance for failure might be a simplistic expression of a bigger problem: science builds unrealistic expectations. Interviewees mentioned that too much was expected from researchers, potentially leading to frustrations, integrity deviations, or even burn out. Different forms of expectations were perceived as being excessive and unrealistic. First, expectations of high yielding results, and extraordinary findings were considered to be embedded in the core of how science is evaluated.

> *"Researchers are incentivized to really get something that is extraordinary, and ground-breaking. And let's face it, all the research in biomedical research is not ground-breaking and extraordinary. Most of it is not." (EP)*

Second, expectations that researchers should work out of passion without personal benefits also surged from our interviews. For example, an institution leader mentioned that institutions "*need people with commitment who <participate in services> for the honour*" without expecting personal gains or compensation. A policy maker added that he "*would never call 'doing science' a job* […] *being passionate about science is almost like being an artist. You live in poverty because you want to pursue your art.*" In other words, research was not seen as a regular career but rather as one that is built on devotion and personal sacrifice for the greater benefit of science. Many interviewees expressed the expectation that researchers work outside ordinary schedules[11], travel abroad regularly, and eventually even rethink their work–life balance.

> *"I think people have to realize when you do a PhD, it's a stressful thing, you really are going to get the highest degree there is at a university, it doesn't fit between 9 and 5." (RIL)*

> *"I'm somewhat older, but I have the impression that younger people have […] somewhat a different work/private balance than I had. And I think that people sometimes could put more energy in their work." (RIL)*

Unfortunately, such expectations of personal sacrifices were not benign on researchers and research students. Researchers and students explained that it was difficult to conjugate their professional and personal life, and that they sometimes felt the need to sacrifice the latter to ensure their professional survival. As we have briefly discussed above, three researchers who changed career mentioned that the difficulty to keep a sane work–life balance played a significant

[11] Working beyond the 9–5 schedule was even seen as a factor for success (Chapter 3).

role in their decision to move away from academia, with some adding that such excessive demands affected their well-being.

> *"I was stressed out completely […] I went to the doctor, [I] was on antidepressants, [I] was in therapy…" (RCC)*

> *"I got a therapist and I worked through it with her and, you know she said… Like whenever she said 'Maybe you want to start thinking about [your work]?' I would just start to cry, so she was like OK, too early! […] Yeah. It was awful." (Another RCC)*

But even those who suffered the effects of excessive expectations tended to perceive "real scientists" as those who should give more than they could. Worrying about this unrealistic perspective and about the implications that unattractive research careers may cause in the long run, one policy maker or influencer advised that researchers should be given equivalent benefits than other individuals on the job market.

> *"Why would I choose to do, to start a career in an area that positions are limited, promotions are limited, high positions are limited, and it's precarious. I have to bring funding, I have to get contracts... So... it's not only the lack of interest — if it is there — from the younger generations, it is also what is the tomorrow. And this is part also, I think, of a scientific governance and a scientific culture issue. Those... We should not consider that researchers are somewhat a different part of the population or that they are saints, that they will sacrifice their wellbeing and their participation in the pleasures of a good economy just because they love science. I think this is very naive." (PMI)*

While it is beyond our purpose to determine whether researchers should, or should not sacrifice their personal life for their career, our interviews show that this expectation is still alive and that it affects researchers' well-being.

Competition, hierarchy, and advantages due to networking

A few additional problems were linked to the social relationships which characterize the academic culture, such as competition, hierarchy, and advantages due to networking.

The issue of competition raised mixed reactions from our participants. On the one hand, some interviewees mentioned that competition was a necessary element of academia as it drove productivity and excellence while imposing limits on the authority of single researchers. But on the other hand, competition also challenged researchers' openness. For one former-researcher, competition was a determining factor for leaving academia. According to this interviewee, competition in academia increased research individualism and dissolved the value of the scientific community.

> *"Competition in the academic world is so strong, so fierce that in the end I experienced it as a… not a war, but a hostile environment." (RCC)*

Not too far from competition, hierarchy was another problem mentioned by some interviewees. In discussing with research students and technicians, we understood that hierarchies were inherent to academia and that they deeply influenced interactions, openness, and integrity. For example, both technicians and PhD students mentioned that they would find it very difficult to openly criticize the conclusion or dubious behaviour of the principal investigator of a laboratory. Most technicians mentioned that they would not dare to flag mistakes and errors because they felt that principal investigators (i.e., researchers) were "*much smarter*" then them. PhD students said that they would refrain from disagreeing with a supervisor's inadequate practice (i.e., we described a case of gifting authorship to a colleague who was not involved in the project) because they were worried that the supervisor would "*make it hard*" on them later on, and might even not allow them to graduate. Researchers, on the other hand, criticized issues linked with inequities in statute and reputation between researchers, saying that because of hierarchy in career achievements, "*the big will <become> bigger and the small will never have a chance.*"

Another relational component raised as a problem from the scientific culture was the issue of unfair advantages from the research network. Networking is an inherent part of science and was mentioned many times as an essential factor for success (Chapter 3). Many researcher and students provided examples in which networking could help them get ahead, and some mentioned that there was comfort in knowing that good relations could bring '*favours*' in case of need.

> *"Once you are in the network, you can also rely upon them and say 'please do me a favour because I did a favour to you." (Researcher)*

> *"…we were going for [a high impact journal], and then we were not writing the paper we were spending all our time trying to get the editor and the reviewers that they knew that the paper was coming so that once the paper was there all these people were involved and engaged*

> *and then it was either accepted or not. So it really depends on who you know and who you don't know. And that's why sometimes I start behaving like that… 'Oh I would like to have a paper in one year in that journal' and I start writing them and seeing them at conferences (laughs) 'Hi, yes I'm thinking about submitting a paper what do you think about the idea' and it really really helps. So it's really not as unbiased as you would like it to be." (PostDoc)*

Nonetheless, most expressed discomfort, frustration, or loss of confidence from the advantages that research networks played on peer review for funding and publication.

> *"I feel like I have less and less confidence in publishing with the fact that 'who is going to be the reviewer?' 'Is he biased?' 'Is it the journal?' It's like some politics that you.... I don't always believe that the best results are published in the best journals." (Researcher)*

> *"…it's just the people that have the money that get the money. Because they're all in the commissions or they have co-workers or close collaborators that are in the commissions, and they just give each other money all the time." (RCC)*

Refuting these ideas, funders and editors explained that their review process was organised to minimise conflicting interests. One editor mentioned that believing that good relationship with editors would help manuscript acceptance was "Wishful thinking". She explained that strict policies against conflicting interests and the weight of external peer-reviewers in the decision cancelled what good networking could have created. To support her claim, she explained that she rejected the manuscript of her best friend not long ago. The discrepancy between the perspectives of researchers and the perspective of funders and editors makes this issue of unfair advantages difficult to resolve.

Punitive, not preventive

Added to the above problems, the worry that the scientific culture focused on punishment rather than prevention was raised by some interviewees. Even though this issue was only mentioned by a few interviewees, their perspectives raised questions which are scarcely addressed in the current integrity discourse. A policy maker or influencer worried about the lack of a second chance for researchers convicted of misconduct. He explained that once misconduct is proven, universities generally ban or shame the convicted researcher without offering any later contact, support, or chance for retaliation. He stated that "*This* *kind of unwillingness of the research system to forgive, not to forget, to forgive, really troubles me.*" This interviewee supported that in some cases, institutions would benefit from rehabilitating deviant researchers and involving them in integrity training later on. He believed that this would lead to higher relevance of integrity training, and would avoid that researchers who committed misconduct simply move on to a new university without any kind of follow up or notice — an issue that often happens in Europe where misconduct cases are not always disclosed publicly.

A GENERAL RESISTANCE TO CHANGE

In the final portion of our interview, we asked researchers '*who* they believe was responsible for promoting integrity'. Although selected actor-specific responsibilities were mentioned, we quickly realised that integrity was generally seen as a shared responsibility in which all actor have a role to play.

> *"So I think that it's a broad ecosystem, and everyone has a role to play in that." (EP)*

> *"I'm not going to say one person. I think it's an extremely complicated theme, and extremely complicated idea, concept… So you cannot focus on one person. You need to target a lot of people." (RCC)*

> *"Everybody [laugh]. Everybody has their share of responsibility, of course" (PMI)*

This sharing of responsibilities, however, appeared to downplay individual responsibilities and to trigger a shared feeling of helplessness. For example, researchers believed that, to survive in the current system, they had to play by the rules of the game, even if they disagreed with such rules. Institutions felt powerless on their own, and some interviewees even believed that it was unrealistic to believe in any drastic improvements.

> *"Everyone is behaving like this. Everyone is saying 'Let's go for the safe road because this is how it is otherwise I will never get funding', so…" (Researcher)*

> *"One institution cannot change that." (RIL)*

> *"I don't think we can expect, realistically speaking — but it's cynical maybe — we can expect the great world change. It couldn't change. You can try to make the ships sail a bit more in another direction but*

> *you cannot turn it. Therefore it's too deep. The idea is… The views on what science is and how people work is too deep. It might be cynical if I'm saying it now." (RCC)*

Lacking the empowerment or hope to take action, interviewees tended to transfer the root of the problems from one to another, creating a circle of blame which fostered frustration and distrust between actor groups. For instance, researchers had to cut corners because universities pressure them to publish; universities had to push researchers to publish because policy makers distribute funding to universities based on publication outputs; policy makers had to distribute funding based on publications because society wants a return on its investment, etc. In other words, each actor appeared to use the failures of higher actor groups to justify its personal inability to endorse best practices. But the complex interplay between actors also led to smaller circular criticism. For example, researchers criticized funders for evaluating them on quantity rather than on quality. But funders explained that even when they have policies in place to ignore quantity, peer-reviewers — who are themselves researchers — tended to cling to old quantitative metrics. Similarly, universities criticized that journals looked for hype rather than quality, but journals believed that the real problem was that universities used selective journals to evaluate researchers, not the decisions that journals take on what they choose to publish. Given that science is built around a community where all actors share the common goal of advancing knowledge, internal distrust and lost hopes for true change are necessarily a worry for the future.

SHORT SUMMARY OF FINDINGS

Our investigation of the problems that affect science and threaten integrity reveals a number of ideas on what needs to change in science. By involving different research actor in our analysis, we were able to discern the perspectives of different actors and to identify conflicting views, both within and between actor groups.

When discussing misconduct, interviewees explained that misconduct was far from black and white. Indeed, the core reasons for condemning misconduct seemed to differ between individuals and actor groups. We also noticed that the jargon which is normally used to discuss misconduct and integrity was not common to all research actors. Finally, although 'excessive pressure' was the factor that was most often mentioned as causing misconduct, many believed that the responsibility of misconduct ultimately resides in the researcher and that pressures cannot become excuses for bad practices.

We did not limit the discussion to strict misconduct. In fact, most interviewees were unfamiliar with genuine misconduct and were thus much more inclined to discuss the general problems which may deter research quality and integrity. In describing such problems, interviewees appeared to point to three general categories: Issues related to (i) personalities and attitudes, to (ii) awareness, and to (iii) the research climate. Issues related to personalities and attitudes were mentioned as potential targets for employers to consider, but were also admitted to be rather immutable. Issues linked to awareness generally discussed inadequate mentorship of research students and insufficient support on how researchers should meet integrity guidance. Finally, issues linked to the research climate highlighted problems which resulted from existing research environments and research cultures. The precariousness and scarcity of research careers, especially problematic for young researchers, were thought to be a major issue which aggravated competition and extended the threats from pressures and perverse incentives. Overspecialisation, lack of collaboration, and expectations that all researchers perform similarly also came into play as constraining the time available for research, intensifying pressures, and reducing the possibility for control and monitoring. Deeper in the cultures attached to research, the care and support given to researchers was also noted as being limited. Researchers were expected to participate in science out of passion, and thus to devote themselves without expecting personal benefits, but this perspective impacted the well-being and personal satisfaction of researchers. A general culture of profit, intolerance for failure, and expectations of extraordinary results added up to fuel a culture of 'publish-or-perish'. The overwhelming pressure to publish further seemed to shape the relationships that researchers have with one another. Competitiveness, hierarchy, and alliances were described and believed to influence how research was planned, performed, and reported.

Finally, when asked about responsibilities for change, interviewees revealed a shared feeling of helplessness towards current problems. They felt that issues were caused by inadequate decisions of different actors, and thus felt frustrated and lost their trust in other actor groups.

DISCUSSION

The present chapter reveals a rich account of various stakeholders' perspectives on misconduct and other problems of research. While it is technically impossible to integrate all diverse and sometimes inconsistent responses in a well-structured discussion, we would like to highlight three main findings which provide insights for the next steps towards better science: revisiting research assessment, empowering researchers, and fostering inter-actor dialogue.

First, the revision of research assessment needs to become central to the integrity discourse. Our respondents clearly indicate that definitions and assessments of success in science are not innocent, and that they impact research practices (Chapter 3). While we understand the strong emphasis on metrics from a pragmatic point of view, in practice, our participants considered reductionistic metrics as imprecise, disruptive, and at the very heart of most problems afflicting science. Without discrediting excellent science that yields remarkable metrics, we must recognise that excellent science does not necessarily translate into such metrics, and probably most often does not. Current output metrics thus provide, at best, a reductionist picture of the qualities and merits of researchers which may dissuade researchers from investing in practices that benefit science without increasing metrics, such as openness, transparency, and collaboration. In fact, researchers who commit to good science regardless of short term high-impact outputs may place their very existence as a scientist (i.e. their scientific career) at risk, rather than advance it. Wide-spread expectations of extraordinary results further add to the problem, not only by suggesting that extraordinary science should be the norm — a paradox in itself — but also by devaluing negative findings and small-steps-science, both of which are key to advancing knowledge. And yet, current assessments were also said to ignore — even inhibit — high risk innovation, originality, and diversity. Considering all this, it is obvious that research assessment must be addressed. A number of recent initiatives, such as the Declaration on Research Assessment (DORA; American Society for Cell Biology, 2013), the Leiden Manifesto (Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015), the Metric Tide (Wilsdon et al., 2015) (Wilsdon et al., 2015), or the Hong Kong Principles for Assessing Researchers (Moher et al., 2019), and numerous scientific editorials and public fora (e.g., Bryce, Dowling, & Lucey, November 26, 2018; Gadd, 27 September 2018; Holtrop, 29 November 2018; "A kinder research culture," 2019) are important pioneers in exposing the challenges of current assessments. Our findings echo these challenges and further link current problems to integrity failures, thereby reinstating that research assessment must become central to the discourse on research integrity.

Second, our findings suggest that approaches to foster integrity should focus on changing research climates rather than solely targeting individual behaviours. Our respondents perceived that research climates play a crucial role on research practices and integrity, a finding that is corroborated in most research on research integrity (Chapter 1). Nonetheless, the majority of approaches aiming to tackle misconduct capitalise on the knowledge and awareness of researchers (Chapter 1). Generally through training or codes of conduct, these approaches aim to discourage scientists from behaving badly. This person-centred perspective has profound implications on the way we perceive integrity. Not only does it ignore the dissonance between what researchers know they should do (i.e., integrity) and what helps them survive in their career (i.e., success), as described above, but it also transfers the burden of integrity on researchers — especially young researchers who are the main target of integrity training. In light of the high pressures, high demands, and lack of support that already afflict young researchers, it seems obvious that approaches to foster integrity need to better consider the climate in which researchers operate, the pressures it exerts, and the conflicts it entails. In this regard, training and education might need to shift their focus from *compliance* (i.e. what not to do) to *empowerment* (i.e., how to do great science) and *resilience* (i.e., how not to give in to cultural pressures that incentivise inferior practices). Training should aim to equip researchers to understand how they can promote good science without jeopardizing their personal success as well as cultivate resilience and skills to allow smoother migrations from research to alternative careers. Likewise, support and consideration of the person behind the research — something that has also been found missing in past works (e.g., Heffernan & Heffernan, 2019) — should be prioritized. Ultimately, if young researchers who are adamant about good science are empowered and resilient to the current issues of academia, they will have a greater chance of surviving the precarious career cycle, of becoming activists for good science, and of shifting the cultures that currently disrupt integrity.

Finally, our findings reinforce the need for inter-actor dialogue in discussions on research integrity. When describing success in science, we argue that a comprehensive inter-actor dialogue is needed to combine different meanings and expectations of scientific success (Chapter 3). Similarly, when discussing problems of science with multiple actors, we understood not only that perspectives differ from one actor to the next, but that the lack of inter-actor discussion leads to a circle of blame in which no one feels able to tackle the problems. Even though actors depend on one another, the opportunities to discuss and share decisions between them are limited, especially for early career researchers. This segregation leads to misunderstandings, false beliefs, and missed opportunity for joint actions. As researchers, we were ourselves surprised to realise that pressures also affect institutions, funders, editors, and policy makers. We thus believe that the best way forward is to create a forum for participatory decisions on topics of success, assessment, climates, and integrity. Prioritising opportunities for inter-actor dialogue and actively seeking the voice of overlooked actors can only help reduce victimisation and blame and promote well-considered joint action.

CONCLUSIONS

Our findings shed light on the complex interplay between success and research integrity. Involving not only researchers, but a wide range of actors who hold different roles in science, we show that there is great tension between what researchers should do to advance science, and what they must do to be successful. This finding resonates with debates that have been taking place in the past few years. But despite heated discussion, initiating changes in research assessments takes time, effort, and broad coordination. Our findings extrapolate a few action points which might help coordinate such changes. First, assessments of success must be tackled and must become central to the integrity debate. Second, approaches to promote better science should be reassessed: not only should they consider the impact of the climate on research practices, but approaches which focus on researchers should also redefine their objective to empower and support researchers rather than to capitalize on their compliance. Finally, inter-actor dialogues and shared decision making are crucial to building joint objectives for change. Such dialogues should actively seek the voices of parties which are forgotten from the current discourse, and should genuinely aim to construct a collective understanding of the problem so that actors can join forces for change.

REFERENCES

Adeleye, O. A., & Adebamowo, C. A. (2012). Factors associated with research wrongdoing in Nigeria. *Journal of Empirical Research on Human Research Ethics, 7*(5), 15-24. doi:10.1525/jer.2012.7.5.15

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics, 13*(4), 437-461. doi:10.1007/s11948-007-9042-5

Antes, A. L., Brown, R. P., Murphy, S. T., Waples, E. P., Mumford, M. D., Connelly, S., & Devenport, L. D. (2007). Personality and Ethical Decision-Making in Research: The Role of Perceptions of Self and Others. *Journal of Empirical Research on Human Research Ethics: An International Journal, 2*(4), 15-34. doi:10.1525/jer.2007.2.4.15

Babalola, Y. T. (2012). Awareness and incidence of plagiarism among undergraduates in a Nigerian Private University. *African Journal of Library Archives and Information Science, 22*(1), 53-60.

Bailey, C. D. (2015). Psychopathy, academic accountants' attitudes toward unethical research practices, and publication success. *Accounting Review, 90*(4), 1307-1332. doi:10.2308/accr-50970

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review, 1*(1), 17. doi:10.1186/s41073-016-0024-5

Brown, R. P., Tamborski, M., Wang, X., Barnes, C. D., Mumford, M. D., Connelly, S., & Devenport, L. D. (2011). Moral credentialing and the rationalization of misconduct. *Ethics and Behavior, 21*(1), 1-12. doi:10.1080/10508422.2011.537566

Bryce, C., Dowling, M., & Lucey, B. (November 26, 2018). To truly judge the quality of research, read it. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/blog/truly-judge-quality-research-read-it

Buljan, I., Barać, L., & Marušić, A. (2018). How researchers perceive research misconduct in biomedicine and how they would prevent it: A qualitative study in a small scientific community. *Accountability in Research, 25*(4), 220-238. doi:10.1080/08989621.2018.1463162

Davies, S. R. (2019). An Ethics of the System: Talking to Scientists About Research Integrity. *Science and engineering ethics, 25*(4), 1235-1253. doi:10.1007/s11948-018-0064-y

Davis, M. S., Riske-Morris, M., & Diaz, S. R. (2007). Causal factors implicated in research misconduct: Evidence from ORI case files. *Science and Engineering Ethics, 13*(4), 395-414. doi:10.1007/s11948-007-9045-2

Davis, M. S., Wester, K. L., & King, B. (2008). Narcissism, entitlement, and questionable research practices in counseling: A pilot study. *Journal of Counseling and Development, 86*(2), 200-210. doi:10.1002/j.1556-6678.2008.tb00498.x

Decreet betreffende de organisatie en financiering van het wetenschaps en innovatiebeleid [Decree concerning the organization and financing of science and innovation policy]. Art. 63/1, Vlaamse Regering [Flemish Government] (2009). Retrieved from https://codex.vlaanderen.be/Portals/Codex/documenten/1018085.html

DuBois, J. M., Anderson, E. E., Chibnall, J., Carroll, K., Gibb, T., Ogbuka, C., & Rubbelke, T. (2013). Understanding Research Misconduct: A Comparative Analysis of 120 Cases of Professional Wrongdoing. *Accountability in Research, 20*(5-6), 320-338. doi:10.1080/08989621.2013.822248

Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE, 4*(5), e5738 (5731–5711). doi:10.1371/journal.pone.0005738

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE, 5*(4), e10271. doi:10.1371/journal.pone.0010271

Fanelli, D., Costas, R., & Larivière, V. (2015). Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity. *PLoS ONE, 10*(6), e0127556. doi:10.1371/journal.pone.0127556

Fang, F. C., Bennett, J. W., & Casadevall, A. (2013). Males Are Overrepresented among Life Science Researchers Committing Scientific Misconduct. *mBio, 4*. doi:10.1128/mBio.00640-12

Franzoni, C., Scellato, G., & Stephan, P. (2011). Changing Incentives to Publish. *Science, 333*(6043), 702–703. doi:10.1126/science.1197286

Gadd, E. (27 September 2018). Better, fairer, more meaningful research evaluation – in seven hashtags. *LSE Impact Blog*. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2018/09/27/better-research-evaluation-in-seven-hashtags/

Ghias, K., Lakho, G. R., Asim, H., Azam, I. S., & Saeed, S. A. (2014). Self-reported attitudes and behaviours of medical students in Pakistan regarding academic misconduct: A cross-sectional study. *BMC Medical Ethics, 15*(1), 43. doi:10.1186/1472-6939-15-43

Heffernan, T. A., & Heffernan, A. (2019). The academic exodus: the role of institutional support in academics leaving universities and the academy. *Professional Development in Education, 45*(1), 102-113. doi:10.1080/19415257.2018.1474491

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature News, 520*, 429-431. doi:10.1038/520429a

Holtrop, T. (29 November 2018). The evaluative inquiry: a new approach to research evaluation. *LSE Impact Blog*. Retrieved from http://blogs.lse.ac.uk/impactofsocialsciences/2018/11/29/the-evaluative-inquiry-a-new-approach-to-research-evaluation/

Hren, D., Vujaklija, A., Ivanišević, R., Knežević, J., Marušić, M., & Marušić, A. (2006). Students' moral reasoning, Machiavellianism and socially desirable responding: Implications for teaching ethics and research integrity. *Medical Education, 40*(3), 269-277. doi:10.1111/j.1365-2929.2006.02391.x

Kaiser, K. A., Cofield, S. S., Fontaine, K. R., Glasser, S. P., Thabane, L., Chu, R., . . . Allison, D. B. (2012). Is funding source related to study reporting quality in obesity or nutrition randomized control trials in top-tier medical journals. *International Journal of Obesity, 36*(7), 977-981. doi:10.1038/ijo.2011.207

A kinder research culture is possible. (2019). *Nature, 574*, 5–6. doi:10.1038/d41586-019-02951-4

Kraemer Diaz, A. E., Spears Johnson, C. R., & Arcury, T. A. (2015). Perceptions That Influence the Maintenance of Scientific Integrity in Community-Based Participatory Research. *Health Education and Behavior, 42*(3), 393-401. doi:10.1177/1090198114560016

Lundh, A., Krogsbøll, L. T., & Gøtzsche, P. C. (2012). Sponsors' participation in conduct and reporting of industry trials: a descriptive study. *Trials, 13*, 146. doi:10.1186/1745-6215-13-146

Martinson, B. C., Anderson, M. S., & De Vries, R. (2005). Scientists behaving badly. *Nature, 435*(7043), 737-738. doi:10.1038/435737a

Miller, A., Shoptaugh, C., & Wooldridge, J. (2011). Reasons not to cheat, academic-integrity responsibility, and frequency of cheating. *Journal of Experimental Education, 79*(2), 169-184. doi:10.1080/00220970903567830

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2019). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *OSF Preprints*. doi:10.31219/osf.io/m9abx

Mumford, M. D., Devenport, L. D., Brown, R. P., Connelly, S., Murphy, S. T., Hill, J. H., & Antes, A. L. (2006). Validation of ethical decision making measures: Evidence for a new set of measures. *Ethics and Behavior, 16*(4), 319-345. doi:10.1207/s15327019eb1604_4

National Institute of Health. (2018). *Research Misconduct - Definitions*. Retrieved from https://grants.nih.gov/policy/research_integrity/definitions.htm

Okonta, P., & Rossouw, T. (2013). Prevalence of scientific misconduct among a group of researchers in Nigeria. *Developing World Bioethics, 13*(3), 149-157. doi:10.1111/j.1471-8847.2012.00339.x

Peters, A. (2019). *Public funding of universities, and indicators used in the funding systems : a Euregional perspective.* Université de Liège, Liège, Belgium. Retrieved from https://matheo.uliege.be/handle/2268.2/8014

Pupovac, V., & Fanelli, D. (2014). Scientists Admitting to Plagiarism: A Meta-analysis of Surveys. *Science and Engineering Ethics, 21*(5), 1331-1352. doi:10.1007/s11948-014-9600-6

Shrader-Frechette, K. (2011). Climate Change, Nuclear Economics, and Conflicts of Interest. *Science and Engineering Ethics, 17*(1), 75-107. doi:10.1007/s11948-009-9181-y

Singh, H. P., & Guram, N. (2014). Knowledge and attitude of dental professionals of North India toward plagiarism. *North American Journal of Medical Sciences, 6*(1), 6-11. doi:10.4103/1947-2714.125854

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication Pressure and Scientific Misconduct in Medical Scientists. *Journal of Empirical Research on Human Research Ethics, 9*(5), 64-71. doi:10.1177/1556264614552421

Wester, K., Willse, J., & Davis, M. (2010). Psychological climate, stress, and research integrity among research counselor educators: A preliminary study. *Counselor Education and Supervision, 50*(1), 39-55. doi:10.1002/j.1556-6978.2010.tb00107.x

Wilsdon, J., Liz Allen, Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from https://re.ukri.org/documents/hefce-documents/metric-tide-2015-pdf/

Wright, D. E., Titus, S. L., & Cornelison, J. B. (2008). Mentoring and research misconduct: An analysis of research mentoring in closed ORI cases. *Science and Engineering Ethics, 14*(3), 323-336. doi:10.1007/s11948-008-9074-5

Zacharewicz, K. J. T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment*. European Commission. [ISBN 978-92-79-68715-0 - ISSN 1831-9424]. Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC101043/kj1a27837enn.pdf

Chapter 5

For the sake of science: Are current research careers changing the norms of science?

CONTRIBUTIONS

Conceptualization: Noémie Aubert Bonn, Wim Pinxten[1], Raymond De Vries[2]

Funding acquisition: Wim Pinxten. Funding granted by the Bijzonder Onderzoeksfonds (BOF) 15NI05

Project administration: Noémie Aubert Bonn, Wim Pinxten

Methodology: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries

Resources: Noémie Aubert Bonn, Wim Pinxten (general); Raymond de Vries, Melissa S. Anderson[3], and Brian C. Martinson[4] (focus group guide); Ines Steffens[5], Inge Thijs[6], and Igna Rutten[6] (focus group organisation and help in recruiting participants)

Investigation: Noémie Aubert Bonn

Data curation: Noémie Aubert Bonn

Formal analysis: Noémie Aubert Bonn

Visualization: Noémie Aubert Bonn

Validation: Noémie Aubert Bonn, Wim Pinxten, Raymond De Vries

Supervision: Wim Pinxten

Writing – original draft: Noémie Aubert Bonn

Writing – review & editing: Noémie Aubert Bonn, Wim Pinxten (intermediate and final versions), Raymond De Vries (intermediate and final versions)

1. Department of Healthcare and Ethics, Hasselt University, Hasselt, Belgium
2. Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor (MI), USA
3. Department of Organizational Leadership, Policy, and Development, University of Minnesota, Minneapolis (MN), USA
4. HealthPartners Institute, Minneapolis VA Medical Center, Center for Care Delivery and Outcomes Research, and University of Minnesota, Department of Medicine, Minneapolis (MN), USA.
5. European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden
6. Faculty of Biomedical and Life Sciences, Hasselt University, Hasselt, Belgium

ABSTRACT

Science has changed tremendously in the past decades. With growing investments in scientific activities, the academic workforce is believed to have nearly doubled in 20 years. This shift has altered the career pathways of researchers and the mechanisms for funding science, alterations that have brought with them changes in the perspectives and ambitions of researchers.

In this chapter, we revisit the norms of science proposed by Robert K. Merton in the 1960's and question the place of the scientific community in today's academia. Using recent interviews and focus groups with diverse research actors from Flanders, Belgium, we show that, although the norms of communalism and disinterestedness remain important today, applying these norms in practice proves more complex. In the present environment, career assessments tend to be based on extraordinary findings, on individual achievements, and on peer-to-peer competition. As a result, researchers often feel compelled to use self-promotion strategies even though these strategies conflict with the norms of communalism and disinterestedness. Not all researchers, however, are willing to change their practices to adapt to the realities of current research careers. Those who refuse to do so struggle with the conflict, lose trust in science, and often decide to leave academia altogether.

The precarious setting of early academic careers paired with inadequate research assessments, has the potential to encourage researchers to make strategic decisions that may undermine the integrity of science. For the sake of maintaining researchers' devotion and selfless motivation for advancing science, a profound revision of the way in which we assess success in research is required.

INTRODUCTION

Academia is growing tremendously. Over the past forty years, investments in Research and Development increased by around 30%, now representing on average 2.4% of the Gross Domestic Product (GDP) of countries included in the Organisation for Economic Co-operation and Development (OECD, 2019). This growth naturally translates in an expanding workforce. Since the 2000's, the number of researchers is estimated to have increased by over 65% in the European Union and to have more than doubled in some Asian countries (National Science Board, 2018). In 20 years, the total number of researchers has nearly doubled worldwide, growing from 4.58 million in 1997 to an impressive 8.68 million in 2017 (UNESCO Institute for Statistics). The fast growth in the research workforce, however, also generated an imbalance between junior and senior positions. In the United States, for instance, only 17% of PhD graduates are likely to obtain tenure or tenure-track positions within five years of obtaining their PhD (National Science Board, 2018), even though the majority aspires to such positions (Sauermann & Roach, 2012). This picture varies between countries, but the chances of obtaining a tenured academic position after completing a PhD remain low worldwide (Anonymous, 2010; Debacker & Vandevelde, 2016; "Many junior scientists," 2017; Martinson, 2011). In such an environment, funders and institutions need to ensure that only the very best make it to the top. Nonetheless, the selection process often fails to honour the intensive and high-risk investments that many researchers have made in the first years of their scientific career. The imbalance between junior and senior positions adds to the frustration of the many excellent scientists who are offered no opportunity in academia and has accelerated demands for transparent and objective selection criteria. As a result, evaluations of researchers quickly moved from the informal renown and peer assessments — the norm fifty years ago — to more seemingly more objective and comparable systems based on pre-defined metrics. Quantifiable metrics that allow clear and easy ranking, such as numbers of publications, citations, and grant success are now generally the norm, or at least play a role in most assessment schemes. But the use of quantifiable metrics has also been criticized, condemned for an overreliance on measures that proved rigid, uninformative, reductionistic, and open to manipulation (American Society for Cell Biology, 2013; Gingras, 2016; Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015; Moher et al., 2019; Wilsdon et al., 2015).

Low chances of success, coupled with constant comparison with one's peers, create a fierce sense of competition among scientists. This fierce, individualized competition distracts researchers from the common pursuit of advancing science; a pursuit based in a shared set of norms and values. The norms of science, as described by the American sociologist Robert K. Merton almost eighty years ago, play a role in nurturing a shared understanding of concepts such as merit, quality, and integrity, in addition to uniting researchers in a common goal of advancing knowledge. Yet, in today’s increasingly selective academia, the communal sense of science conflicts with the rivalry that governs the interactions between scientists. Most specifically, current methods of evaluating researchers, with its focus on exceptional findings, individual performance, and competition, poses a threat to researchers’ sense of community and their appreciation of the shared norms of science.

In the current chapter, we explore the sense of community and the norms and values that researchers share with one another in today’s academic climates. Using qualitative data gathered between 2018 and 2019 in the Flemish research system, we compare current views to historical accounts from Robert K. Merton and Warren O. Hagstrom on the norms of disinterestedness and communalism[12]. Showing how the norms of science have shifted in the context of current methods of assessing research careers, we then explore the impact of these changes on the scientists of today and on the future of science.

HISTORICAL ACCOUNTS OF THE COMMUNITY IN SCIENCE

In 1942, Robert K. Merton proposed that four norms appeared to guide researchers in their profession: universalism, organized skepticism,

[12] In the original documents, Merton refers to ‘communism’. To avoid confusion with the political meaning of this term, we will refer to Merton’s ‘communism’ using the term ‘communalism’ in the present thesis.

disinterestedness, and communalism (Merton, 1942). Despite having been disputed by some, Merton's works on the sociology of science are still extensively cited today, and the four norms appear to have stood the test of time (Anderson, Ronning, De Vries, & Martinson, 2010). Two of these norms, disinterestedness (with its counternorm of self-interestedness) and communism (or communalism, with its counternorm of secrecy or solitariness), reinforce the idea that scientists work towards a common goal that is shared within the scientific a community.

Barely a decade later, Warren O. Hagstrom complemented Merton's normative works by discussing the place of the community in science. Using interviews he conducted with American scientists from the 'exact' sciences, Hagstrom further exemplifies the place that Mertonian norms held in researchers of this period (Hagstrom, 1975).

Hagstrom and Merton both describe several behaviours and attitudes which capture researchers' endorsement of the norms of communalism and disinterestedness. For instance, they describe that researchers express disapproval for self-interest through a disdain for the desire for recognition and praise. Nonetheless, Hagstrom and Merton also concede that scientists' behaviours are often driven by a desire for recognition and priority of discovery. In acknowledging this conflict, Merton's and Hagstrom acknowledge the ambivalence between the desire for recognition and the selflessness which is expected of scientists.

Merton describes this inner conflict as a "contest between the values of recognized originality and of modesty", asserting that "great modesty may elicit respect, but great originality promises everlasting fame". (Merton, 1957) p. 308. In line with this view, Hagstrom notes that scientists usually feel that it improper or even degrading to seek reward, yet they still appear to be driven by a desire for recognition.

Interestingly however, neither Merton nor Hagstrom condemn the desire for recognition. Instead, both sociologists state that the desire to be recognized in the scientific community may strengthen the bond that keeps scientists together. For instance, Merton suggests that the desire for recognition is "anything but personal and individual" (pp. 400), rather serving as a mechanism of social validation. In seeking the approbation from the scientific community, scientists strengthen their bond with it. Hagstrom offers a similar perspective. In his view, scientific publications act as a proof of acceptance and praise of one's ideas within the scientific community. Following this logic, Hagstrom suggests that research institutions honour communal recognition by rewarding publication, and that "*organizational pressures on scientists to publish are more likely to reinforce the power of the scientific community than to supplant it*" (p. 54). According to him, the fact that scientists continue to publish even when they are not compelled to (e.g., after tenure) suggests that publications are a proxy for internal recognition and that they reinforce the sense of community. Nonetheless, Hagstrom is suspicious of other forms of reward in science. Cash incentives, for instance, do not bind scientists to one another since they allow personal gratification outside of the academic community (i.e., through the purchase of goods and services). In such cases, scientists have a contractual relationship with research, pursuing science in exchange for personal benefits (i.e., external reward). In this sense, Hagstrom is concerned that financial incentives or other incentives which do not rely on internal recognition may disrupt the sense of community and may lead to science which is ultimately performed for personal motives rather than for common ones.

In the following sections, we question whether science of the 21st century continues to nurtures a sense of community by valuing the norms of disinterestedness and communalism. Using data from interviews and focus groups with a variety of research actors from the Flemish research scene, we compare the views and perspectives that govern current science to the historical perspectives highlighted by Merton and Hagstrom. Full methods and additional findings are available in Chapters 3 and 4.

CURRENT VIEWS ON THE 'REAL SCIENTIST'

In early discussions, Merton and Hagstrom noted that researchers expressed disapproval of the thirst for praise and recognition. Our interviews and focus groups provide evidence to support that, even in the 21st century, a scientist is thought to be someone interested in advancing science rather than in gaining personal benefit. Indeed, scientists were described as "*intrinsically motivated to do good science*" (PhD student) and as driven by a passion for, and devotion to, the greater goal of advancing knowledge (Chapter 3). Some interviewees even expected researchers to be so passionate about science that they should be willing to sacrifice their personal life and comfort for the advancement of science.

> *"I am quite amazed when I listen to some discussions that are triggered by PhDs students and PostDoc, where they consider their PhD work to be kind of a job. This is not a job. This is an opportunity to pursue a scientific interest. I would never call 'doing science' a job. […] Because, yeah, being passionate about science is almost like being an artist. You live in poverty because you want to pursue your art." (Policy maker/influencer)*

The passion and disinterestedness described by Merton and Hagstrom thus appeared to reach a second level in current academia: Researchers should not only control their desire for praise and reward, but they should also accept to undermine their personal comfort to pursue science.

Not all respondents agreed with this view. Researchers, research students, and former researchers were particularly vocal about their disagreement, arguing that passion and devotion alone are unsustainable and that the lack of stability and benefits of current research careers is worrisome. One funder explained that, in the Flemish system (as in many other international research systems), structural funding for research is limited, leaving the researchers to depend more and more on external competitive funding. Yet, most competitive funding schemes provide funding for only two to four years, after which researchers must apply for new funding to continue the work or even maintain their employment. This continuous cycle was mentioned many times during our interviews, including by research institutions leaders who were unable to find an alternative to solve the issue. The very low percentage of young researchers who are able to maintain academic careers was another recurrent topic of discussion. One respondent explained that only 10 to 20% of graduating PhD students will continue in academia, even though most aspire to it. Post-doctoral researchers, who felt they were at a "tipping point" between academic and non-academic careers, admitted that the insecurity of their position was stressful and frustrating, especially since many have to care for a young family. Former researchers mentioned the lingering insecurity, daunting competition, and high demands of research careers as the main reason for leaving academia.

> *"If I go for the academic world, I'm going to have to tell my wife, that was pregnant [at the time…], I have to tell my wife "well we're going to a financially uncertain situation for at least 10 to 15 years. And maybe when I'm 30 or 35 and I have said no to you an enormous amount of times, I'm going to be so successful that I can say 'It's ok now, we can pay the bills.' But I'm still going to say no to you because I have to compete with the other people. Whereas if I choose another life or career, you get, for example a contract that lasts for your entire life, and you can build your life. You can start building your life. You can settle in a way, you can… You can make plans. Whereas in the academic world you can only make plans for 2 or 3 years. And that was the kind of life that I didn't want to <live>. And that was the kind of life that I didn't want to live. And I also didn't find that the reward in the academic world was sufficient. The price was too high for what you get. It was too expensive. So I chose something that was economically more balanced." (Researcher who changed career)*

Looking at these perspectives, together with the idea that scientists should be driven by selfless devotion to build knowledge, offers an alternative view of the tension between personal and communal gain. In current academia, researchers are asked to give up minimal comfort and stability for a career in which their ultimate odds of success almost seem negligible. Far from being selfish and self-driven, young researchers simply appear to consider that their passion for science does not justify the sacrifices they would need to make to (attempt to) remain in academia. This perspective was typically expressed by young researchers, research students, and former researchers, but policy makers or influencers and other actors who were aware of the problem also reflected on the difficult conditions of current academic careers.

> *"I have a choice. Why would I go there. Why would I choose to do, to start a career in an area that positions are limited, promotions are limited, high positions are limited, and it's precarious? I have to bring funding, I have to get contracts... So... it's not only the lack of interest — if it is there — from the younger generations, it is also what is the tomorrow. And this is part also, I think, of a scientific governance and a scientific culture issue. Those... We should not consider that researchers are somewhat a different part of the population or that they are saints, that they will sacrifice their wellbeing and their participation in the pleasures of a good economy just because they love science. I think this is very naive." (Policy maker/influencer)*

The increasing insecurity and precariousness of current research careers appear to have opened a new discussion around expectations of disinterestedness. Although the norm is still expected of scientists today, the harsh and insecure living conditions imposed on young researchers now also raise concern. This new perspective may result from changes in the way research careers are organised. For instance, it is possible that research careers increasingly involve individuals who need resources to support their non-academic life (e.g., young parents, individuals from varied socio-economic backgrounds, etc.; Maher & Sureda Anfres, 2016). It is also possible that attitudes have changed, and that the 'sanctity' of research and academia is being questioned, or that there is a growing awareness of the impact that precarious and insecure employment may have on the wellbeing of researchers (see for example Anonymous Academic, 2018a; Evans, Bira, Gastelum, Weiss, & Vanderford, 2018; Levecque, Anseel, De Beuckelaer, Van der Heyden, & Gisle, 2017; "The mental health of PhD," 2019; Pain, 2017; Woolston, 2019). Regardless of the reason driving this shift, it is clear that the contractual theory advanced by Hagstrom — where scientists gain external benefits from performing science (i.e., an attractive salary, stability, family benefits, etc.) — is being discussed more openly now than it was half a century ago.

CURRENT RESEARCH ASSESSMENTS AND THE RESEARCH COMMUNITY

We have shown that even though scientists are still expected to embody the norm of disinterestedness, the difficult conditions of early academic careers raised concerns about the extent of the sacrifice that they are expected to make. One factor behind the precarious conditions of early academic careers comes from the way in which researchers are assessed and resources are distributed. Research assessments have been heavily criticized in the past years, especially because of the reductionistic metrics they often use. But reductionistic metrics are not the only aspect of research assessments that can be problematic.

In this section, we question the impact of three key features of current research assessments on the sense of community shared between researchers. Specifically, we will look at the expectation of extraordinary results, the focus on the individual, and the competitiveness of current research assessments.

Expecting the exceptional

There is wide agreement that excellence is what researchers should, and must, aim for. Yet despite being extensively used in any funding or assessment scheme, the term 'excellence' is vague and rarely explicitly defined (Moore, Neylon, Paul Eve, Paul O'Donnell, & Pattinson, 2017). Despite being vague and rarely explicitly defined, the term 'excellence' is used extensively in any funding or assessment scheme. Excellence is what researchers should, and must, aim for. Our interviews and focus groups revealed that, in today's academia, excellent research is often a synonym for 'ground-breaking', 'innovative', and 'exceptional' science (Chapter 3). Editors and publishers discuss the wish to publish research that is "*novel enough*" and that provides "*a really big step forward*", funders emphasize the place of "innovation" in successful applications, and university leaders concede that exceptional publication records (i.e., papers published in highly prestigious journals) can compensate for lower productivity indicators. Conceiving excellence as something that 'stands out' is nothing new. Merton and Hagstrom also described the passion that scientists have in pursuing big and challenging questions, as well as their natural drive towards extraordinary discoveries. For example, Hagstrom points out that in fields like physics and mathematics, dozens of research teams have been tackling the same questions for years in the hope of uncovering the elements which evaded everyone else. And indeed, big leaps of knowledge do advance science, and in some way, they may strengthen its community by forcing broader collaboration. Yet, by tying 'exceptionality' to all forms of research funding and position, we build an unrealistic threshold that detaches researchers from reality. The exception becomes the norm — a contradiction in itself — and 'good but ordinary research' becomes unworthy of reward or support.

This perspective creates a tension with the goal of building, validating, and rigorously advancing knowledge together as a community. In focusing on the exceptional, assessments ignore and stigmatize the small steps that are essential for good science.

> *"The problem we have with the current research assessment system is the fact that, you know, it's really… success is measured by publishing in very selective journals that are looking for very important ground-breaking claims. You are incentivized to find these ground-breaking claims… And so you are incentivized to really get something*

> *that is extraordinary, and ground-breaking […] the problem we have is that at the moment it's only the big leaps that are being rewarded. So people are incentivized to find these big bold claims […] why would you go through the effort of really verifying your claims until this effect becomes very small at which point you can't publish it anymore because you really wanted the big effect to publish in the first place… These are the kinds of perverse incentives that we have in the system at the moment." (Editor or Publisher)*

> *"There is tension, there is obvious tension between the kudos, and the whole system that has been put in place where it's… you have to be special. It doesn't fit. It doesn't fit with the kudos! It doesn't fit with the universalism etc. So I think that that's where something is wrong. I don't have the solution, but that's what needs to be addressed! […] I would try to solve that tension that exists right there, to be able to go back to the other communalism, to the universalism etc. You know, the kudos itself. […] It's beyond community! It is also science for the sake of science. I mean let's face it! Here – and I understand perfectly – we're human beings again, people are doing science as a career. I'm sure that they still believe strongly in it, and I'm sure that they're there to change society, but in the end, the system makes it that they have to be good to get the funding to be able to do it. They have to be excellent. So again, it's back under this tension." (Editor or Publisher)*

In discarding the consideration for 'normal science', the lack of interest for negative results was also raised as a big problem for the quality and the advancement of science (see section on intolerance for failure in Chapter 4). Indeed, defining excellence as 'striking findings' was said to stigmatize unimpressive and negative findings, thereby contributing to selective reporting and to the artificial inflation of positive findings.

In short, current research assessments seem to evaluate researchers on their ability to generate extraordinary findings. In doing so, they ignore the bulk of 'normal but good science' and they fail to reward behaviours that contribute to the joint and communal advancement of knowledge.

VALUING INDIVIDUALS

A second problematic feature of current strategies of assessment is their tendency to value scientists individually even though scientific advances are generally the result of complex team efforts. Some distribution schemes — such as the REF (Research Excellence Framework) in the UK and the BOF (Bijzonder Onderzoeksfonds, special research funds) in Belgium — attempted to tackle this issue by assessing the number of publications and the cumulative impact factors produced in institutions or in departments rather than at the researcher level.

However, the indicators used in these distribution schemes tend to be reflected in the institution's expected key performance indicators at the researcher level. Indeed, interviewees explained that institutions often turn group indicators into individual Key Performance Indicators (KPI) that researchers must achieve to secure and advance their career. Focusing achievements at the researcher level, rather than on the team or the science itself, ignores the communal and collaborative nature of research and encourages researchers to think about their own Curriculum Vitae rather than about the functioning of their team. In our interviews, one former researcher even believed that the individualistic structure of academia placed so much emphasis on independence that researchers themselves lost interest in collaborating.

> *"In Dutch we say ZAP, zelfstandig academisch personeel [independent academic staff]. So they're academic staff, but independent… Independent academic staff. They're individuals who choose to sometimes work together. That's it. It's an individualistic group that, because of the necessity of their work, sometimes collaborate. But if they can, they will not…" (Researcher who changed career)*

The individual character of the current system was also criticized for isolating researchers from one another. Researchers and research students often felt "*on an island*" where they had to master every task on their own even though team efforts would be much more effective. They complained that current environments discouraged them from pairing efforts with their other researchers, even with colleagues down the hall.

> *"The advantage of academics is that you have many tasks, but this is also a disadvantage. Sometimes you have to do everything, you have to be good <in> English and <grammar>, in statistics, in everything, and... which is not always our expertise and also neither our interest…" (Researchers)*

> *"…there's not enough collaboration in research. […] It would be better if we could collaborate and think out good methodologies to do it on larger population and so on. But everyone is doing like "No, it is mine, it's mine" And then, yeah... I think it's a pity because the research would be better when there would be more collaborations. For me personally, to give a better feeling, but also for the population that we are doing it for." (PhD student)*

> *"[… In discussing with my supervisor] I said 'Ah maybe it's interesting to collaborate with this person because his field of expertise is in that' and [my supervisor] was like 'No, we're going to try first on our own.' And I was a little bit disappointed with this reaction because I thought the other researcher was always saying like 'Ah you're also doing the*

> *things that I actually also would like to do' and like, it was perfect to get to work together on it. But my promoter said No. Maybe if we really need him for the knowledge or the content then... So I was a little bit disappointed about that, but, yeah..." (PhD student)*

> *"Nobody feels like they're part of this team because of course… it didn't really feel like it was competition but more like everybody was just doing their own thing and surviving." (Researcher who changed career)*

Beyond encouraging isolation, a focus on individual achievements creates homogeneity and expectations that are the same for all researchers. Each researcher needs to publish a certain number of papers as first or last author, to obtain a certain level of impact factors, to supervise a certain number of PhD students, to teach a certain number of courses, and to win a certain number of grants. All researchers are thus asked to fit in the same box and to fulfil the same requirements. This approach disregards personal skills, preferences, and talents that could benefit teams and the overall productivity of research.

> *"If I was the boss in the university I would give the tenure tracks or the appointed researchers that are… that have a big talent for research, I would give them more research tasks and the people that are good at teaching which we also need because if we don't have students we don't have any money, I would give them less research tasks. Why do we all have to be… we can be complimentary yeah, we don't have to be all copies of each other." (Researcher who changed career)*

In our focus groups, one post-doctoral researcher had the luck of being in an atypical team which did not assess researchers individually, relying instead on measures of the outputs of the entire research group. According to his experience, not only did the team do equally well on the joint indicators they had to reach, but allowing researchers to 'be themselves' also added to the research process and the collegiality.

> *"I don't think we have strict KPIs. At least not in [my group] we don't have them, and I think it's very beautiful because we have [many] PhD students and they're all, or 99% of them are successful, but they are so different in being successful. Some are really being successful in the number of publications, some of them are really successful in the network they have with other companies, with other research institutes, some of them are really successful in the perseverance to do something really new and to make it happen […] So, they're so successful on so many different levels and I really like the fact that we don't judge them all in the same way because they can be themselves and be successful in the way that they want to be successful." (PostDoc)*

In comparison, teams where assessments discouraged personal differences were said to undermine the role of important yet non-leading actors in teams. One interviewee believed that good team players should also be given the chance to advance their careers without needing to become leaders.

> *"You need a lot of people who are not so very up front but they are also very important, and also in giving permanent positions, you should take them into account." (Research integrity office member)*

Laboratory technicians with whom we held a focus group expressed that, even though they are responsible for a lot of the work that happens in a research project, they virtually never receive authorship for it. Some were grateful for being in the acknowledgements and believed that authorship would not be useful to them, but others found it more problematic. For instance, they noted that "*the PI is always mentioned,* [even when] *he <sic> didn't do anything...*", while they need to demonstrate that they provided innovative ideas to be considered for authorship, on top of taking care of the animals, assisting with or performing the manipulations, reading and helping with the English, helping with the analysis, etc. The lack of formal recognition for the essential work of technicians and other research staff further confirms that praise and recognition of efforts are easier to receive for leading actors than for non-leading team members, regardless of the work involvement. This one-sided view, unfortunately, further increases the sense of hierarchy that governs academia and likely increases the competition for leadership positions. Internal competition, in turn, impacts the sense of community that is shared between the different members of a research team.

To summarize, targeting homogenized assessments at individual researchers ignores the essential place of team members, unique talents, and individuality in knowledge production. Every researcher must fit the box and fulfil the same requirements in order to be recognized. In addition, leadership and hierarchy come into play, bringing along frustrations and internal competition.

Competitive environment

The last feature of the current assessment strategy that has the potential to disrupt the norms of disinterestedness and communalism is the omnipresent reliance on competition. Merton and Hagstrom see competition as the desire for priority of discovery. In today's academia however, competition expanded far beyond the desire for intellectual priority as a fight for career opportunities developed. Indeed, even though the past decades yielded a tremendous growth of the research workforce, the growth largely ignored senior positions and tenured posts (Anonymous, 2010). As a result of this imbalance, early career scientists are in fierce competition for the limited permanent positions available. Even though competition was initially described as a driver to increase the pace of scientific discovery, its purpose and impact are now openly questioned (Anderson, Ronning, De Vries, & Martinson, 2007; Anonymous Academic, 2018b). Our interviewees worried about the impacts of excessive competition. Many described competition in science as '*fierce*', '*enormous*', or '*omnipresent*', even within departments and research teams. The resulting climate was described as a '*fight*', a '*struggle*', or a '*hostile environment'*.

One important problem with the omnipresence of competition in science is its encouragement of secrecy, the counternorm of communalism. To secure priority of discovery and opportunities, scientists refrain from sharing their data, ideas, and findings before they have formalized their contribution (e.g., through publication, patents, etc.). Young researchers expressed strong frustration with the lack of openness that competitive environments encouraged. Although they most often attributed the choice of secrecy to their supervisors, PhD students acknowledged being torn between their desire to be open and the fear of having their work stolen by competitors.

> "Participant A: *The problem with research is also, it's really a competition in research. I also have it now that I can't present on a congress because there are only three articles published on the subject I'm studying, so the supervisors are scared if I make a poster or I present that other researchers will get interested in the same topic, and then, if they publish first all I'm doing is a waste of time... […] And I think it's also a problem that no one wants to share their unpublished data because they are scared that someone else will go and take the data and will publish first and then, you don't have it anymore.*
> Participant B: *Yeah I completely understand the feeling because what we are doing it's also new so it's never been done and my promotor is always so reluctant to let me go and show the data to other people. […] he is always so scared that other people are going to steal his ideas... Sometimes I do understand, but sometimes I'm also like, I don't really like this kind of environment, it struggles with my personality a lot, I think. I have sometimes a very hard time with it." (PhD Students)*

Later in this same discussion, we directly probed the tension expressed between the desire to advance science and help patients, and the desire to be recognized for their work. In their responses, PhD students explicitly expressed the tension and the ambivalence they share between both values.

> "Interviewer: *I remember at the beginning, many of you talked about the translation of the results, well if another team does it, it's going to go to the patient faster right? In the end.*
> Participant B: *Yeah that's true.*
> Participant C: *But then I think you… you also said you have to publish to get credits like a researcher, so if someone else publishes your idea you don't get the recognition for it.*
> Participant A: *Or also your article will not be accepted because they can say 'Oh no, they showed it already' and then yeah... It's not that you don't work or that your work is lost, but… sometimes it feels a bit like <that> I think. Yeah, I'm at the beginning I didn't publish anything out of it, but…*
> Participant C: *Yeah, I think it's maybe the wrong motivator, that you need to publish, but I think you cannot deny that it exists.*
> Participant A: *And also, yeah your research also focus on data that's already published so publishing is an important part of research. Because if you don't have any literature to focus on it would be really difficult to start any research I think.*
> Participant B: *But it's definitely true, if another research group can do it ten times more fast then you and it gets out there five years earlier than yours, then why should you stop it, indeed it's completely true. These patients will benefit from it very well...*
> Participant D: *The patients benefit...*
> Participant A: *It's better to collaborate than just make a competition about it.*
> Participant B: *Yeah, indeed." (PhD Students)*

This tension seemed less prominent with more senior researchers, for whom fulfilment of KPIs could provide satisfaction without necessarily needing to yield translation into health or practice (Chapter 3). Although we remain careful about group comparison, the decreasing interest in research translation could suggest that those who survive longer in academia learn to value priority of discovery and competition over the desire to benefit the society. For example, one former researcher said that competition becomes "*inherent*" to researchers, "*It's part of who they are*". Expanding on this idea, post-doctoral researchers explained that the relationship between the researcher and the research changes over time. These relationships, be they collaborative or competitive, influence behaviours, opportunities, and openness.

> *"As a PhD student it's different because you don't really know a lot of people and you just have all those papers that you think 'Ah those are very nice' but you don't really know the research group behind it.*

> *Whereas that is the longer you are in academia these people either become competitors or collaborators and it becomes really difficult to position yourself into that community and still… Because in the end you still would like to present at the conferences and people are either accepting or not accepting based on whether or not they like you. So, yeah… I think it's difficult." (PostDoc)*

In short, the wish for priority of discovery has long existed and has been discussed as a trigger for competition by both Merton and Hagstrom. In the past few years, however, competition has become omnipresent in science, infiltrating institutions and even research teams. This increasing competition appears to dampen the openness of researchers, who tend to become secretive and defensive with their competitors. In other words, the omnipresence of competition now appears to provide legitimacy for secrecy and lack of transparency.

ACCEPTANCE, LOSS OF TRUST, OR WITHDRAWAL

We have shown that the place of the scientific community has changed considerably in the past 50 years. Many still believe that scientists should be devoid of self-interest, yet an increasing awareness of the reality of current scientific careers and their impact on the wellbeing of researchers raised concerns about this expectation. Young researchers seemed particularly discouraged from devoting themselves blindly to a career that promises very little in return. Research assessments also played an important role in remodeling the role of communalism and disinterestedness in science. The expectation of exceptional findings discourages researchers from being realistic in reporting their findings, while the focus on individuals and the reliance on competition dissuade them from collaborating and from being open. It appears that science is changing and with it, the commitment to the norms of communalism and disinterestedness. For our last point, we explore the reactions of scientists to these changes and to the growing ambivalence between the desire to work for the advancement of science and the need to secure one's career.

To begin, many respondents reported that scientists have to 'play with the rules of the game'. In other words, respondents agreed that communalism and disinterestedness are important, but they also implied that to survive the system, scientists need to adopt strategies or behaviours that deviate from ideal, selfless science. Unlike 'good science', 'strategic science' can inflate success regardless of the quality, value, or true worth of the science advanced. Some examples of the strategies mentioned include building a strong network, selectively reporting one's results to inflate their value, lobbying funders, nurturing a strong and active self-image on social media, etc. While some interviewees qualified these strategies as 'questionable', others considered simple strategies to be acceptable or even essential for success.

> *"For example, it just a silly example, I don't know whether it matters, but… I carefully considered to which journal I would send my first paper. [Interviewee describes the strategy involving networking with editors and timing of submission]. And it got published. […] So it's also that. It's also the networking. […] 'if you can't play by the rules you don't have to play the game'. Those are the rules. That's rules of the game. […] It's a game! It's a dance in a way. It's something you do, interact with people, it's not solely your message…" (Researcher who changed career)*

> *"I have now this 65-year-old sponsor who says 'Yeah, I have an idea, the paper is almost finished, maybe we can put you as last author, it will help your CV'. OK! Let's do it. I will change a comma (laughs.). So you're playing the game then as well right? He offered, so…" (PostDoc)*

> *"Whatever [research students] learn in university, is only the first step in one element in a much broader aspect of all that they need to do to be successful. So yes of course you need to be excellent. […] But next to that you need a lot of other things. […] You need to be able to play somehow the politics that are taking place, and the further you go along in your career, the more you have to do that. Don't come and try to convince me that the most successful researchers are not the ones who can lobby the funding agencies, who can create great relationships with editors, for example… I mean, <to> know managing editor of a journal, to actually be in a good position to get the funding, to get the paper, etc. It's part of the game." (Editor or publisher)*

Conversely, other interviewees disagreed with, or despised, strategic science, much like researchers in the time of Merton and Hagstrom despised the desire for praise and reward. Among them, we noticed two different ways of expressing this disagreement. On the one hand, researchers and research students who disagreed sometimes expressed a feeling of helplessness with the current system (see more on this discussion in Chapter 4). These researchers wished the system was different, but since strategic science was their only chance of survival, they went along with it.

> *"I don't really like the idea that only the publication rate will get you far in research career life. But I think at the moment this is how it is." (PhD student)*

> *"Everyone is sticking to the system because the system is like it is." (Researcher)*

On the other hand, others expressed stronger disagreement with the current dynamics of science and refused to endorse strategic science. This refusal led some to lose trust in science entirely. Indeed, some interviewees said that they doubted the reliability of scientific results and colleagues, and even believed academia was a futile enterprise. In many cases, this perspective was expressed by those who left academia for a different career, and was cited as a reason for their departure.

> *"In research what I find the biggest frustration, what I think, it's that sometimes, I got… I feel to lose trust in it. Because if I know, if I realise that research <is> only publishing the positive results and all those things... if it happens, in all the fields, also in the life threatening fields[…] then I'm sometimes afraid really, just in daily life to take medication because then I think 'Is it also based on research, that kind of research that we are doing in academia?' It's very hard to say. It's not that I don't believe in our kind of research, but because if it's so <often> based on research that is positive, results that are published, then I'm a little bit afraid sometimes of things that have been done to myself or to my kids or whatever. Because it's based on so many assumptions…" (Researcher)*

> *"But I think the real problem in science, the problem in my experience is that we do not trust each other totally. We are apparently not capable of saying 'well this research happened, it probably is true so we can elaborate further on it. It isn't true. A lot of research has shown that certain publications are not reproducible, or that certain findings are not published." (Researcher who changed career)*

> *"And also in the end I even started doubting 'What the fuck am I doing research for?', 'What's the difference?' All the money that we are putting in it, all the rats that we are killing doing it, all the hours you spend counting and statisticising <sic* [in the sense of analysing data]>*, and writing… why? What's the use? I mean let me teach at least I know I am training the doctor of tomorrow. I'm not saying research is useless, but research at the level where I was doing it, I don't know if it has a big use. For an article in a paper that no one ever reads, let's be honest, the other five idiots that are trying to do the same thing but… I am sounding very bitter now I realize, but…" (Researcher who changed career)*

The conflict between the idealistic vision of research and the realities of research careers raised different reactions among our informants. Some accepted and openly acknowledged that scientists need — and should be taught — strategic skills to survive in the system. Others disliked the need for strategies but believed that there was no alternative to survive in the current system. Finally, others lost their trust in science altogether, started doubting the methods and results of colleagues, or even questioned the purpose of science itself. Among these last participants, some were unable to reconcile the conflict between the type of scientist they wished to be with the scientist they had to be to survive and succeed. For these, leaving academia was the only way to stay true to their personal integrity.

In the early 2000's, Anderson and her colleagues assessed American researchers' subscription to Mertonian norms and compared such self-declared normative subscription to ratings of their behaviours and the behaviours of their colleagues (Anderson, Martinson, & De Vries, 2007). They found that scientists tend to rate their subscription to norms significantly higher than their own behaviours, expressing a normative dissonance which resonates with our findings. In addition, they found that those who most strongly believed in the norms were most likely to identify their colleagues' behaviour as counternormative. Pairing our findings with this important conclusion, it is possible that those who subscribe and believe most strongly in the norms of science are most disappointed in the behaviours of their colleagues and in what the system asks of them. Although these researchers may also need to adopt strategic science and counternorms to survive, they risk being dissatisfied with the current system. This disappointment may explain why some of our interviewees who were vocal about the importance of the norms of science lost trust in science or decided to leave academia. Yet, by draining out these researchers, the current research system supports the position that individual visibility and prestige prevail over adherence to the norms of communality and selfless collaboration. The current system supports the idea that researchers should essentially focus on building a strong Curriculum Vitae, assembling a strategic network, and ticking boxes of required deliverables instead of focusing on the fundamental objective of advancing science together.

CONCLUSIONS

Nearly eighty years ago, Merton introduced terms to describe the norms that appeared to govern the life of scientists. Beyond putting the words on concepts that have been implicitly guiding scientists for decades, the norms also re-established the importance of the sense of community between scientists. The norms of communalism and disinterestedness, in particular, imply that the knowledge that scientists produce belongs to the community, and that scientists are interested in advancing science and not in personal benefit. But science has evolved tremendously in the past century, and new demands and environments challenge the practicality of these norms. In the face of new research environments, the norms introduced by Merton have been questioned by several authors (see a discussion on the criticism towards the norms in Anderson et al., 2010) who doubt that the norms are still realistic in the face of new demands, heightened bureaucracy, and increasingly applied research fields (Ziman, 1999). A number of additional norms have been proposed, but the original four norms of Merton — often referred to as the CUDOs — are still among the most frequently used in discussions of research integrity.

Using recent interviews and focus groups with diverse research actors in Flanders, Belgium, we show that the norms of disinterestedness and communalism remain important in the current vision of science. Discussions around the importance of devotion and passion, but also around the crucial nature of collaborations, openness, and collegiality provide evidence of the key place that the two Mertonian norms have in shaping 21st century views of academia. Much like Hagstrom and Merton described decades ago, interviewees also condemned the counternorms of self-interest and secrecy. Indeed, interviewees shamed or stigmatized the desire for fame and recognition, with some even asserting that 'real' researchers should be willing to give up a healthy work-life balance, and should remain geographically mobile to pursue their passion. Some interviewees, however, disagreed with this last perspective. They claimed that today's academic realities expected scientists to sacrifice too much while offering too little in return and believed that early career scientists should be allowed to find a healthier work-life balance. This perspective has found a place in the media and scientific literature of the past few years. Research on the wellbeing of early career researcher has shown that young researchers face a disproportionately high risk of mental health issues, stress and depression (Levecque et al., 2017; Pain, 2017; Woolston, 2017). Corroborating what we learned in our interviews, many link this increased vulnerability to precarious employment driven by to the scarcity of permanent positions (Anonymous, 2010; Debacker & Vandevelde, 2016; "Many junior scientists," 2017) and to the strong competition and pressures imposed on early career researchers (Anderson, Ronning, et al., 2007; Anonymous Academic, 2018a; Maher & Sureda Anfres, 2016; Powell, 2016). Both of these issues share a common factor: they depend on the way in which researchers are assessed and promoted.

Current strategies for evaluating researchers are problematic. First, current research assessments connect excellence with positive and exceptional findings, thereby ignoring the importance of negative or inconclusive findings and small-step-research. Several scholars have expressed similar concerns, pointing out that publication biases decrease the reliability of published results and waste research resources (some well-known examples include Chalmers & Glasziou, 2009; Dwan et al., 2008; Fanelli, 2012; Ioannidis, 2005), with some even claiming that selective reporting is among the most damaging issue of current science (Bouter, Tijdink, Axelsen, Martinson, & ter Riet, 2016). Research assessments also tend to evaluate researchers individually, or at least to be translated in individual success requirements. In doing so, assessments risk undermining the value of collaboration and teamwork, as well as reducing the role of scientists to a one-size-fits-all. The problem of individual evaluations has been decried in the past (e.g., Bothwell, 2019; Mishra, 2015) and some initiatives such as CRediT or collaborative evaluation proposals (Mazumdar et al., 2015) have been put in place to increase the visibility of non-leading team members. Nonetheless, individual performance requirements remain the standard and feed into the related issue of hyper-competition. Indeed, the third problem with current methods for evaluating researchers, described above, is the omnipresence of fierce competition between scientists. Our interviewees described the complex role of competition in science since, given that it is thought to be both a stimulant and a risk to scientific advancement. In the literature however, the negative effects of competition on the integrity of science and on the collegiality of scientists are more apparent than its positive effects (see for e.g., Anderson, Ronning, et al., 2007; Fanelli, 2010; Fang & Casadevall, 2015; Tijdink et al., 2016).

Our interviews also suggest that individually focused and competitive research climates impact the way in which research is performed. In today's academia, researchers who ignore the importance of competition or who refuse to use strategies to succeed and advance their career place their career in serious jeopardy. Assessment tools thus appear to justify strategic and self-interested science, at the expense of communalism and disinterestedness.

In sum, the sense of community remains important to the ideology of the 'good researcher', but such ideologies struggle to survive in a reality where modest, collective, and collaborative science are not encouraged. Since we are currently witnessing an era of change, we should question the long-term impacts of research assessments that value extraordinary, individual, and competitive science.

Years ago, long before the onset of our project, a call for action had raised similar concerns, proposing that the overproduction of PhD graduates was a driver of many of academia's current problems (Martinson, 2011). While there have been some efforts to address this problem, the situation has not changed much in the past 10 years and only 3 to 20% of PhD students are able to continue in academia (Anonymous, 2010; "Many junior scientists," 2017), even though some estimates show that almost 60% aspire to it (Sauermann & Roach, 2012). From our interviews, we understood that Flanders, one of the regions with the highest ratio of PhD graduates to academic positions, still falls short on what the OECD (Organisation for Economic Co-operation and Development) recommends in numbers of PhD graduates. We also understood that having numerous PhD graduates is determinant for structural funding of Flemish institutions. This overproduction of PhD graduates however, fosters an hypercompetitive and precarious environment in which young researchers' wellbeing is put at risk. Young scientists must build strong research skills, but also need to learn the strategies that will ensure that they can get ahead in their career. While strategic science and integrity are not necessarily opposed, we found that some norms expressed in strategic science (e.g., self-interest, secrecy, interestedness, etc.) are in direct conflict with the largely agreed upon norms of disinterestedness and communalism. The uneasiness about this conflict, felt by some of our interviewees, is reassuring, affirming the ideological place of communalism and disinterestedness in current academia. Yet, the growing acceptance of this conflict suggests that we might be one of the last generations of scientists to feel the ambivalence between 'what we believe we should do as scientists', and 'what we must do in order to survive as scientists. Although it is outside of the scope of this chapter to address how communalism and disinterestedness can be fostered in the current academic climate, our results suggest that a profound revision of the way we evaluate researchers is necessary to restore — or at least maintain — the academic spirit that Merton observed nearly 80 years ago.

REFERENCES

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Anderson, M. S., Martinson, B. C., & De Vries, R. (2007). Normative dissonance in science: Results from a national survey of US scientists. *Journal of Empirical Research on Human Research Ethics, 2*(4), 3-14. doi:10.1525/jerhre.2007.2.4.3

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics, 13*(4), 437-461. doi:10.1007/s11948-007-9042-5

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2010). Extending the Mertonian Norms: Scientists' Subscription to Norms of Research. *Journal of Higher Education, 81*(3), 366-393. doi:10.1353/jhe.0.0095

Anonymous. (2010, December 18). The disposable academic. *The Economist*. Retrieved from https://www.economist.com/christmas-specials/2010/12/16/the-disposable-academic.

Anonymous Academic. (2018a, 22 June). I just got a permanent academic job – but I'm not celebrating. *The Guardian*. Retrieved from https://www.theguardian.com/higher-education-network/2018/jun/22/permanent-academic-job-university-system-unfair-exploitative.

Anonymous Academic. (2018b, 15 June). Is competition driving innovation or damaging scientific research? *The Guardian*. Retrieved from https://www.theguardian.com/higher-education-network/2018/jun/15/is-competition-driving-innovation-or-damaging-scientific-research.

Bothwell, E. (2019, 14 October). Award Nobels to teams, not individual 'heroes', say scientists. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/news/award-nobels-teams-not-individual-heroes-say-scientists.

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review, 1*(1), 17. doi:10.1186/s41073-016-0024-5

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet, 374*(9683), 86-89. doi:10.1016/s0140-6736(09)60329-9

Debacker, N., & Vandevelde, K. (2016). From PhD to professor in Flanders. *ECOOM Brief (no. 11)*. Retrieved from https://biblio.ugent.be/publication/8043010.

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A. W., Cronin, E., . . . Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE, 3*(8). doi:10.1371/journal.pone.0003081

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*, 282. doi:10.1038/nbt.4089

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE, 5*(4), e10271. doi:10.1371/journal.pone.0010271

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90*(3), 891-904. doi:10.1007/s11192-011-0494-7

Fang, F. C., & Casadevall, A. (2015). Competitive Science: Is Competition Ruining Science? *Infection and Immunity, 83*(4), 1229-1233. doi:10.1128/iai.02939-14

Gingras, Y. (2016). *Bibliometrics and Research Evaluation: Use and Abuses*. Cambridge (MA): MIT Press.

Hagstrom, W. O. (1975). *The scientific community*. Carbondale (IL): Southern Illinois University Press.

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature News, 520*, 429-431. doi:10.1038/520429a

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), 696-201. doi:10.1371/journal.pmed.0020124

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868-879. doi:10.1016/j.respol.2017.02.008

Maher, B., & Sureda Anfres, M. (2016). Young scientists under pressure: what the data show. *Nature, 538*, 444-445. doi:10.1038/538444a

Many junior scientists need to take a hard look at their job prospects. (2017). *Nature, 550*, 429. doi:doi:10.1038/550429a

Martinson, B. C. (2011). The academic birth rate. Production and reproduction of the research work force, and its effect on innovation and research misconduct. *EMBO Reports, 12*(8), 758-761. doi:10.1038/embor.2011.142

Mazumdar, M., Messinger, S., Finkelstein, D. M., Goldberg, J. D., Lindsell, C. J., Morton, S. C., . . . Translational Science Awards, C. (2015). Evaluating Academic Scientists Collaborating in Team-Based Research: A Proposed Framework. *Academic medicine : journal of the Association of American Medical Colleges, 90*(10), 1302-1308. doi:10.1097/ACM.0000000000000759

The mental health of PhD researchers demands urgent attention. (2019). *Nature, 575*, 257-258. doi:10.1038/d41586-019-03489-1

Merton, R. K. (1942). The Normative Structure of Science. In N. W. Storer (Ed.), *The Sociology of science : theoretical and empirical investigations (1973)*. Chicago (IL): University of Chicago Press.

Merton, R. K. (1957). Priorities in Scientific Discovery. In N. W. Storer (Ed.), *The Sociology of science : theoretical and empirical investigations (1973)*. Chicago (IL): University of Chicago Press.

Mishra, J. (2015, 9 November). Choosing team over tenure. *Science*. Retrieved from https://www.sciencemag.org/careers/2015/11/choosing-team-over-tenure. doi:10.1126/science.caredit.a1500258

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2019). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *OSF Preprints*. doi:10.31219/osf.io/m9abx

Moore, S., Neylon, C., Paul Eve, M., Paul O'Donnell, D., & Pattinson, D. (2017). "Excellence R Us": university research and the fetishisation of excellence. *Palgrave Communications, 3*, 16105. doi:10.1057/palcomms.2016.105

National Science Board. (2018). *Chapter 3 - Science and Engineering Labor Force*. National Science Foundation. [Digest (NSB-2018-2)]. Retrieved from https://nsf.gov/statistics/2018/nsb20181/report/sections/science-and-engineering-labor-force/global-s-e-labor-force

OECD. (2019). *Main Science and Technology Indicators*. [MSTI 2019/1]. Retrieved from http://oe.cd/msti

Pain, E. (2017, April 4th). Ph.D. students face significant mental health challenges. *Science*. Retrieved from https://www.sciencemag.org/careers/2017/04/phd-students-face-significant-mental-health-challenges. doi:10.1126/science.caredit.a1700028

Powell, K. (2016). Young, talented and fed-up: scientists tell their stories. *Nature, 538*, 446-449. doi:10.1038/538446a

Sauermann, H., & Roach, M. (2012). Science PhD Career Preferences: Levels, Changes, and Advisor Encouragement. *PLoS ONE, 7*(5), e36307. doi:10.1371/journal.pone.0036307

Tijdink, J. K., Schipper, K., Bouter, L. M., Maclaine Pont, P., de Jonge, J., & Smulders, Y. M. (2016). How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers. *BMJ Open, 6*(2), e008681. doi:10.1136/bmjopen-2015-008681

UNESCO Institute for Statistics. *Science, technology and innovation*. Retrieved from http://data.uis.unesco.org

Wilsdon, J., Liz Allen, Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from https://re.ukri.org/documents/hefce-documents/metric-tide-2015-pdf/

Woolston, C. (2017). Graduate survey: A love–hurt relationship. *Nature, 550*(7677), 549-552. doi:10.1038/nj7677-549a

Woolston, C. (2019). PhDs: the tortuous truth. *Nature, 575*, 403-407. doi:10.1038/d41586-019-03459-7

Ziman, J. (1999). Rules of the game of doing science. *Nature, 400*(6746), 721-721. doi:10.1038/23379

Chapter 6

Advancing science or advancing careers? Results from a survey with researchers

An adaptation of this chapter has been submitted to the journal *PLOS ONE* and is available as a preprint at biorXiv at https://doi.org/10.1101/2020.06.22.165654.

CONTRIBUTIONS

Conceptualization: Noémie Aubert Bonn, Wim Pinxten[1]

Funding acquisition: Wim Pinxten. Funding granted by the Bijzonder Onderzoeksfonds (BOF) 15NI05

Project administration: Noémie Aubert Bonn, Wim Pinxten

Methodology: Noémie Aubert Bonn and Wim Pinxten built the survey; Vincent Larivière[2] provided guidance on the survey content, terminology, and recruitment methods; Patricia Tielens[3] helped us understand how to distribute the survey in respect with GDPR; Raymond De Vries[4], Søren Holm[5], and Daniele Fanelli[6] provided feedback on earlier versions of the survey; Dana Hawwash[7], Paolo Corsico[5], and Audrey Wolff[8] provided feedback on final versions of the survey.

Resources: Many people helped share the survey, thereby helping us find participants. These include Deans, Directors of doctoral school, and secretaries of universities we contacted, especially those of UHasselt and UAntwerpen who shared the survey internally, Raffaella Ravinetto[9], Hannelore Storms[1], Carl Lachat[7], Stefanie Van der Burght[10], and many more people who shared or re-tweeted our survey. We also wish to thank the participants themselves for their time, efforts, and willingness to share their thoughts.

Investigation: Noémie Aubert Bonn

Data curation: Noémie Aubert Bonn

Formal analysis: Geert Molenberghs[11] provided guidance and continued support on the appropriate statistical analyses to use; Noémie Aubert Bonn performed the analysis.

Visualization: Noémie Aubert Bonn

Validation: Noémie Aubert Bonn, Wim Pinxten

Supervision: Wim Pinxten

Writing – original draft: Noémie Aubert Bonn

Writing – review & editing: Noémie Aubert Bonn, Wim Pinxten

1. Department of Healthcare and Ethics, Hasselt University, Hasselt, Belgium
2. École de bibliothéconomie et des sciences de l'information, Université de Montréal, Montréal, Canada
3. Jurist Data Protection Officer, Hasselt University, Hasselt, Belgium
4. Center for Bioethics and Social Sciences in Medicine, University of Michigan Medical School, Ann Arbor (MI), USA
5. Centre for Social Ethics and Policy, University of Manchester, Manchester, UK
6. Department of Methodology, London School of Economics and Political Sciences, London, UK
7. Department of Food technology, Safety and Health, Universiteit Gent, Ghent, Belgium
8. GIGA Consciousness - Coma Science Group, Université de Liège, Liège, Belgium
9. Public Health Department, Institute of Tropical Medicine Antwerp, Antwerp, Belgium
10. Research Co-ordination Office, Universiteit Gent, Ghent, Belgum
11. L-BioStat, KU Leuven, Leuven, and Data Science Institute, Hasselt University, Belgium

ABSTRACT

The way in which we assess researchers has been under the radar in the past few years. Critics argue that current research assessments focus on productivity and that they increase unhealthy pressures on scientists. Yet, the precise ways in which assessments should change is still open for debate. We circulated a survey with Flemish researchers to understand how they work and how they would rate the relevance of specific indicators used in research assessments. We found that most researchers worked far beyond their expected working schedule. We also found that, although they spent most of their time doing research, respondents wished they could dedicate more time to it and less time to other activities such as administrative duties and meetings. When looking at success indicators, we found that indicators related to openness, transparency, quality, and innovation were perceived as highly important in advancing science, but as relatively overlooked in career advancement. Conversely, indicators which denote of prestige and competition were generally rated as important to career advancement, but irrelevant or even detrimental in advancing science. Open comments from respondents further revealed that, although indicators which indicate openness, transparency, and quality (e.g., publishing open access, publishing negative findings, sharing data, etc.) should ultimately be valued more in research assessments, the resources and support currently in place were insufficient to allow researchers to endorse such practices. In other words, current research assessments are inadequate and ignore practices which are essential in contributing to the advancement of science. Yet, before we change the way in which researchers are being assessed, supporting infrastructures must be put in place to ensure that researchers are able to commit to the activities that may benefit the advancement of science.

INTRODUCTION

The way we define and evaluate scientific success impacts the way in which research is performed (Butler, 2003). Yet, definitions of success in science are ambiguous and have raised many debates in the past few years. The San Francisco Declaration on Research Assessments (American Society for Cell Biology, 2013), the Leiden Manifesto (Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015), or the Metric Tide (Wilsdon et al., 2015) are all examples that denounced the inadequacy of the metrics currently used for research assessments. Most critics argue that current metrics are reductionistic and inappropriate for individual evaluations. But despite increasing criticism, alternative assessments are difficult to find (e.g., Tregoning, 2018). Some argue that narratives and subjectivity must be reintroduced in research assessments, while other support the need for new metrics to broaden the scope of evaluations. As the approaches that need to be taken are still disputed, a few institutions and funding agencies have taken the lead in exploring new ways to evaluate researchers (see for e.g., DORA, 2020; *Ghent University is changing course*, 7 December 2018; "VSNU, NWO, NFU and ZonMw," 2018).

Nevertheless, research assessments do not only depend on institutions and funding agencies. Indeed, even when assessments are organised by institutions and agencies, researchers are often the ones who act as referee. Consequently, changing research assessments does not only require new regulations, guidance, and policies from institutions, but also a cultural reform on the research floor.

Last year, we explored definitions of success in biomedical research in Flanders, Belgium (Chapter 3). Using interviews and focus groups, we captured the perspectives of different research actors on what determines success in science. Actors included research institutions, research funders, scientific editors or publishers, researchers, research students, and several other actors who play an important role in academic research. Although interviewees largely agreed on what currently constitutes success in science, their opinions also conflicted in some aspects. Oftentimes, the same indicator of success would raise opposite reactions, with some proposing that one indicator was paramount to advancing science and others arguing that the same indicator threatened research integrity.

Using the indicators of success which were mentioned in the interviews and focus groups, we built a brief survey to better understand the importance of each indicator in advancing science, advancing one's career, and in yielding personal satisfaction.

METHODS

TOOL

We built the survey from themes extracted in past interviews and focus groups (Chapter 3) using Qualtrics XM. More specifically, we built the survey questions on the indicators which were thought to play a role in acquiring success in science, but which also raised conflicting opinions. The resulting survey assesses the impact of each indicator on career advancement, scientific advancement, and personal satisfaction. We refined the survey questions and statements several times by consulting a few experts and colleagues as well as researchers with experience in similar questionnaires. We tested the survey with close colleagues (4 PhD students, and 4 senior researchers) to ensure clarity and relevance of the questions. The final version of the survey includes 18 statements whose impact is assessed on each of the three pillars. Table 1 showcases the final statements which figured in the survey alongside the available answer options.

Beyond the 18 statements included in the survey, we asked a few demographic questions, such as gender, university and faculty of affiliation, current position and seniority, and number of publication. We also included a series of questions on time management which asked the average number of hours worked each week; the percentage of time dedicated to research, teaching, and 'other' tasks; and in greater detail, the percentage of time dedicated to direct student supervision, hands on research work, staying up to date, writing papers, reviewing, grants writing, and other tasks. The full printout of the survey is available in Appendix 10.

Table 1. Statements included in the survey and answer options

Publishing papers is…	□ essential	
Publishing in high impact journals is…	□ important	…in advancing my career
Publishing commentaries or editorials is…	□ irrelevant	
Publishing more papers than others is…	□ unfavorable	
Publishing open access is…	□ detrimental	
Peer reviewing is…		
Replicating past research is…	□ essential	
Publishing findings that did not work (i.e., negative findings) is…	□ important	…in advancing science
Sharing your full data and detailed methods is…	□ irrelevant	
Reviewing raw data from students and collaborators is…	□ unfavorable	
Conducting innovative research with a high risks of failure is…	□ detrimental	
Connecting with renowned researchers is…		
Collaborating across borders, disciplines, and sectors is…	□ essential	
Getting cited in scientific literature is…	□ important	…to my personal satisfaction
Having your papers read and downloaded is…	□ irrelevant	
Having public outreach (e.g., social media, news, etc.) is…	□ unfavorable	
Having your results used or implemented in practice is…	□ detrimental	
Having luck is…		

Recruitment

To enable broad recruitment in the Flemish academic research landscape, we initially considered sharing the survey using Web of Science corresponding emails from Flemish authors available in the database. Nevertheless, after consulting the Data Protection Officer at our institution, we found that this recruitment method might not be fully compliant with the new European General Data Protection Regulation (GDPR; i.e., the purpose for sharing one's email address when publishing as first author does not imply an agreement to receive invitations to research surveys). Consequently, in respect of GDPR, we downsized our prospective sample and directly contacted the faculties of medicine and life science of Flemish universities or equivalent to ask them to circulate our survey within their faculty.

We contacted Deans and Directors of doctoral schools from all five Flemish universities, namely Universiteit Antwerpen (UAntwerpen), Universiteit Gent (UGent), Universiteit Hasselt (UHasselt), Katholieke Universiteit Leuven (KU Leuven), and Vrije Universiteit Brussels (VUB). We further reached out to the Institute of Tropical Medicine Antwerp (ITM) and to the Interuniversity Microelectronics Centre (IMEC) using contacts we knew from within the institutes. Two universities distributed the survey with the entire faculty via email (UHasselt and UAntwerpen) and ITM agreed to distribute the survey internally to its researchers. One university preferred not to share the survey within its institution (KU Leuven) and our emails to one university remained unanswered (VUB). One university agreed to distribute the survey, be it not by distribution via a mailing list but by social media invitation to our survey (UGent). Given the latter, we composed an invitation which was shared by UHasselt's social media accounts (Twitter and Linked'in), and later shared by UGent's social media accounts. By sharing the survey publicly, we allowed anyone interested to participate, whether they were affiliated with a Flemish institution or not. We encouraged any re-tweets, likes, and shares to promote participation via snowballing strategies. Select contacts at IMEC and UGent also shared the survey within colleague groups. Finally, we further transferred the survey to select research groups and mailing lists with whom we are acquainted in Flanders. As a result, our participant group is diverse and spread out but also uneven, with numerous responses from institutions who shared the survey internally and few responses from institutions who relied on social media or snowballing. The survey was open from the 8th to the 31st October 2019.

The project was approved by the committee for medical ethics (Comité voor Medische Ethiek) of the Faculty of Medicine and Life Science of Hasselt University, protocol number cME2019/O3s.

DATA ANALYSIS

In order to maximally exploit our data, we conducted statistical analyses that may help compare different categories and dimensions of answers. Considering the exploratory nature of our study as well as the possible biases and lack of generalizability of our sample however, we encourage readers to interpret the statistical results conscientiously. We join measures of central tendency and statistical results are available in Appendices.

In the time allocation questions, we used paired t tests to compare the actual time spent on different research activities to the ideal time respondents would like to spend on such activities. For these analyses, we only included researchers who declared working full time as researchers. Since many of the responses on time distribution were not normally distributed, we transformed all time values to logarithmic values, adding 0.5% to all values to avoid having to deal with zeros.

For the dimensions of success indicators, we captured the views of participants using Likert scales with five options to rate the importance that each success indicator had on 'advancing science', 'advancing careers', and 'contributing to personal satisfaction' (table 1). Before analysing our data statistically, we transformed the answer options to numerical values (i.e., detrimental = 1, unfavorable = 2, irrelevant = 3, important = 4, essential = 5). The possibility of interpreting Likert Scale data as continuous data raises controversies, and our findings should be interpreted with caution. Indeed, since Likert scales provide no guarantee that the distance between each category is equal, many propose that they should be treated as ordinal data (Allen & Seaman, 2007; Mangiafico, 2016; Sullivan & Artino, 2013). We decided to use repeated measures ANOVAs with Bonferroni post hoc test to compare the average ratings of each statement's importance in i) advancing one's career, ii) advancing science, and iii) advancing one's personal satisfaction, but to also provide thorough visual depictions of our data (made using Tableau Desktop 2018.1 and Excel) and access to full data files to allow re-analysis and re-interpretation of our findings. Our survey data are available as in the Thesis Online Material folder that is on our Open Science Framework registration https://osf.io/ap4kn/.

RESULTS

DATA AVAILABILITY

The datasets are available online as 'Chapter 6 – 2. Survey data' in our Open Science Framework registration https://osf.io/ap4kn/. To ensure confidentiality and to avoid inter-university comparisons (which we believe would provide no useful information at this point), we extracted the detailed affiliation from the dataset, leaving only the information stating whether the respondent was affiliated with a Flemish institution or not.

PARTICIPANTS

In total, 126 participants completed the survey, two-thirds of which were either PhD students or post-doctoral / non-tenure-track researchers (Table 2). The gender distribution was well balanced (64 females, 60 males, and 2 prefer not to disclose). Almost 90% of participants were affiliated with Flemish institutions (n=112), with the Universiteit Antwerpen, Universiteit Hasselt, and the Institute of Tropical Medicine Antwerp being the three most represented (Table 2). A few international participants also contributed to the survey. Most participants had below ten published papers and around three quarter of respondents had below 30 publications. Yet, the distribution of publication profiles was broad and ranged until the maximum option available in the survey of 'over 210 published papers'.

Table 2. General demographics

POSITION	
PhD Student	48
PostDoc / Non-tenure-track researcher	36
Tenure-track researcher / Professor	13
Tenured researcher / Full professor	13
Emeritus Professor	3
Former researcher	6
Other	7
GENDER	
Female	64
Male	60
Prefer not to say	2
AFFILIATION	
Affiliated with a Flemish institution	113
UAntwerpen	31
UHasselt	28
ITM	18
KU Leuven	11
UGent	9
IMEC	9
VUB	4
Other	3
Affiliated with an institution outside Flanders	13
PUBLICATION PROFILE	
<10 peer-reviewed papers	58
10–30 peer-reviewed papers	40
30–60 peer-reviewed papers	9
60–90 peer-reviewed papers	4
90–120 peer-reviewed papers	3
120–150 peer-reviewed papers	6
150–180 peer-reviewed papers	0
180–210 peer-reviewed papers	3
>210 peer-reviewed papers	3
TOTAL NUMBER OF PARTICIPANTS	126

TIME MANAGEMENT

Almost three quarter of the respondents (n=93) stated that they worked full time as a researcher or PhD student. On average, respondents who declared working full time worked 46.91 hours per week (median 46) but the distribution was very wide (Table 3). Among those who declared working full time as a researcher or a research student, 73 (78.5%) said that they worked more than 40 hours per week, from whom 41 (44.1%) declared that they worked more than the European Union directive maximum of 48 hours per week (Directive 2003/88/EC, 2003), and 11 (11.8%) declared that they worked 60 hours per week or more until a maximum of 80 hours (fixed maximum in the survey). When including respondents who were not declared as full time researcher, the number of respondents declaring to work 60 or more hours per week rose to 18.

Table 3.Hours worked per week for respondents who declared working full time

	N	Average time	Median	Min	Max
PhD student	40 (39)	46.30 (46.72)	45 (45)	30 (38)	60
Post-doctoral / Non-tenure track position	27	46.59	46	40	70
Tenure-track researcher / Professor	9 (8)	48.00 (54.00)	50 (55)	0 (42)	65
Tenured researcher / Full professor	8 (7)	51.13 (57.29)	54.50 (59)	8 (42)	80
Researcher in the past, but moved to another career	4	48.75	49	46	51
Other	5	43.40	45	36	51
	Total N 93 (90)	Overall average 46.91 (48.06)	Overall median 46 (46.5)	Overall min. 0 (36)	Overall max. 80

Note: Numbers in parentheses exclude three answers which were not above 30h per week (i.e., 0, 8, and 30). We deemed that these answers could be deleted since respondents confirmed being employed full time as researchers, and thus the answers may reflect a misunderstanding with our interpretation of full time research employment

In the next questions, we targeted more specific research activities to understand how researchers distribute their research time. A first question asked respondents to distribute their time between three main pillars, namely *Teaching*, *Research*, and *Other*. A second question targeted more specific activities, namely *direct student supervision*, *hands-on research work (e.g., lab work, data analysis)*, *staying up-to-date (e.g., reading, listening, building skills, etc.)*, *writing papers*, *reviewing*, *grant writing*, and *anything else (e.g., administration, meetings, etc.).* For both questions, we asked respondents to tell us the percentage of time they really spent on each activity, as well as the percentage of time they would like to attribute to each activity if they were in an ideal world.

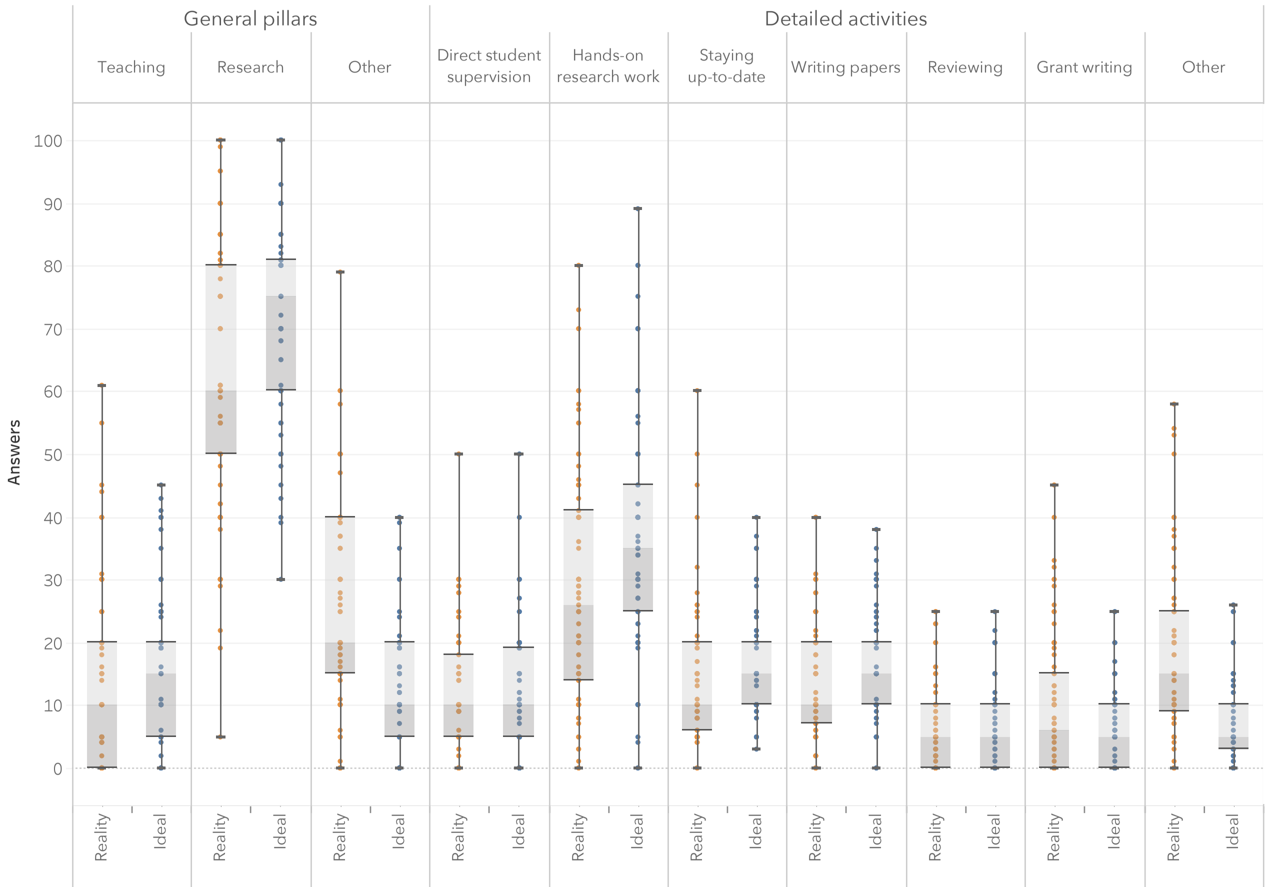
The range of answers was very broad, as could be expected for this type of question. Since some categories raised different responses between respondents who worked full time as researchers and those who did not (i.e., 'Research' and 'Other' from the main pillars, and 'Research work' and 'Anything else' from the detailed categories), we decided to keep only full-time researchers (n=93) for analyses of the questions on time distribution. Since many of the responses were not normally distributed, we transformed all values to logarithmic values, adding 0.5% to all values avoid zeros. All means, medians, and statistical results may be seen in Appendix 12.

From the general pillars, we found that participants wished they could dedicate more time to 'Teaching' and especially to 'Research', but less time to 'Other' activities. When looking at the detailed activities, we could see that participants generally hoped they could spend more time on *hands-on research work*, '*staying up to date'*, and *writing papers*. On the other hand, respondents wished to spend much less time doing '*anything else*'. The differences between real and ideal times for *'supervision', 'reviewing',* and *'grant writing'* were not significant. Fig 1 illustrates the distribution of answer for each activity.

IMPACT OF ACTIVITIES AND INDICATORS

In the last section of the survey, we asked participants to indicate the impact of 18 different research activities on i) advancing their career, ii) on advancing science, and iii) on their personal satisfaction (table 1). Since our data comes from Likert scales we must be careful in interpreting our findings statistically (see

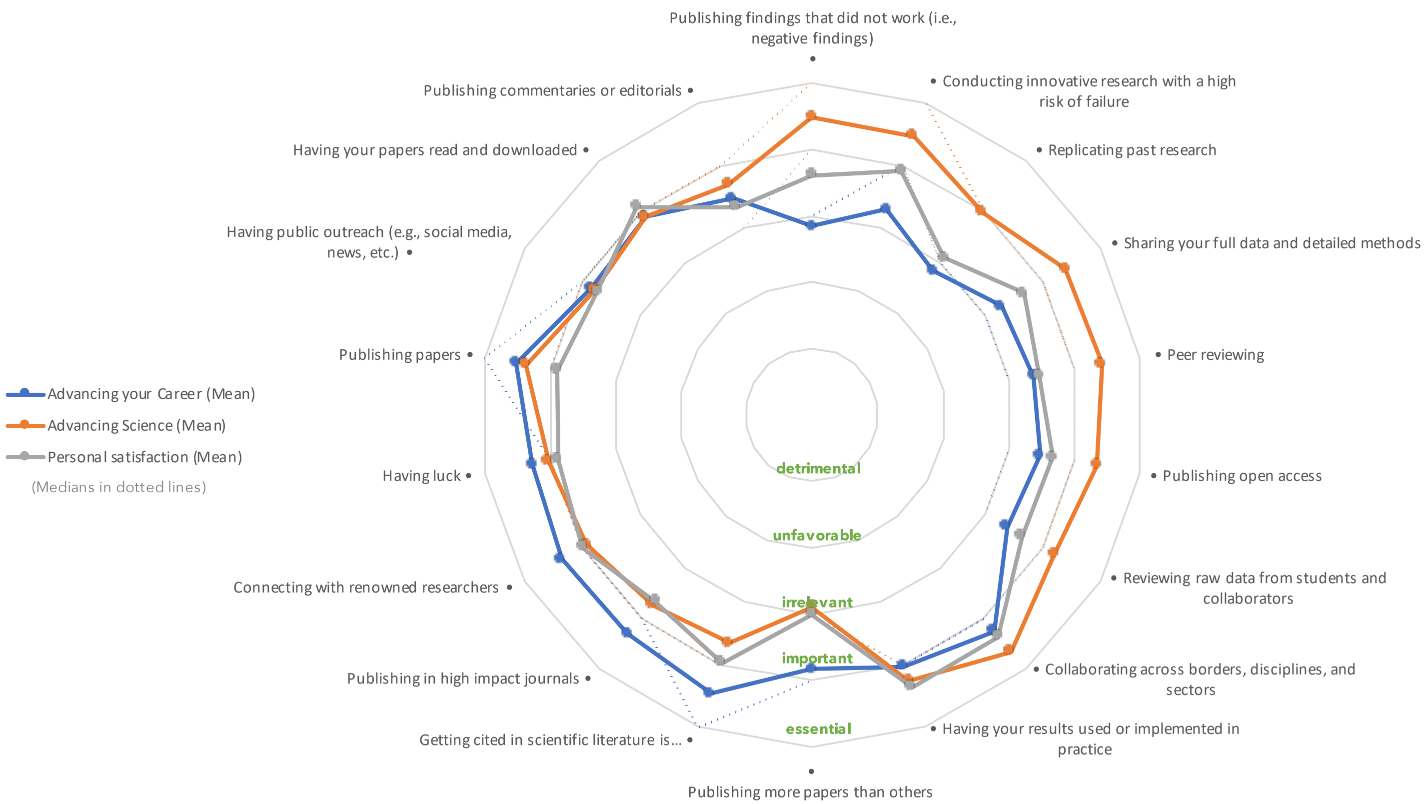
Figure 1. Self-reported percentage of time spent on different research activities in reality and in an ideal world for respondents who declared working full time (n=93).



more information about this in the methods section). Yet, we believed it worthy to exploit our data as much as possible and conducted repeated measures ANOVAs while carefully complementing our analyses with thorough visual representations of the answers gathered so as to increase the comprehensibility of the data.

In figure 2, we show the mean and median scores gathered for each of the 18 statements. We organized statements to allow a quick visual inspection of statements which were rated as highly important in advancing science, but less relevant in advancing one's career (North-East quadrant), and statements which were considered essential in advancing one's career, but of lesser importance in advancing science (South-Western quadrant). Except for the statement "Having public outreach (e.g., social media, news, etc.)", all main effects were significant, meaning that at least two of the dimensions (advancing science, advancing careers, or personal satisfaction) differed from one another. Appendix 11 shows all means and statistical results. Bonferroni post hoc test revealed that, in most cases, activities were rated differently on their impact in advancing science than on their impact in advancing one's career. Only one statement yielded similar scores in advancing one's career and in advancing science, namely "Having your papers read and downloaded". "Publishing papers" was thought to be slightly more important in advancing one's career than in advancing science, but the difference was not so distinct (means of 4.52 and 4.37 respectively, 95% CI 0.001 – 0.301). Personal satisfaction, on the other hand, was at times closer to the impact on one's career, and at other times closer to the impact in advancing science. Specifically, "Peer reviewing" and "Collaborating across borders, disciplines, and sectors" were rated similarly on their importance to researchers' careers and personal satisfaction. While "Peer reviewing" was low on both career and personal satisfaction "Collaborating across borders, disciplines, and sectors" was thought to be highly important for both. "Publishing in high impact journals", "Publishing more papers than others", "Connecting with renowned researchers", and "Having luck" were rated as contributing more to advancing researchers' career than to either science and personal satisfaction. Finally, "Having results used or implemented in practice" was rated higher on both personal satisfaction and scientific advancement than on the impact it has on one's career.

Figure 2. Mean and median ratings on each dimension for each statement



Taking our statements on a second level, a few trends become visible. First, it appears that various activities meant to promote openness and transparency (i.e., "Publishing findings that did not work"; "Sharing your full data and detailed methods"; "Publishing open access"), quality assurance (i.e., "Replicating past research"; "Peer reviewing", and "Reviewing raw data from students and collaborators"), and innovation (i.e., "Conducting innovative research with a high risk of failure") were thought to be important to advancing science but significantly less important in advancing researchers' careers. Looking at the full range of answers in figure 3, we can see that several respondents classified some of these activities as unfavorable or even detrimental in advancing their career. On the other hand, a few statements were rated as more important in advancing researcher's career than in advancing science. Among those, statements which relate to competition (i.e., "publishing more papers than others") and prestige (i.e., "getting cited in the literature", "publishing in high impact journals", and "connecting with renowned researchers") were most notable. As a general rule, respondents appeared to have most satisfaction from "Collaborating across borders, disciplines, and sectors", and from 'Having their results implemented in practice'.

The survey also allowed participants to comment after each item. These comment boxes were rarely used, but the few answers collected provide richer insights about some of the elements included in the survey.

Publications and metrics. Comments added to the statements on the importance of "Publishing papers" proposed that publications were important for science, but were not necessarily used properly. Some comments mentioned that publications were a better indicator of the status and resources of a laboratory than they are of the '*actual research capabilities*' of researchers, while others stated that publishing should aim to share a message, not to increase metrics.

> *"It depends on what is in the paper. When we really have something to say, we should say it. On the other hand, publishing for the sake of publishing is very detrimental for science as well as to my personal satisfaction."*

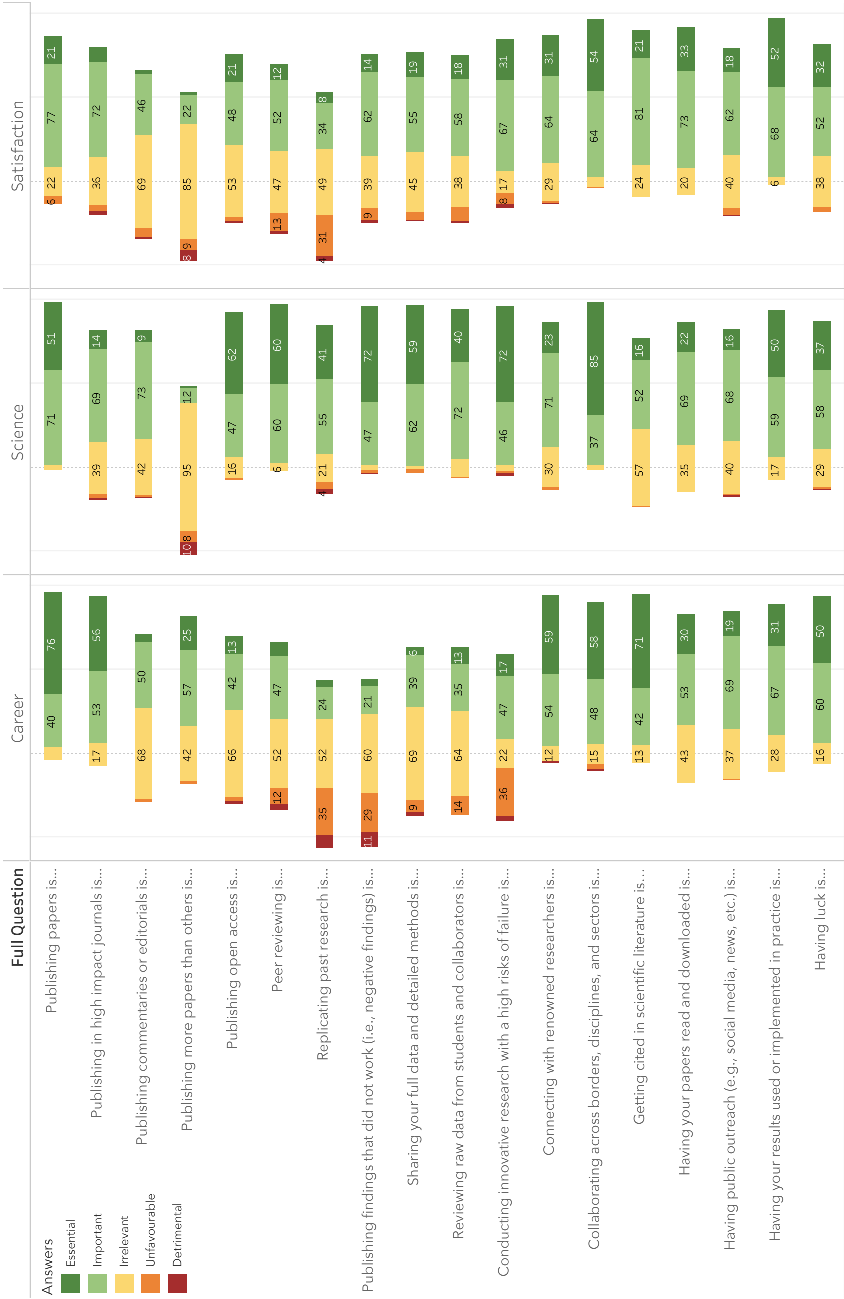


Figure 3. Distribution of answers for each dimension of each statement

When asked further about the impact of “Publishing in high impact journals”, respondents mentioned that expectations of high impact added pressures, but that publishing in high impact journals was also satisfying for researchers since high impact was perceived as a mark of quality.

> *"I really believe that it should be irrelevant, but if I am honest, I admit that it is somewhat important to my personal satisfaction. I am proud if I am able to publish a paper (for which I have really done my best) in a good quality journal."*

Others, on the other hand, worried that focusing on high impact journals prevented smaller or local journals from developing.

Openness and transparency. A few statements related to openness also provoked comments. “Publishing open access” raised a few controversial reactions. One respondent stated being generally ‘suspicious’ of open access journals for asking researchers to cover the publication costs, but most other comments rather mentioned that they would support open access but lacked the funds to do so. Comments on the importance of “Publishing findings that did not work (i.e., negative results)” explained that negative results must be more visible in published literature — one comment even stated that "*publication bias is the single most detrimental issue to modern science*". Yet respondents recognised that publishing negative results was rarely recognised. One respondent proposed that ‘sneaking’ negative results into papers with positive results could help counter publication biases. Another respondent, however stated that despite supporting the importance of publishing negative results, her precarious situation — accentuated by gender and seniority inequalities — made it difficult to withhold her convictions and to feel satisfied about publishing negative findings.

> *"Even though I have been publishing and speaking about my failures, I also find it difficult to do so, as I am in a precarious situation as a woman and early career researcher, which of course limits how much I feel personally satisfied..."*

Finally, the importance of “Sharing full data and methods” yielded diverging opinions. One respondent disclosed that, having experienced plagiarism in the past, he preferred to share data in personal discussions than to open it to the world. Others added that sharing raw data implied a lot of extra work and ethical issues which they were not ready to deal with. Finally, another three respondents raised doubts about the true benefits of sharing data for promoting integrity.

> *"I believe that it is a (naïve) illusion to think that putting full data online is the solution for research misconduct and sloppy research. I can think of (because I often see) many ways to ruin a dataset before it is shared online."*

Reviewing. From the open comments, we found that some respondents mistrusted the value of peer-review with some believing that it was governed by a principle of "*I scratch your back if you scratch mine*". Yet, given that most respondents rated "Peer-reviewing" as essential or important in advancing science, doubt on its value may not have been generalized. Indeed, other comments proclaimed that peer review was very important in advancing science in an unselfish way and that researchers "*who publish papers but refuse to invest time in reviewing rig the system*". Beyond the value of peer-review itself, others stated that although they appreciated peer-review, they felt exploited by big publishers when donating their time to it.

> *"I like staying up to date and be challenged in my thinking because of reviewing, but I don't like the fact that my free labour helps to increase profit for major publishers."*

The same respondent explained that peer review could be a satisfying experience if researchers were *"able to spend time on it during working hours (and not just on top of everything else)".* Along the same lines, responses on the importance of "Reviewing raw data from students and collaborators" also suggest that reviewing data was important for the quality of the work, but that it would take time which was necessary elsewhere, even if for the wrong reasons.

> *"Unfavorable in advancing my career because it takes time and my supervisors want me to spend more time in grant writing. This comment goes for all the activities that really advance science but are difficult to measure. It is the "doing the best I can" that is not acceptable to my supervisors. They want productivity (in the past this meant getting my name on many publications, nowadays it increasingly means getting grant money)."*

LIMITATIONS

Our findings are preliminary and contain important limitations which must be considered when interpreting the results.

The first limitation concerns our recruitment strategy, which was changed and challenged many times to respect the new General Data Protection Regulation. Initially aiming for a controlled recruitment through university contacts and direct emails, we expected to be able to calculate response rates and to have control over the profile of participants who would respond to our survey. Nevertheless, given our inability to reach all research institutions, we shared the survey online to allow for additional institutions to distribute it. Although most broad surveys are now shared using social media and snowballing, this choice inevitably influenced the pool of participants who had access to our survey, and may have increased participation bias. Sharing the survey online also removed the possibility of calculating a clear response rate. Adding to this first concern, our decision to focus largely on the Flemish region of Belgium may have led to answers that are not generalizable elsewhere. Our findings should therefore be interpreted with caution and in consideration of the limited number of principally Flemish participants.

Second, asking participants to estimate their working hours and time distribution relies on precision of recall and accuracy of self-report; two aspects we had no means to verify in the current work. Assessments of the reliability of self-reported working hours are largely absent from the literature. We only found one published paper to support the correlation between recorded and self-reported working hours, but it concerned Japanese workers highly aware of their working schedules (Imai et al., 2016). We cannot assume that similar findings would be observed in academic researchers whose working hours vary greatly and whose task concentration changes between academic year periods. Past works and popular surveys investigating the time allocation of scientists found different distribution of work allocation than those we found in our survey, generally reporting a higher proportion of time spent teaching (Matthews, 2018; Ziker, 2014). This difference, which is probably due to the high representation of PhD students among our participants, suggests that our findings might not be representative of other settings and populations and should be interpreted with caution. Despite this limitation, our findings coincide with other works in stating that researchers report working overtime (Barnett, Mewburn, & Schroter, 2019; Bothwell, 2018; Kinman & Wray, 2013; Koens & Jonge, 2018; Mckenna, 2018; Powell, 2016; Ziker, 2014), that they are subject to heavy administrative burden (Koens & Jonge, 2018; Schneider, Ness, Shaver, & Brutkiewicz, 2014), and that they wish they could dedicate more time to research (Mergaert & Raeymaekers, 2017).

Another important limitation comes from the formulation of the questions within the survey itself. In our attempt to make a manageable and coherent survey, we preferred to formulate simple questions than to try to describe the full complexity of the terms used. As a result, respondents may have interpreted concepts differently depending on their experience and personal views. We consider that this rich and diverse interpretation of terms and concepts, however, is closer to the reality of research assessments than one in which clear and precise definitions are provided (i.e., evaluation committees rarely have clear definitions of the concepts of ‘innovative’ or ‘excellent’ and are generally left to their own interpretations). Nevertheless, we concede that different interpretation of terms may have influenced responses.

Finally, a respondent suggested that "*the categories were not refined enough*" and that we should have included more options or a numerical scale to allow for some nuance. Indeed, our initial idea was to use a slider scale, but given the poor and flimsy rendering of this option on a touchscreen, we opted for pre-defined Likert options. We would probably choose otherwise if we are to pursue this survey further in the future.

DISCUSSION

In the past few years, research careers raised worrying concerns. Indeed, we know that researchers generally work more hours than they are paid for (Kinman & Wray, 2013; Koens & Jonge, 2018), face high performance pressure (Maher & Sureda Anfres, 2016; Tijdink et al., 2016), are at high risks of stress and mental health issues (Evans, Bira, Gastelum, Weiss, & Vanderford, 2018; Kinman &

Wray, 2013; Levecque, Anseel, De Beuckelaer, Van der Heyden, & Gisle, 2017; Woolston, 2017), and often experience burnout (Padilla & Thompson, 2016). This grim portrait of academic careers has raised the alert in the scientific community (e.g., Farrar, 2019; "The mental health of PhD," 2019; Stroobants, Godecharle, & Brouwers, 2013) and reinforced the need to join efforts in order to address unhealthy dynamics. One recurrent issue thought to play a key role in this problematic climate is the inadequacy of current research assessments. Indeed, the perceived inadequacy of current research assessments is such that a few organisations and movements have already issued recommendations to encourage changes (e.g., American Society for Cell Biology, 2013; Hicks et al., 2015; Moher et al., 2019; Saenen, Morais, Gaillard, & Borrell-Damián, 2019; Wilsdon et al., 2015). Nevertheless, a hefty debate remains, with some thoroughly approving of the points raised in those recommendations, and others also finding value in the current methods (Saha, Saint, & Christakis, 2003; Traag & Waltman, 2019; Tregoning, 2018). In our failure to find and agree on an alternative, research assessments are most often left untouched.

Our findings add to existing insights on the habits, wishes, and perspectives that researchers hold towards research and research assessments. In particular, our results provide a more granular understanding of specific indicators used to assess success in science and detail whether these indicators are believed to help advance science, to help fulfil personal satisfaction, or simply to advance one's career without equally contributing to scientific advancement or personal satisfaction.

Overworked and still lacking time for research

Almost 80% of full-time researchers who responded to our survey report to work more than 40 hours per week, with 44% stating that they work more than the weekly maximum authorised by the European Union (Directive 2003/88/EC, 2003). This finding is no surprise since researchers are known to work overtime and outside office hours (Barnett et al., 2019; Kinman & Wray, 2013; Koens & Jonge, 2018). Our findings also reveal that researchers are unsatisfied with the ways in which they need to distribute their time. Respondents wished they could dedicate more of their time to teaching and research, especially to tasks such as "Hands on research work", "Staying up to date", and "Writing papers", a finding that corroborates with similar works (Koens & Jonge, 2018; Mergaert & Raeymaekers, 2017). On the other hand, respondents wished they could spend less of their time performing other activities such as administration and meetings. This apparent struggle resonates with past works that expose the significant administrative burden of current research careers (Matthews, 2018; Schneider et al., 2014; Ziker, 2014) and with the interviews and focus groups that shaped the current survey in which different research actors target the lack of time for research as an important problem which ultimately lead to a number of issues that jeopardized the integrity and the quality of their work (Chapter 4).

Lack of reward for openness, transparency, quality, and high risk research

Our findings on specific research assessment indicators provide an overview of the areas which have more importance on advancing science and those which, in turn weigh more on career advancement without necessarily helping to advance science. We were not surprised to find that practices meant to promote openness and transparency (i.e., publishing findings that did not work; sharing your full data and detailed methods; publishing open access), quality (i.e., replicating past research; peer reviewing, and reviewing raw data from students and collaborators), and high risk research (i.e., conducting innovative research with a high risk of failure), were often described to be important or even essential for advancing science, but irrelevant, unfavorable, or sometimes detrimental in advancing one's career. This perspective is shared by several of the important documents and works on research assessments (American Society for Cell Biology, 2013; Hicks et al., 2015; Moher et al., 2019; Wilsdon et al., 2015), and was also thoroughly visible in our qualitative works (Chapter 3 and 4). Following this finding, it seems obvious that research assessments need to stimulate openness and quality, as well as to accept the importance of failure. Nonetheless, our survey also captured nuances which could lead to potential barriers in these areas. First, the distribution of answers and the open comments allowed us to grasp that not everyone is convinced of the value of open access, nor of the added benefit of openly sharing data and methods. Some respondents rated open access as detrimental in advancing science, while some assumed that its profit model based on publication implied bigger biases and lower quality assurance. This finding, which was also echoed in our qualitative interviews (Chapter 4), highlights that the current views on what constitute best practices are not yet uniform, and that greater and more generalised awareness is needed before customs and cultures can change. Other respondents mentioned that, although they would support open access in theory, the lack of funding for article processing charges prevented them from publishing in open access journals. Similar issues were noted when discussing publication of negative findings and data sharing, where respondents explained that such tasks come with an added burden and new ethical challenges to which they have no time to dedicate. Given that researchers mentioned lacking both the support to tackle ethical challenges (Chapter 4) and the time to undertake new research tasks (Chapter 4), valuing openness in research assessments needs a restructuration that goes far beyond research assessments. If research assessments are to formally value openness, researchers must be given the resources, infrastructures, and potentially even the workforce necessary to undertake such practices without increasing already existing burdens. Valuing openness without providing such resources risks increasing inequalities by further benefiting already successful research groups and disadvantaging young researchers, small institutes, and divergent research fields.

An overemphasis on competition and prestige

Our findings also help exemplify the overemphasis of current research assessments on competition and prestige. In fact, respondents stated that it was important for their career to publish more papers than others, to publish in high impact journals, to be cited, and to have a strong network of renowned researchers. These indicators, however, were said to be of lesser importance in advancing science and in contributing to respondents' personal satisfaction. In today's academia, researchers are expected to be excellent, yet their excellence is only recognized if they are highly productive, visible, and impactful (Chapter 3 and 4), three characteristics which, when added to the scarcity of senior positions available (Anonymous, 2010; Debacker & Vandevelde, 2016), nurture very competitive climates (Martinson, 2011). The tight competition forces researchers to spend a lot of their research time writing grants to compensate for chances of success which are often negligible. In turn, the colossal demands for research money also adds pressure to the funders who face an overload of applications to revise (Chapter 4). Paradoxically however, funders also need to ask researchers to peer-review and judge applications, further reducing the time that researchers have available for conducting good research.

CONCLUSIONS

Our survey grasps the perspective of researchers on the value that different research activities have in advancing science, in advancing research careers, and in contributing to researchers’ personal satisfaction. We found that respondents would like to be able to dedicate more time on direct research activities such as writing papers and performing hands-on research work, and wish they could dedicate less time to other tasks such as meetings and administration. Our survey also reveals that many research practices related to openness, transparency, quality, and acceptance of failure are perceived as important or even essential in advancing science, but are seen as irrelevant or even sometimes detrimental in advancing researchers' careers. Conversely, some practices which inflate the prestige, visibility, and competitiveness of researchers are seen as important in career assessments, but much less relevant in advancing science. It is important to consider that our survey captured the perspectives of a limited sample of predominantly Flemish researchers and may thus be of limited generalisability. Nonetheless, our findings align with a growing body of international works, declarations, and reports on the topic (see for example American Society for Cell Biology, 2013; Hicks et al., 2015; Moher et al., 2020; Moher et al., 2018; Nuffield Council of Bioethics, 2014; Saenen et al., 2019). Together with this growing body of literature, our findings support that research assessments need to be addressed so that researchers' careers consider activities that pursue the genuine advancement of science. Yet, our findings also show that there are nuances and disagreements on the impact of specific practices in advancing science. To ensure that changes to research assessments benefit, rather than worsen, research practices and researchers' working conditions, a thorough restructuration of the resources and infrastructures needs to take place. Beyond recognizing the importance of openness, transparency, and quality, institutions and funders should work together to enable the establishment of local resources that assist and support researchers in fostering these values.

REFERENCES

Allen, E., & Seaman, C. A. (2007). Likert Scales and Data Analyses. *Quality Progress, 40*(7), 64-65.

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Anonymous. (2010, December 18). The disposable academic. *The Economist*. Retrieved from https://www.economist.com/christmas-specials/2010/12/16/the-disposable-academic.

Barnett, A., Mewburn, I., & Schroter, S. (2019). Working 9 to 5, not the way to make an academic living: observational analysis of manuscript and peer review submissions over time. *BMJ, 367*.

Bothwell, E. (2018, 8 February). Work-life balance survey 2018: long hours take their toll on academics. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/features/work-life-balance-survey-2018-long-hours-take-their-toll-academics.

Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation, 12*(1), 39-46. doi:10.3152/147154403781776780

Debacker, N., & Vandevelde, K. (2016). From PhD to professor in Flanders. *ECOOM Brief (no. 11)*. Retrieved from https://biblio.ugent.be/publication/8043010.

Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time E. Union (2003). Retrieved from http://data.europa.eu/eli/dir/2003/88/oj

DORA. (2020). *Good Practices: Research Institutes*. Retrieved from https://sfdora.org/good-practices/research-institutes/

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*, 282. doi:10.1038/nbt.4089

Farrar, J. (2019). Why we need to reimagine how we do research. *Wellcome - News and views*. Retrieved from https://wellcome.ac.uk/news/why-we-need-reimagine-how-we-do-research

Ghent University is changing course with a new career model for professorial staff. (7 December 2018). Retrieved from https://www.ugent.be/en/news-events/new-career-model-professorial-staff.htm

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature News, 520*, 429-431. doi:10.1038/520429a

Imai, T., Kuwahara, K., Miyamoto, T., Okazaki, H., Nishihara, A., Kabe, I., . . . Japan Epidemiology Collaboration on Occupational Health Study, G. (2016). Validity and reproducibility of self-reported working hours among Japanese male employees. *Journal of occupational health, 58*(4), 340-346. doi:10.1539/joh.15-0260-OA

Kinman, G., & Wray, S. (2013). *Higher stress: A survey of stress and well-being among staff in higher education*. University College Union (UCU). Retrieved from https://www.ucu.org.uk/media/5911/Higher-stress-a-survey-of-stress-and-well-being-among-staff-in-higher-education-Jul-13/pdf/HE_stress_report_July_2013.pdf

Koens, L., R., & Jonge, H. a. J. d. (2018). *What motivates researchers? Research excellence is still a priority*. Rathenau Instituut. Retrieved from https://www.rathenau.nl/sites/default/files/2018-07/What%20motivates%20researchers.pdf

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868-879. doi:10.1016/j.respol.2017.02.008

Maher, B., & Sureda Anfres, M. (2016). Young scientists under pressure: what the data show. *Nature, 538*, 444-445. doi:10.1038/538444a

Mangiafico, S. S. (2016). Introduction to Likert Data. In *Summary and Analysis of Extension Program Evaluation in R* (pp. 152-197). New Brunswick (NJ): Rutgers Cooperative Extension.

Martinson, B. C. (2011). The academic birth rate. Production and reproduction of the research work force, and its effect on innovation and research misconduct. *EMBO Reports, 12*(8), 758-761. doi:10.1038/embor.2011.142

Matthews, D. (2018, 8 November). If you love research, academia may not be for you. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/blog/if-you-love-research-academia-may-not-be-you.

Mckenna, L. (2018, 7 February). How Hard Do Professors Actually Work? *The Atlantic*. Retrieved from https://www.theatlantic.com/education/archive/2018/02/how-hard-do-professors-actually-work/552698/.

The mental health of PhD researchers demands urgent attention. (2019). *Nature, 575*, 257-258. doi:10.1038/d41586-019-03489-1

Mergaert, L., & Raeymaekers, P. (2017). *Researchers at Belgian Universities: What drives them? Which obstacles do they encounter?* Kind Baudouin Foundation. [ISBN: D/2893/2017/16]. Retrieved from https://www.kbs-frb.be/en/Virtual-Library/2017/20171113PP

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2019). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *OSF Preprints*. doi:10.31219/osf.io/m9abx

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS Biology, 18*(7), e3000737. doi:10.1371/journal.pbio.3000737

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology, 16*(3), e2004089. doi:10.1371/journal.pbio.2004089

Nuffield Council of Bioethics. (2014). *The culture of scientific research in the UK*. Retrieved from https://nuffieldbioethics.org/publications/the-culture-of-scientific-research

Padilla, M. A., & Thompson, J. N. (2016). Burning Out Faculty at Doctoral Research Universities. *Stress Health, 32*(5), 551-558. doi:10.1002/smi.2661

Powell, K. (2016). Young, talented and fed-up: scientists tell their stories. *Nature, 538*, 446-449. doi:10.1038/538446a

Saenen, B., Morais, R., Gaillard, V., & Borrell-Damián, L. (2019). *Research Assessment in the Transition to Open Science: 2019 EUA Open Science and Access Survey Results*. European Universities Association. Retrieved from https://www.eua.eu/downloads/publications/research%20assessment%20in%20the%20transition%20to%20open%20science.pdf

Saha, S., Saint, S., & Christakis, D. A. (2003). Impact factor: a valid measure of journal quality? *Journal of the Medical Library Association : JMLA, 91*(1), 42-46.

Schneider, S. L., Ness, K. K., Shaver, K., & Brutkiewicz, R. (2014). *Federal Demonstration Partnership 2012 Faculty Workload Survey - Research Report*. Retrieved from https://osp.od.nih.gov/wp-content/uploads/SMRB_May_2014_2012_Faculty_Workload_Survey_Research_Report.pdf

Stroobants, K., Godecharle, S., & Brouwers, S. (2013). Flanders overrates impact factors. *Nature, 500*(7460), 29-29. doi:10.1038/500029b

Sullivan, G. M., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education, 5*(4), 541-542. doi:10.4300/JGME-5-4-18

Tijdink, J. K., Schipper, K., Bouter, L. M., Maclaine Pont, P., de Jonge, J., & Smulders, Y. M. (2016). How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers. *BMJ Open, 6*(2), e008681. doi:10.1136/bmjopen-2015-008681

Traag, V. A., & Waltman, L. (2019). Systematic analysis of agreement between metrics and peer review in the UK REF. *Palgrave Communications, 5*(1), 29. doi:10.1057/s41599-019-0233-x

Tregoning, J. (2018). How will you judge me if not by impact factor? *Nature, 558*(7710), 345-345. doi:10.1038/d41586-018-05467-5

VSNU, NWO, NFU and ZonMw drive change in recognition and reward of academics. (2018, 26 November). Retrieved from https://vsnu.nl/en_GB/news-items/nieuwsbericht/489-vsnu-nwo-nfu-and-zonmw-drive-change-in-recognition-and-reward-of-academics.html.

Wilsdon, J., Liz Allen, Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from https://re.ukri.org/documents/hefce-documents/metric-tide-2015-pdf/

Woolston, C. (2017). Graduate survey: A love–hurt relationship. *Nature, 550*(7677), 549-552. doi:10.1038/nj7677-549a

Ziker, J. (2014, 31 March). The Long, Lonely Job of Homo Academicus. *The Blue Review*. Retrieved from https://www.boisestate.edu/bluereview/faculty-time-allocation/

General discussion

The present thesis reflects on success in science and questions its interaction with research integrity. In this PhD, I explored the landscape of research on research integrity and built a comprehensive perspective of current issues that threaten the integrity of science. Using three different methodologies — a thorough literature analysis, semi-structured interviews and focus groups, and a survey — I found that research assessments are particularly problematic and need to be remodeled to promote better integrity in science. In this last chapter, I combine all our findings to propose four major recommendations which I believe could help academia foster better research. Before expanding on these recommendations, let me revise our findings briefly.

In our literature analysis (Chapter 1), we reviewed ten years of research on research integrity. In doing so, we extracted the topics of nearly a thousand articles on research integrity, and extracted the methods, approaches, and focus of over 300 empirical articles. We found that empirical research on research integrity extensively involves researchers and research students, but rarely involves other research actors, such as institution leaders, funders, or policy makers. We also found that empirical research on research integrity most often identifies problems from the system, such as pressures, perverse incentives, and competition, as causes for deviations from integrity. Yet, most empirical research proposing approaches to foster integrity targets individual researchers, generally through training modules or guidelines, rather than known problems from system.

In order to remedy these two inconsistencies from past research on research integrity, we decided to build a project in which we would involve actors beyond researchers and research students, and in which we would tackle an issue that is embedded in the research system and is tightly linked to pressures, incentives, and competition: the attribution of success in science. We conducted interviews and focus groups with many different research actors to discuss success, integrity, and responsibilities in science. Resonating with past research (e.g., Anderson, Ronning, De Vries, & Martinson, 2007; Davies, 2019; Fanelli, 2010; Singh & Guram, 2014; Tijdink, Verbeke, & Smulders, 2014; Wester, Willse, & Davis, 2010), our findings reiterate that problems in the system play an important part in explaining integrity failures. In particular, our interviews and focus groups suggest that the way in which researchers are being assessed, promoted, and rewarded may actually be the root of the problem by increasing pressures and competition. Our findings also added a new dimension to the well-known concept of pressures by showing that pressures do not only affect researchers, but also permeate through several layers of the research system to impact other actors such as funders, institutions, and publishers. In addition, our findings further reinforced the importance of tackling the precariousness of early research careers, suggesting that the highly selective career structure further impresses and solidifies the problems of current research assessments. Finally, although most agreed that important changes are needed to improve research climates, responsibilities for change were passed from one actor to the next leading to a circle of blame, hopelessness, and inaction.

The final step of our project consisted of a survey to capture researchers' working habits as well as the importance of specific success indicators in advancing science, in advancing careers, and in contributing to personal satisfaction. We found that most respondents declared working overtime, many even working more than the maximum European Working Time Directive. We also found that despite spending most of their time doing research, researchers wished they could dedicate more time to it and spend less time doing 'other tasks' such as administration and meetings. The survey further revealed that many activities considered important or essential in advancing science, such as activities that contribute to openness, transparency, quality, and high-risk innovative research, were considered largely insignificant in advancing research careers. Conversely, certain indicators which were considered important or essential in advancing one's career, such as indicators of prestige and productivity, were thought to be of little importance or even to be detrimental in advancing science. Finally, the survey revealed that although indicators that foster openness, transparency, and quality should be recognised in researcher's careers, adequate infrastructures and resources must be put in place if we wish to avoid further inequalities and pressures on researchers.

In light of our combined findings, it is obvious that there are issues in the core organisation of science and scientific careers. Combining the key messages of the different steps of our project, I built four broad recommendations for change which I believe should be prioritized to help promote better science.

FOUR RECOMMENDATIONS FOR CHANGE

The following recommendations are grounded on the results of our empirical works and constitute a logical response to the challenges we have observed in our study. Given the profound multidimensionality of science, these four recommendations do not address single actors, but rather address changes which need to take place concurrently, or at least collaboratively, with several stakeholders. These recommendations also do not necessarily tackle the depth of the issues that have been discussed in the past, but they reflect the direct findings of the current thesis. I will briefly place these recommendations in context in the next section to highlight examples of initiatives that are already moving science ahead.

1. LOOK BEYOND THE RESEARCHER

One of the most important points that transpires from this thesis is that research integrity does not only depend on researchers, but also hinges heavily on the system and the cultures in place. Even though blatant misconduct was described as resulting from misplaced egos and morals, less obvious detrimental research practices — which are thought to have an even greater cumulative impact on the integrity of science (Bouter, Tijdink, Axelsen, Martinson, & ter Riet, 2016) — were most frequently described as resulting from the inadequacy of the demands and climates in place (Chapter 4). Our findings are far from the first to highlight the crucial role that research climates play on research practices and integrity (Anderson et al., 2007; Davies, 2019; DuBois et al., 2013; Fanelli, 2010; Kaiser et al., 2012; Lundh, Krogsbøll, & Gøtzsche, 2012; Shrader-Frechette, 2011; Singh & Guram, 2014; Tijdink et al., 2014; Wester et al., 2010). Still, in today's academia, the majority of approaches aiming to tackle misconduct capitalise on personal behaviours rather than on resolving the systemic problems behind faulty research climates (Chapter 1).

This person-centred perspective has profound implications on the way we perceive integrity. On the one hand, it implies that research integrity is predominantly a responsibility of researchers rather than a shared responsibility of different stakeholders — a point we describe further in Chapter 1 as possibly being rooted in the evolution of the discourse on research misconduct. In fact, codes of conduct, integrity courses, whistleblowing channels, and internal oversight all point to individual researchers. Without discrediting the value of these approaches in building a solid culture of integrity among researchers, our findings support that these approaches are only a temporary relief from a problem whose roots extend deeper within the research system, and that other actors also have a crucial role to play in reshaping research climates. On the other hand, by centering our efforts on integrity training without changing inadequate demands, we increase the dilemma that researchers already endure between what they are trained to do to maximise research integrity and what they need to do to survive and succeed in academia (Chapter 5). Beyond the potential stress that this dissonance imposes on researchers, it also risks risks creating a system in which transgressions are normalised (Chapter 4) and in which success depends on strategic science (Chapter 5).

Consequently, our finding highlight the need for research integrity to move beyond the researcher and to study, understand, and tackle systemic problems. In doing so, research on research integrity needs to build strong connections with other fields of meta research such as research on research assessments, scientometrics, innovative publishing models, open science, responsible research innovations, and research on human resources and professional wellbeing. Interdisciplinary collaboration with these fields, and a careful consideration of the empirical evidence available from past research on research integrity are essential to determine the elements that motivate and drive researchers and to propose sustainable approaches that extend beyond awareness and compliance.

2. Rethink research assessments

It is paramount that research assessments are addressed and tackled. I have mentioned this point a number of times in the present thesis, but I believe that it merits being reiterated, once more, as something that must become a priority for the future of science. On this recommendation, five main issues require explicit attention. This recommendation, as I address it, predominantly targets research institutes and research funders, as well as researchers who act as referees or peer-experts. Yet, as I will point out in the fifth issue, this recommendation also addresses actors who assess the performance of universities, be it for ranking, assessment, or resource distribution purposes.

First, **research assessments must be transparent and reflective in their use of metrics**. Although metrics may add value when used properly, they are too often used in situations for which they are unfit and ultimately uninformative (Van Noorden, 2010). The impression of objectivity and detail that they provide (see for instance the discussion on the three decimal point of the impact factor in Gingras, 2016) is often tainted with important drawbacks which make metrics no more objective than peer-review. In addition, the limited information they provide creates a static definition of success, one which has been criticized for ignoring social value and innovation (Lebel & McLean, 2018; Schmidt, 2020). Several international efforts already expressed their concerns on the threat of misused metrics. The San Francisco Declaration on Research Assessment (DORA; American Society for Cell Biology, 2013), the Metric Tide (Wilsdon et al., 2015), the Leiden Manifesto (Hicks, Wouters, Waltman, Rijcke, & Rafols, 2015), and the Hong Kong Principles for Assessing Researchers (Moher et al., 2019), for example, explicitly worry about the dangers of overusing maladapted metrics. These important efforts further support that research assessments should not rely on metrics alone, but should also involve careful and reflective human review to ensure an appropriate translation and interpretation of indicators. Research institutions and research funders in particular have a role to play in acknowledging the impact of current approaches and in reflecting on responsible ways to use metrics in their assessments.

Another issue is the fact that **current assessment predominantly target individual researchers rather than teams, departments, or the research itself**. As I have described in Chapter 3, even when research assessments target departments or institutions (for instance the Research Excellence Framework (REF) in the UK or the Bijzondere Onderzoeksfonds [special research funds] (BOF) in Flanders), they tend to be reflected in individual key performance indicators. In addition, I have discussed that, even in mandates where funding is evaluating project proposals, researchers' profiles and past success play an important part in the decisions (Chapter 2 and 3). Consequently, individual performance is always at the heart of grants attribution, employments, awards, tenure and promotions. Focusing on individual researchers, however, ignores the essential role of teams in research and knowledge production (Chapter 5). It expects that every researcher excels in similar ways and in all steps of research, teaching, and services. In reality, we know that individuals have specific skills, preferences, and talents which may not always equate those of others. Some excel at communicating the results to the public, some excel at scientific writing, others are rigorous data analysts, others have strong networks which ensure broad collaborations, etc. Expecting that all researchers perform in the same way and deliver the same outputs indubitably challenges the practices and overall productivity of research. It makes every researcher a competitor to others rather than a team or a community member who shares personal skills for the common benefit of advancing science (see Chapter 5). A few movements are starting to recognize the importance of collaborative efforts (see for example Bothwell, 2019; CASRAI; "VSNU, NWO, NFU and ZonMw," 2018), but much more needs to happen to de-individualize research achievements. These impact of individual assessments must be carefully considered by institutions and other actors. For instance, publishers could consider using contributorship on top of (or even in place of) authorship (McNutt et al., 2018), universities could consider aggregating individual requirements to permit specialisation of researchers in internal teams (Mishra, 2015), and research assessments could allow for individuals to set their own goals of success rather than to have them imposed to them based on a fixed and universal view of success. All of these initiatives are starting to appear (see examples in Moher et al., 2020), but they need to become commonplace.

A third issue of current assessments is their **reliance on outputs rather than on the processes and practices attached to the research**. Past outputs are paramount to tenure and promotion, but also play into grants and scholarships where the profiles of researchers are evaluated. The idea behind this reliance on outputs is that past successes are generally a good indicator of future performance (Chapter 3). Yet, reliance on outputs comes with two important challenges. First, it encourages researchers to focus their efforts on producing outputs rather than on developing good processes and embracing good research practices (e.g., openness, transparency, collegiality, etc.); a trend which encourages quantity rather quality. Second, it further increases the gap between early career and senior researchers, making the former vulnerable and the latter virtually invincible. As a result, early career researchers not only struggle and often fail to pursue an academic career, but they also often need to abandon innovative disruptive ideas and high risk research to be able to stand a chance amongst established researchers that surround and evaluate them (Maher & Sureda Anfres, 2016; Oni, Sciarrino, Adesso, & Knight, 2016). A restructuration of research assessments to consider the processes behind the output would help increase the visibility of openness (open access, open data, open codes, etc.), transparency (registration of studies, replicability, dissemination of negative findings, etc.), collegiality (peer review, acknowledgements, collaborations), and other aspects that are known to be important in advancing research (Chapter 6). Nevertheless in changing research assessments, the infrastructures and the resources in place must be carefully adapted. For instance, we found that although many researchers support open access and would be willing to share their data online to promote the transparency and reuse of the work, they lack the funds, time, support, and expertise to adopt such practices (Chapter 6). In changing research assessments, it is thus essential that infrastructures are adapted to ensure a smooth transition that benefits science without increasing inequalities and undue burden on researchers.

A fourth key problem with research assessments is the **lack of realism in what is expected from researchers**. I have highlighted the fact that current assessments focus on the individual and on outputs, thereby expecting researchers to be equally excellent at every step of the research process. But assessments also expect researchers to devote themselves to science at the expense of their personal life and mental health. As I have shown in Chapter 5, the perspective of the devoted, selfless researcher has historical roots and was a key aspect described by Merton decades ago (Merton, 1973). In current research careers however, expecting researchers to do more for less creates a tension which has important repercussions. First, many researchers work much more than they are paid for (Chapter 6), a habit which is thought to threaten their wellbeing and possibly lead to burn out (Evans, Bira, Gastelum, Weiss, & Vanderford, 2018; Levecque, Anseel, De Beuckelaer, Van der Heyden, & Gisle, 2017). But second, the realities of researchers have changed tremendously in the past decades. Researchers are less likely to benefit from support staff, and are more likely to have young children and working spouses (Maher & Sureda Anfres, 2016). Expecting researchers to work beyond schedule and to travel on request thus risks deepening the gender gap and the social inequalities which are already associated with research recognition. It is thus primordial that institutions, but also supervisors respect the boundaries and the work-life balance that are necessary for a healthy future of academia.

Finally, assessments for research need to be improved **on all levels of funding**. Although most of this thesis addresses research institutions' urgent responsibility in tackling research assessments, it is important to consider that institutions often build internal assessments on indicators for which they are evaluated themselves. The funding models used to fund universities often depend on performance indicators (Zacharewicz, 2016). As we have described in Chapter 2, the university funding model in Flanders is highly dependent on performance, especially on output indicators such as publication metrics (Peters, 2019). In an attempt to maximise efficiency of resources, several universities reuse these performance indicators to distribute resources within the institution (i.e., on a faculty, department, and researcher level). This transfer of indicators from university-assessments to researcher-assessment is a phenomenon that happens (and will continue to happen) even if those in charge of high level funding models explicitly warn against such use (Engels & Guns, 2018). In other words, even if public distribution keys do not directly address research departments and researchers, high level performance based funding models will always impact the researchers and practices in one way or another. A clear consideration of the impact of high level performance based funding models and their translation in university is needed, and a thorough adaptation of such funding models to promote best practices is instrumental in improving research assessments. Those in charge of determining how funding is distributed between institutions have a great and powerful opportunity to make science better. Nevertheless, research institutions also have an opportunity to assert their values in taking the lead to changing their own assessments in spite of high level funding models — an opportunity that can bring visibility, respect, and recognition (see for instance *Ghent University is changing course*, 7 December 2018).

3. Rethink the research career path

The third point I found essential to address regards academic career structures. Predominantly targeting research institutions, this point also addresses policy makers who often set the employment objectives and tie structural funding to the research workforce in place in research institutions. The European Commission's Code of Conduct for the Recruitment of Researchers states that "*Employers and/or funders should ensure that the performance of researchers is not undermined by instability of employment contracts, and should therefore commit themselves as far as possible to improving the stability of employment conditions for researchers*" (European Commission, 2014, p. 17). Still, precariousness and career insecurity were recurrent themes in our interviews and focus groups (Chapter 4 and 5). Interviewees explained that academic careers have a funnel structure in which only one tenth of PhD student will be able to secure a tenured career in academia. This proportion varies globally, generally ranging between 3% and 20%, but the problem appears to be recognised worldwide (Alberts, Kirschner, Tilghman, & Varmus, 2014; Anonymous, 2010; Debacker & Vandevelde, 2016; "Many junior scientists," 2017; Martinson, 2011). The yearning for a stable position is such that it is even thought to have misplaced tenure as the 'end goal' of science rather than as a means for academic freedom (Holbrook, 2017).

The root of the problem, at least within the biomedical sciences, is thought to be the lack of adjustment to research structures after the booming federal investments in research that many countries benefited from in the middle of the twentieth century (Alberts, Kirschner, Tilghman, & Varmus, 2015). A failure to adjust the entrance workforce to the slowing expansion of the research system in more recent years has led to an overproduction of aspiring scientists with limited opportunities for permanent careers in academia.

While a failure of adjustment may be expected for sudden changes and variations in research expenditures, the problem has been raised for over twenty years with very little noticeable improvement (see Alberts, 1999; Marincola & Solomon, 1998 for early works noticing the issue). The difficulty to justify the lingering lack of adjustment led some researchers to assume that the imbalance between junior and senior opportunities was of little worry to research managers and policy makers. In fact, a perspective that is increasingly popular in the media (see for example Hall, 2019; The disposable academic, 2016) proposes that the current overrepresentation of PhD students may be explained by the fact that junior trainees are much cheaper to maintain (i.e., their salary is most often secured externally through scholarships or is otherwise untaxed and not subject to pension schemes, making it at least half the investment of a post-doctoral researcher) while they still contribute to the university output measures and research productivity (i.e., in Flanders, most faculties of medicine ask that PhD students publish a minimum of three papers before they can defend, and the expectation persists even in institutions where this requirement is not explicit or has been removed). Another perspective however, explains that a key of mission of universities is to educate and form highly skilled workers. PhD students are thus not only part of the workforce of universities, but also beneficiaries from the education provided. Yet, interviewees reported that PhD training is often inadequate to prepare students to non-academic work, generally leaving students with a narrow area of expertise and failing to provide them with strategies and relational skills needed in other careers (also see Alberts, 1999; Heffernan & Heffernan, 2019; Van de Velde, Levecque, Mortier, & De Beuckelaer, 2019; Woolston, 2017). Past research has also found that careers outside academia are rarely discussed between students and their supervisors, and that students often feel that their supervisors look down on external career options (Woolston, 2017). Furthermore, by including researchers who left academia among our interviewees, we understood that leaving academia can leave a vivid wound, if not a feeling of failure or grief (Chapter 4). Considering that past research in Flanders, the Netherlands, and the USA has found that the majority of PhD aspire to continue in academia (Debacker & Vandevelde, 2016; Sauermann & Roach, 2012; van der Weijden, Teelken, de Boer, & Drost, 2016; also see Woolston, 2017 which surveyed readers from Springer Nature), a big proportion of young researchers will be forced to face the pain of having to leave academia against their will. Beyond the emotional distress of those who need to leave academia, the slim chance of success also means that young researchers are constantly competing with one another for securing research resources and permanent positions. To survive, they must move ahead of their colleagues, beat them at metrics, produce more outputs than them, build bigger networks then them. Such competition does not enhance the productivity of scientists in advancing science, it simply encourages young researchers to adopt competitive strategies to maximise their chances of survival (Chapter 5). It incites researchers to focus on outputs, to ignore unrewarded processes which are essential to ensure the quality of science, and to compete, mistrust, and hide rather than collaborate, trust, and share. And as if the impact on researchers and research outputs was not enough, highly selective research careers also impede the opportunity for systemic change. Indeed, because of the tight bottleneck between junior and permanent positions, only those who 'fit the cast' of current assessments will survive and succeed, while those who disagree with the demands of the system are unlikely to strive and survive (Chapter 5). As a result, research cultures are built from a selective minority of researchers who — although they might not always agree with the system as it currently stands — were able to master the skills and the tactics it takes to survive it. Those researchers will naturally be more inclined to teach and share those tactics to younger generations to help them get ahead, thereby perpetuating and reinforcing a static culture where little change can happen. In other words, on top of impacting researchers' satisfaction and research practices, highly selective careers encourage competition and maintain the uniformity of the highly-criticized research culture.

Going back to the European Charter, it then seems obvious that the current lack of stability in early stage careers *does* affect research performance. It cultivates fierce competition which further accentuates the range of problems attached to research assessments, and it must be addressed. Changes could include a re-equilibration of junior-to-senior positions by having lower PhD entrance levels, they could include changing the structure of the PhD to have fewer academia-oriented PhDs and more industry-oriented PhDs, or they could simply allow for more differentiation on what is expected from research careers, including through the creation of new, highly skilled professional roles within academia where unique skills are advanced to foster collaborative research teams (e.g., professional peer-reviewers, professional protocol designers, professional data analysts, or simply permanent research staff; Alberts, 1999; Alberts et al., 2015). One thing is certain: we cannot keep this status quo for another twenty years.

4. Foster and encourage inter-actor discussion

The last point I wish to raise is, at first glance, very simple: the discussion on these issues needs to involve the entire array of research actors.

As I have shown in Chapter 1, research on research integrity has grown substantially in the past decade, gaining visibility, credence, and momentum. But even though most research explaining why misconduct happens point to the system (i.e., pressures, competition, perverse incentives), most research proposing approaches to foster integrity focuses on the researcher's awareness and compliance rather than on changing the faulty research system. This inconsistency indicates that, even within a small field such as research on research integrity, we fail to listen and inform each other. Within the four years of this project, I have already witnessed great improvements in building a better, more interactive community for research integrity. The development of working groups and networks such as the World Conferences on Research Integrity Foundation, the Embassy of Good Science, national integrity offices as well as numerous highly visible Horizon 2020 collaborative projects pave the road to a better dialogue between research integrity experts. But even in this new dialogue, the stakeholders involved are often largely uniform. Exceptions do exist — I will review some of these exceptions in the next section — but the rule of thumb is that networks and groups tend to center around a dominant actor group, rarely a balanced mix of different actors. In addition, certain essential actors are repeatedly overlooked. In our interviews, the unfamiliarity of many key actors with the integrity jargon (Chapter 4) supports that discussions may operate in separate funnels. Non-researcher stakeholders are generally forgotten in empirical research seeking perspectives on integrity and misconduct (Chapter 1), while former-researchers and early career scientists are rarely heard in policy building and scientific publishing (again, exceptions *do* exist) even though their perspectives may be very different than that of those who survived and succeeded in the current system (Chapter 5). As a result, the discourse on research integrity risks remaining a collection of individual voices rather than a collective perspective.

We argue that the academic system needs to change (i.e., approaches to integrity, research assessments, and career structures) to promote better research integrity (Chapters 1 and 4). But systems are the result of a complex interplay of interacting, interrelated, and interdependent bodies (Kim, 1999). In hearing the voices of multiple actors, we realised that perspectives of success, integrity, and misconduct differ between individuals and actor groups (Chapter 3 and 4). Furthermore, the responsibility for change is passed from one actor to the next, creating a stagnant system that often resounds of blame and hopelessness (Chapter 4). For broad, systemic changes to be operationalized, we need to understand the dynamics and the relationships at play in the current problems of science. We need to dig deeper in the spaces and responsibilities that link different actors and that build the foundations of our shared concepts of excellence and integrity. In turn, we need to provide an open and inclusive forum for discussion between actor groups. Some propose that we need a convening body with enough authority to provide an impartial discussion forum and ensure that the propositions are taken up and implemented in a coordinated fashion (Gunsalus, McNutt, Martinson, Faulkner, & Nerem, 2019). The establishment of such a convening body would undoubtedly encourage a rich inter-actor dialogue, but in the meantime, existing resources may also be exploited. Broad expert groups such as the European Commission policy platforms or expert groups in Scientific Societies and Academies, for example, already provide a platform where the opinions of different actors meet and influence science policies. Ensuring that these platforms include the voices of all research actors at all seniority levels would be the next logical step into ensuring a rich and realistic dialogue.

Considering these four recommendations, it is clear that the very foundations of scientific careers, cultures, and integrity need to be addressed. While my thesis proposes areas in which we should focus our attention, it does not offer specific answers on how these elements should be implemented in practice. Revisiting research careers, cultures, and integrity using a Systems Thinking approach could help determine actionable ways of restructuring science (Kim, 1999; Stroh, 2015). In turn, implementing these changes and empirically examining their impact on research practices, research cultures, and researchers should become a core objective of academia.

THE RECOMMENDATIONS IN CONTEXT

Throughout this thesis, I talk about problems of science that need to be addressed. It was pointed out by one of the jury members that I failed to mention initiatives that are taking the lead in making science better. Indeed, remarkable changes are taking place and ambitious groups are actively tackling the issues of science as we speak. While it is extremely challenging to discuss examples of initiatives since I can only provide a deeply biased and profoundly incomplete sample, I also realise that concrete examples can help to inspire future action.

First, several excellent reports played a pivotal role in building awareness around the problems that science is currently facing. I mentioned the San Francisco Declaration on Research Assessment (DORA; American Society for Cell Biology, 2013), the Leiden Manifesto (Hicks et al., 2015), and the Hong Kong Principles for Assessing Researchers (Moher et al., 2019) many times in this thesis. These reports were mostly led by researcher and editors and have greatly advanced the discourse around research assessments. Research policy groups, funders, and academies also joined the debate. For example, the Metric Tide and the EUA Briefings on Research assessments continued the important discussion of research assessments (Saenen & Borell-Damián, 2019; Wilsdon et al., 2015), alongside several others, many of which are available in the review by Moher and colleagues (Moher et al., 2018). Other reports such as those from the Nuffield Council of Bioethics, The Wellcome Trust, the Royal Society and more recently Vitae in partnership with the UK Research Integrity Office (UKRIO) and the UK Reproducibility Network (UKRN) addressed the culture of research, raising issues not only towards pressures and incentives but also towards workloads, diversity, and bullying and harassment (Janet Metcalfe, Katie Wheat, Munafò, & Parry, 2020; Nuffield Council of Bioethics, 2014; The Wellcome Trust and Shift Learning, 2020). The US National Academies 'Fostering Integrity in Research' and former editions of the work are also a gold mine that has a tangible influence in the field of research integrity (NASEM, 2017).

Besides reports on the direct topics of this thesis, concrete initiatives are also taking place. Broad scale initiatives on **Open Science**, for example, are appearing all around the globe. Open science is intimately interlaced with integrity, adhering to principles such as transparency, collaboration, reproducibility, and quality.

Open science initiatives join the efforts of funders, policy makers, publishers, libraries, infrastructure providers, institutions, researchers, and several other research actors, often bringing them to work together to promote system changes. The European Commission Open Science Policy Platform Recommendations nicely illustrate the coordinated actions of the different actors involved in moving open science ahead (Open Science Policy Platform, 2017). Other actions such as Data sharing platforms, FAIR Data Principles, the Center for Open Science, cOAlition S, preprint servers, and select publishers such as the Public Library for Open Science, BioMed Central and several others are only a minute selection of the important actions that are currently changing science for the best. Another important area of initiative is **Reproducibility.** The debate around the reproducibility crisis (now also called reproducibility opportunity; Brock, 2018) led to the development of numerous interdisciplinary research centers entirely focused on making science more reproducible. Related initiatives such as registered reports, reporting checklists, and changes in editorial policies to support the publication of negative results or mandate the publication of data also led to increasing transparency in data reporting by combining the efforts of a whole array of research actors. Initiatives to involve **early career researchers** in decisions for system are also booming, with journals building early career communities (e.g., eLife early career advisory group) different actors involving early career researchers in their board decisions or in panel discussions, and young researchers creating their own groups to enter the policy debates (e.g., PhD networks, EuroDoc, Young Academy). **Responsible metrics** are also rising, with new metrics being engineered to encompass broader impact activities (e.g., Altmetrics) and several expert groups on next generation metrics (e.g., Miedema, Mayer, Holmberg, & Leonelli, 2018; Wilsdon, Bar-Ilan, Frodeman, Peters, & Wouters, 2017). **University libraries** are also taking the lead in making concrete changes in their institutions, for example through data stewards (e.g., Delft University of Technology) or in their promotion of open access (e.g., University of California and Elsevier). More concrete changes on **research assessments** are also worth mentioning. Powerful ideas such as PubPeer, open peer-review, and Contributor Role Taxonomy (CRediT) provide a great potential for change. Structural changes in assessment models are also starting to appear (e.g., Ghent University and in the Association of Universities in the Netherlands). A whole array of additional examples are detailed in the Hong Kong Principles "Current implementation" sections (Moher et al., 2020) and in the blog curated by DORA.

I could go on at least as long as this entire thesis, but I think that these few example already illustrate that there is great action towards better science and that broad inter-actor changes, although daunting, are possible.

CONCLUDING REMARKS

Looking back at our project, I realise that our findings are not new nor revolutionary. They largely align with past and current discussions on research integrity. Yet, our findings also provide a decent empirical basis to support and verify what we already know about the problems of science. Several of our findings and conclusions also add to the current discourse to establish a better understanding of the dynamics in place for different actor levels. For instance, the understanding that pressures do not only affect researchers is an important issue which has rarely been discussed in research on research integrity. Along the same line, although mistrust and blame may have served in raising awareness and in mobilising researchers in the past, inter-actor dialogues and collaboration are now necessary to initiate complex systemic changes.

In the four years that this project took place a lot has happened in the field of research integrity and especially in the topic of research assessments. As we were conducting our research, new developments, new assessment initiatives, and influential opinions on the topic would hit the news at least every week, sometimes daily. Although challenging to follow when writing a thesis, such a vivid interest in the topic is highly stimulating and very reassuring since it indicates that the research community is aware of the problems and is increasingly ready to change. There is much reason to be optimistic for the future. Still, our research suggests that topical initiatives will only grow to their full potential and change research culture if they become instrumental to broad and coordinated approaches for change. Using all these combined efforts, the time may be right to shift our attention from what needs to change to who will stand up for the change.

REFERENCES

Alberts, B. (1999). Are our universities producing too many PhDs? *Trends in Cell Biology, 9*(12), M73-M75. doi:https://doi.org/10.1016/S0962-8924(99)01686-4

Alberts, B., Kirschner, M. W., Tilghman, S., & Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences, 111*(16), 5773. doi:10.1073/pnas.1404402111

Alberts, B., Kirschner, M. W., Tilghman, S., & Varmus, H. (2015). Opinion: Addressing systemic problems in the biomedical research enterprise. *Proceedings of the National Academy of Sciences, 112*(7), 1912. doi:10.1073/pnas.1500969112

American Society for Cell Biology. (2013). *San Francisco Declaration on Research Assessment*. Retrieved from https://sfdora.org/read/

Anderson, M. S., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007). The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics, 13*(4), 437-461. doi:10.1007/s11948-007-9042-5

Anonymous. (2010, December 18). The disposable academic. *The Economist*. Retrieved from https://www.economist.com/christmas-specials/2010/12/16/the-disposable-academic.

Bothwell, E. (2019, 14 October). Award Nobels to teams, not individual 'heroes', say scientists. *Times Higher Education*. Retrieved from https://www.timeshighereducation.com/news/award-nobels-teams-not-individual-heroes-say-scientists.

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review, 1*(1), 17. doi:10.1186/s41073-016-0024-5

Brock, J. (2018, 17 December). "It's not a replication crisis. It's an innovation opportunity". *Medium*. Retrieved from https://medium.com/dr-jon-brock/its-not-a-replication-crisis-it-s-an-innovation-opportunity-915bc058683b.

CASRAI. *CRediT – Contributor Roles Taxonomy*. Retrieved from https://casrai.org/credit/

Davies, S. R. (2019). An Ethics of the System: Talking to Scientists About Research Integrity. *Science and engineering ethics, 25*(4), 1235-1253. doi:10.1007/s11948-018-0064-y

Debacker, N., & Vandevelde, K. (2016). From PhD to professor in Flanders. *ECOOM Brief (no. 11)*. Retrieved from https://biblio.ugent.be/publication/8043010.

DuBois, J. M., Anderson, E. E., Chibnall, J., Carroll, K., Gibb, T., Ogbuka, C., & Rubbelke, T. (2013). Understanding Research Misconduct: A Comparative Analysis of 120 Cases of Professional Wrongdoing. *Accountability in Research, 20*(5-6), 320-338. doi:10.1080/08989621.2013.822248

Engels, T. C. E., & Guns, R. (2018). The Flemish Performance-based Research Funding System: A Unique Variant of the Norwegian Model. *Journal of Data and Information Science, 3*(4), 45-60. doi:https://doi.org/10.2478/jdis-2018-0020

European Commission. (2014). *The European Charter for Researchers*. Retrieved from http://ec.europa.eu/euraxess/index.cfm/rights/europeanCharter

Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology, 36*, 282. doi:10.1038/nbt.4089

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS ONE, 5*(4), e10271. doi:10.1371/journal.pone.0010271

Ghent University is changing course with a new career model for professorial staff. (7 December 2018). Retrieved from https://www.ugent.be/en/news-events/new-career-model-professorial-staff.htm

Gingras, Y. (2016). *Bibliometrics and Research Evaluation: Use and Abuses*. Cambridge (MA): MIT Press.

Gunsalus, C. K., McNutt, M. K., Martinson, B. C., Faulkner, L. R., & Nerem, R. M. (2019). Overdue: a US advisory board for research integrity. *Nature, 566*, 173-175. doi:doi: 10.1038/d41586-019-00519-w

Hall, R. (2019, 1 April). Are PhDs just cheap labour for universities? *The Guardian*. Retrieved from https://www.theguardian.com/education/2019/apr/01/are-phds-just-cheap-labour-for-universities.

Heffernan, T. A., & Heffernan, A. (2019). The academic exodus: the role of institutional support in academics leaving universities and the academy. *Professional Development in Education, 45*(1), 102-113. doi:10.1080/19415257.2018.1474491

Hicks, D., Wouters, P., Waltman, L., Rijcke, S. d., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature News, 520*, 429-431. doi:10.1038/520429a

Holbrook, J. B. (2017). The future of the impact agenda depends on the revaluation of academic freedom. *Palgrave Communications, 3*(1), 39. doi:10.1057/s41599-017-0041-0

Janet Metcalfe, Katie Wheat, Munafò, M., & Parry, J. (2020). *Research integrity: a landscape study*. Retrieved from https://www.vitae.ac.uk/vitae-publications/reports/research-integrity-a-landscape-study

Kaiser, K. A., Cofield, S. S., Fontaine, K. R., Glasser, S. P., Thabane, L., Chu, R., . . . Allison, D. B. (2012). Is funding source related to study reporting quality in obesity or nutrition randomized control trials in top-tier medical journals. *International Journal of Obesity, 36*(7), 977-981. doi:10.1038/ijo.2011.207

Kim, D. H. (1999). *Introduction to Systems Thinking*. Cambridge, UK: 1999 by Pegasus Communications, Inc.

Lebel, J., & McLean, R. (2018). A better measure of research from the global south. *Nature, 559*, 23-26. doi:10.1038/d41586-018-05581-4

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., & Gisle, L. (2017). Work organization and mental health problems in PhD students. *Research Policy, 46*(4), 868-879. doi:10.1016/j.respol.2017.02.008

Lundh, A., Krogsbøll, L. T., & Gøtzsche, P. C. (2012). Sponsors' participation in conduct and reporting of industry trials: a descriptive study. *Trials, 13*, 146. doi:10.1186/1745-6215-13-146

Maher, B., & Sureda Anfres, M. (2016). Young scientists under pressure: what the data show. *Nature, 538*, 444-445. doi:10.1038/538444a

Many junior scientists need to take a hard look at their job prospects. (2017). *Nature, 550*, 429. doi:doi:10.1038/550429a

Marincola, E., & Solomon, F. (1998). The career structure in biomedical research: implications for training and trainees. The american society for cell biology survey on the state of the profession. *Molecular biology of the cell, 9*(11), 3003-3006. doi:10.1091/mbc.9.11.3003

Martinson, B. C. (2011). The academic birth rate. Production and reproduction of the research work force, and its effect on innovation and research misconduct. *EMBO Reports, 12*(8), 758-761. doi:10.1038/embor.2011.142

McNutt, M. K., Bradford, M., Drazen, J. M., Hanson, B., Howard, B., Jamieson, K. H., . . . Verma, I. M. (2018). Transparency in authors' contributions and responsibilities to promote integrity in scientific publication. *Proceedings of the National Academy of Sciences, 115*(11), 2557. doi:10.1073/pnas.1715374115

Merton, R. K. (1973). *The Sociology of science : theoretical and empirical investigations*. Chicago (IL): University of Chicago Press.

Miedema, F., Mayer, K., Holmberg, K., & Leonelli, S. (2018). *Mutual Learning Exercise - Open Science: Altmetrics and Rewards*. European Commission. Retrieved from http://publications.europa.eu/resource/cellar/449cc187-693f-11e8-ab9c-01aa75ed71a1.0001.03/DOC_2

Mishra, J. (2015, 9 November). Choosing team over tenure. *Science*. Retrieved from https://www.sciencemag.org/careers/2015/11/choosing-team-over-tenure. doi:10.1126/science.caredit.a1500258

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2019). The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity. *OSF Preprints*. doi:10.31219/osf.io/m9abx

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., . . . Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS Biology, 18*(7), e3000737. doi:10.1371/journal.pbio.3000737

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLoS Biology, 16*(3), e2004089. doi:10.1371/journal.pbio.2004089

NASEM. (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press.

Nuffield Council of Bioethics. (2014). *The culture of scientific research in the UK*. Retrieved from https://nuffieldbioethics.org/publications/the-culture-of-scientific-research

Oni, T., Sciarrino, F., Adesso, G., & Knight, R. (2016). Let researchers try new paths. *Nature, 538*(7626), 451-453. doi:10.1038/538451a

Open Science Policy Platform. (2017). *OSPP-REC*. European Commission. [ISBN 978-92-79-88333-0]

Peters, A. (2019). *Public funding of universities, and indicators used in the funding systems : a Euregional perspective.* Université de Liège, Liège, Belgium. Retrieved from https://matheo.uliege.be/handle/2268.2/8014

Saenen, B., & Borell-Damián, L. (2019). *EUA Briefing - Reflections on University Research Assessment: Key concepts, issues and actors*. European Universities Association. Retrieved from https://eua.eu/resources/publications/825:reflections-on-university-research-assessment-key-concepts,-issues-and-actors.html

Sauermann, H., & Roach, M. (2012). Science PhD Career Preferences: Levels, Changes, and Advisor Encouragement. *PLoS ONE, 7*(5), e36307. doi:10.1371/journal.pone.0036307

Schmidt, R. (2020, 24 August). The Benefits of Statistical Noise. *Behavioral Scientist*. Retrieved from https://behavioralscientist.org/the-benefits-of-statistical-noise/

Shrader-Frechette, K. (2011). Climate Change, Nuclear Economics, and Conflicts of Interest. *Science and Engineering Ethics, 17*(1), 75-107. doi:10.1007/s11948-009-9181-y

Singh, H. P., & Guram, N. (2014). Knowledge and attitude of dental professionals of North India toward plagiarism. *North American Journal of Medical Sciences, 6*(1), 6-11. doi:10.4103/1947-2714.125854

Stroh, D. P. (2015). *Systems Thinking for Social Change*. Hartford (VT) USA: Chelsea Green Publishing Company

The disposable academic. (2016, 27 December). Why doing a PhD is often a waste of time. *The Economist*. Retrieved from https://medium.economist.com/why-doing-a-phd-is-often-a-waste-of-time-349206f9addb.

The Wellcome Trust and Shift Learning. (2020). *What Researchers Think About the Culture They Work In*. [MC-7198/01-2020/BG]. Retrieved from https://wellcome.ac.uk/reports/what-researchers-think-about-research-culture

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication Pressure and Scientific Misconduct in Medical Scientists. *Journal of Empirical Research on Human Research Ethics, 9*(5), 64-71. doi:10.1177/1556264614552421

Van de Velde, J., Levecque, K., Mortier, A., & De Beuckelaer, A. (2019, September). Waarom doctorandi in Vlaanderen denken aan stoppen met doctoreren [Why PhD students in Flanders think about stopping their PhDs]. *ECOOM Brief (no. 20)*. Retrieved from http://hdl.handle.net/1854/LU-8630419.

van der Weijden, I., Teelken, C., de Boer, M., & Drost, M. (2016). Career satisfaction of postdoctoral researchers in relation to their expectations for the future. *Higher Education, 72*(1), 25-40. doi:10.1007/s10734-015-9936-0

Van Noorden, R. (2010). Metrics: A profusion of measures. *Nature, 465*(7300), 864-866. doi:10.1038/465864a

VSNU, NWO, NFU and ZonMw drive change in recognition and reward of academics. (2018, 26 November). Retrieved from https://vsnu.nl/en_GB/news-items/nieuwsbericht/489-vsnu-nwo-nfu-and-zonmw-drive-change-in-recognition-and-reward-of-academics.html.

Wester, K., Willse, J., & Davis, M. (2010). Psychological climate, stress, and research integrity among research counselor educators: A preliminary study. *Counselor Education and Supervision, 50*(1), 39-55. doi:10.1002/j.1556-6978.2010.tb00107.x

Wilsdon, J., Bar-Ilan, J., Frodeman, R., Peters, I., & Wouters, P. (2017). *Next-generation metrics: Responsible metrics and evaluation for open science*. European Commission.

Wilsdon, J., Liz Allen, Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Retrieved from https://re.ukri.org/documents/hefce-documents/metric-tide-2015-pdf/

Woolston, C. (2017). Graduate survey: A love–hurt relationship. *Nature, 550*(7677), 549-552. doi:10.1038/nj7677-549a

Zacharewicz, K. J. T. (2016). *Research Performance Based Funding Systems: a Comparative Assessment*. European Commission. [ISBN 978-92-79-68715-0 - ISSN 1831-9424]. Retrieved from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC101043/kj1a27837enn.pdf

Additional outcomes from the current thesis

TEACHING AND EDUCATION

An important accomplishment of the current project was its inspiration in developing educational material to help other researchers understand research integrity and question research cultures. The courses that resulted from this project are not typical ‘good practice’ courses, but rather reflexive courses in which the limits of science are openly discussed. The courses address some of the issues that were raised in my findings and aim to raise awareness, better understanding, and resilience on the flaws that are inherent to current research systems.

Scientific publishing

Hasselt University; spring 2017, 2018, 2019; course intended for 3rd year bachelors’ student of Medicine.

> Intensive one-week course as part of the UHasselt GLW exploration week. In this short course, I introduce students to scientific publishing, showing them how scientific publishing works and helping them write a simple academic abstract in English. Although the course targets technical aspects of scientific publishing such as the documentation, submission, review, and writing processes, I spend one day to focus on the ethics of scientific publishing. Throughout the course, I also emphasise the importance of communicating science transparently, of sharing data and methods, of listening, discussing, and being constructively critical with one another, and of being transparent, honest, and comprehensible throughout the research process. This course equips students to better understand how to publish, but also to capture and confront the imperfections of the current system.

Publication Ethics: Where are the gaps and how can we address them

Hasselt University; summer 2020; course intended for PhD students.

> This 6-hour course introduces topics of publication ethics and provides a toolbox to allow PhD students to tackle some of the issues in the current research system. Half of he course is comprised of discussions in which I invite students to reflect and discuss the foundation of these problems and the actions that they can take to help change the problematic culture of science. The course covers topics of authorship, originality (plagiarism, duplication, salami slicing), transparency (pre-registration, selective reporting, corrections, retractions), openness (preprints, open access, open data, predatory publishers), impact (publication metrics and their limits), and peer-review. The course has been digitalized to adapt to the COVID-19 pandemic, and has already been offered to five groups, with more groups planned for the fall 2020.

Publication strategies — Great science and high impact factors

Hasselt University; spring 2019; course intended for PhD students and post-doctoral researchers.

This 3.5-hour course was organized with Prof. dr. Dominique Hansen and Prof. dr. Sven Hendrix in Hasselt University. The course targets publication strategies and the relation between excellence, impact factors, and ethical challenges. In my section, I discuss 'the good, the bad, and the ugly' of the Journal Impact Factor. I introduce students to the value and the limits of this highly contradicted but poorly understood metric. I introduce DORA (the Declaration on research assessments) and explain how the impact factor is truly measured, showing why we should be careful about using it to compare fields or evaluate individual papers or researchers.

CONFERENCE PRESENTATIONS

In a field of general interest like research integrity and research assessments, conference presentations are also important to raise awareness. During my PhD, I even had the impression that oral presentations were imperative in reaching key audiences who would be unlikely to read published papers on the topic. Since conference presentations come at a cost in time and in environmental imprint, I found important to reflect on each of them to illustrate how they helped to shape and share the present thesis.

The future of science: The role of young researchers in shaping research climates. Invited oral presentation and panel discussion. Open Access Ambassadors Conference from the Max Planck Digital Library (MPDL) and Max Planck PhDnet, Berlin (BE), 10-11 December 2019.

Description: In this presentation, I highlighted four main findings from our project. First, I described the issue in the way we assess scientists (Chapter 3). Second, I described how research assessments shape our perspective of success (Chapter 6). Third, I argued that the problem faces a vicious circle (Chapter 5). And finally, I highlighted how we all play a role in the problem (Chapter 4). I then described existing initiatives and presented ideas for young researchers to make science better.

Personal reflection: This conference was the last one of my PhD and it was one of the most gratifying. The conference was organised and almost exclusively attended by Master's and PhD students from Max Planck institutes who acted as ambassadors for open science. I was truly impressed by the level of reflection, interest, hopes, and eagerness of so many captivating young researchers. Not only were these students reflective, but they were also genuine activists in the field. They dared to ask questions which forced authoritative figures to be transparent, and they had the courage to question even their own thoughts and behaviours. It left me very impressed and hopeful for the future of science.

Enough complaining, it's time for change! Academic Culture from a graduate student's perspective. Invited oral presentation and panel discussion. *2019 OASPA (Open Access Scholarly Publishing Association) Conference on Open Access Scholarly Publishing,* Copenhagen (DK), 24–26 September 2019.

> **Description:** In this presentation, I discussed the impact of current research cultures on research integrity and on early career researchers by highlighting that research assessments are inadequate (see Chapter 3 and 4) and that the current problems get accentuated in the culture of science (see Chapter 5). I also showed that most research on research integrity focuses on the exposing the problem but that very little research targets the solution (Chapter 1). Commented slides are available at https://osf.io/m796s.
>
> **Personal reflection:** This conference mainly involved publishers, editors, and librarians. The reception of my open criticism of the system, and the questions raised in the discussion panel opened my eyes to a new perspective on the topic. Publishers are, in the end, businesses, and their openness to change was very encouraging. But interestingly, many attendees were also researchers who changed career, and many changed career because of the problems and unrealistic demands of the system. The discussions I had at coffee breaks were very rich and personal, and made this conference one of the most insightful of my PhD. It also made me realise that the target audience of our findings are not only researchers specialising on research integrity, but all research actors.

Are success indicators threatening integrity? Results from interviews and focus groups with diverse research actors. Poster. *6th World Conference on Research Integrity*, Hong Kong (HK), 2–5 June 2019.

> **Description:** In this poster, I introduced a broad overview of the themes and ideas captured in our interviews and focus group (Chapters 3, 4, and 5). In particular, I highlighted that current definitions of success are often linked with problems which are thought to threaten integrity, and that, although most participants agreed that some changes need to happen in how we scientific success is assessed, no one feels able to initiate the change on their own. An electronic version of our poster is available at https://re-sinc.wixsite.com/project/posterwcri6.
>
> **Personal reflection:** This poster presentation was rewarding since it allowed me to have in depth discussion with experts and young researchers who do research which is very similar to ours on the international scene.

We need to talk: why improving inter-actor communication may be crucial to advance research assessments. Oral presentation and panel discussion. *EUA (European Universities Association) Workshop on Research Assessment in the Transition to Open Science*, Brussels (BE), 14 May 2019.

> **Description:** Short presentation on our findings to argue that that a lack of communication between research actors may lead to inaction, distrust and frustration, and may paralyse the changes needed to improve research assessments. The slides are available at https://osf.io/kyuth.

Personal reflection: This presentation was very daunting. Most of the audience were university leaders and directors, and my presentation was directly targeting the inadequacies of current assessments, most of which are articulated by research institutions. As a young researcher representative, I also participated in a panel at the end of the conference where I discussed my personal views of being powerless in a system which forces us to perform in a way we disagree with. To my surprise, the short talk was received very well, and the panel was truly inquisitive and open to our different perspectives. A number of leaders from all over Europe even came to speak to me after the conference to know more about the project and to ask how they can make things better. This honest openness for change made me very eager to continue our work.

Let's Get Back to Earth! Why Researchers Are—and Must Be Seen As—Normal Human Beings. Oral presentation. *ASBH (American Society for Bioethics and Humanities) Annual Meeting: The Future is Now* – Anaheim, CA (USA), 18–21 October 2018.

Description: In this presentation, I discussed preliminary findings from our qualitative works (especially Chapter 5) and showed that we hold an unrealistic vision of researchers as individuals that are 'beyond' human beings. I explained that endorsing an unrealistic view of researchers reduces critical thinking, creates expectations which may harm researchers and submit them to mental strain, and stigmatizes uncertainty and failure.

Personal reflection: The ASBH is largely composed of philosophers, clinical ethicists, and researchers in bioethics. I realised that my presentation was outside the expertise of most, but it raised interest with PhD students in the audience. Instead of appealing to their expertise, I felt that I was appealing to their personal experience as researchers. The questions were very personal and interesting since they also allowed me to see differences in how research assessments are made between medical sciences (i.e., what my thesis focused on) and humanities (i.e., where most of the attendees came from).

Tackling hierarchies in academia: a proposal for promoting integrity in research. Oral presentation. *EACME (European Association of Centers of Medical Ethics) Annual Conference: Ethics in action* – Amsterdam (NL), 6–8 September 2018.

Description: In this presentation, I discussed preliminary findings from our qualitative works to highlight that hierarchies present in the current research system have a damaging effect on whistleblowing, open criticism, and responsibilities for integrity.

Personal reflection: My presentation was somewhat outside of the topic covered in the session in which I presented, and in a way, I worried that my topic would be of little interest to the small audience. Interestingly, patient representatives and co-design participants expressed vivid interest in the idea of hierarchy in research and reached out to discuss this concept further after the conference. The discussion also allowed me to understand that Belgium is perceived to have a much higher hierarchy and competition culture than other countries such as the Netherlands, where most participants came from. This is a point which is important to keep in mind when discussing hierarchies in international audiences.

Research on research integrity: publishing patterns, trends, and impact… a focus on Editors. Invited oral presentation. *EASE (European Association of Science Editors) Conference*, Bucharest (RO), 7-10 June 2018.

Description: I presented the results of our literature analysis about research on research integrity, focusing on the role of editors in the discourse. I also provided a short introduction for the qualitative interviews and focus groups planned, explaining how it can serve as a platform to hear the views of editors on topics of integrity.

Personal reflection: Beyond the presentation, the discussions that this event allowed me to have made me discover a new face to the integrity discourse: the open science movement. A lot of the talks addressed open science, open peer-review, open data, and hearing the struggle that editors face in such advances was quite enlightening. As researchers, we tend to criticize editors and publishers without really considering their side of the struggle.

Success, integrity, and cultures in academia: Voices of Belgian researchers, research students, and other key research actors. Oral presentation. *PRINTEGER European Conference on Research Integrity*, Bonn (DE), 5-7 February 2018.

Description: In this presentation, I highlighted preliminary findings and gathered insights from the expert audience on how to maximize the value of our methods.

Personal reflection: This conference was largely constituted of experts on research integrity. It was a very rewarding experience as I was able to chat with some influential actors who are behind big decisions in the field, but was also limiting in terms of personal impact for our work.

Research on research integrity: publishing patterns, trends, and impact. Oral presentation. *5th World Conference on Research Integrity*, Amsterdam (NL), 28-31 May 2017.

Description: I presented the preliminary results of our literature analysis about research on research integrity (Chapter 1).

Personal reflection: This conference was the first to target findings from my PhD project. The crowd was largely composed of experts and researchers doing research on research integrity. One thing I remember from the questions of the audience was the overarching question 'Who finances research on research integrity'. There was a call and a need for more research on the topic. The conference in itself was very rewarding since it kept me up-to-date on the latest advances in the field. I realised however that being experts among experts also limited our reach to non-researcher research actors, and I kept this thought to guide my choice of audiences later in my project.

OUTREACH

It was also important for me to reach out to the community, both of scientists, and of non-scientists. I plan on fostering more outreach activities once more chapters are published, but I also undertook some small activities during the PhD project. Some initiatives have already taken place, but we have plans for further communications later when more sections of the project will be made public.

TrueTuesday series

Together with Martijn Peters, I issued a weekly Instagram series entitled TrueTuesday through the 2018-2019 academic year. In this small series, I conducted mini-interviews with experts on topics of research integrity and research culture. The rationale behind this series was the realisation that there was still a general lack of awareness on new initiative taking place in research integrity. The series also meant to raise awareness and acceptation of some topics which are often kept untold, such as mistakes and failure, researchers' mental health, or the debated impact metrics.

OASPA webinar

In December 2019, I presented a summary of our findings in an online webinar from the Open Access Scholarly Publishers Association (OASPA) on Early Career Researcher's take on Academic Culture and Openness. The seminar attracted individuals from many different expertise. The recording and summary of the webinar are available at https://oaspa.org/oaspa-webinar-phd-students-take-on-openness-and-academic-culture-webinar-key-takeaways/.

Max Planck PhDnet Offspring Magazine podcast

In the summer of 2020, I was also approached by members of Max Planck PhDnet to discuss my PhD project in a podcast. Nikolai Hörmann and Srinath Ramkumar, both enthusiastic and ambitious PhD students, organised the podcast. I was amazed at their professionalism and knowledge of the topic, and felt honoured I could take part in this activity with them. The episode is available on their podcast channel, alongside many other episodes that present fascinating topics and excellent speakers.

Appendices

APPENDIX 1

BUILDING THE CLASSIFICATIONS

To build the classifications for our research, we used an inductive process (Elo & Kyngas, 2008) based on the findings from the first set of papers retrieved (i.e., the SCOPUS search). An inductive process means that we started with the general idea of describing research, and that we decided on which categories and classification options we should include based on what we found in the abstracts and papers assessed. In an inductive process, categories are highly mobile in the beginning of the analysis, but as we carry on the analysis, recurrent themes and groups of themes slowly solidify the nodes that we choose to look into. As the analysis advances, the process becomes more and more deductive, with new nodes being created when new information is not covered in the past nodes. At the end of the process, the nodes are looked at all together, and through axial coding and connection of concepts, merged in fewer nodes. For our particular work, we proceeded as follows:

We started with the broad, general idea of describing the literature. Initially, we had a column for topic, and a column for whether the study was empirical or not, as these two points were interesting to us from the onset. We classified papers in these two columns, and, when the study was empirical, we wrote, in another column, one or two sentences to describe what the study was about. We rapidly realised that a series of aspects described in the papers were recurrent, namely, the topic, the methodology, the population, etc. We then found that the same methodology (e.g., content and textual analysis) could be used with different populations (e.g., researchers vs. editors), or with a different data source (e.g., retractions, notes, or bibliometric data vs. guidelines, policies, or university requirements). We thus added the 'studied population' and the 'source of data' categories, and added classification options as we continued to review the literature. We also noted that, in the column in which we described the studies, we hinted on the 'objective' of the researchers (e.g., assessing an approach, describing, exploring, or quantifying a problem, etc.), so we added this category. Realising a trend in the 'objective' categories, we then decided to group those objectives in the 'focus', to express whether the study was describing the causes, the problem, the approaches to the problem, or the consequences of the problem. This grouping allowed a better overview of the research that had been done.

In this process, NAB first read and classified the papers. After one first ‘round’ of classification, she met with WP and discussed the (then very numerous) nodes she extracted. Together, NAB and WP classified these nodes into broader nodes in order to have fewer, more encompassing classifications. NAB then looked back at the specific abstracts and papers to make sure that the new broader classifications grasped the specific differences accurately. In this process, she sometimes had to create new classifications, or to adapt the wording of the classifications to allow for ambiguous papers to fit in. NAB and WP met again, and repeated this process, until both were satisfied with the overview and simplicity that the categories allowed while retaining enough details to yield accurate knowledge. This process was key to building a comprehensive categorisation of articles. For instance, NAB initially had over 160 different ‘topics of interest’. After repeatedly sitting down with WP to look at these topic while investigating for more details in the abstracts and the papers, they managed to reduce these topics in the final 17 topic classifications.

It is important to note that inductive coding is, by definition, dependent on the reviewers. In other words, different reviewers may look at the same data and build different categories, but once the categories are built, deductive analysis (i.e., placing the articles in the respective categories) should yield similar results regardless of the coders. Unfortunately, since this project was only a first step to a bigger project, only NAB went through the full coding process. Nevertheless, our goal being to build simple groups which could help us make sense of the broad diversity of the literature on research integrity, we believe that this method served our purpose.

Elo, S., & Kyngas, H. (2008). The qualitative content analysis process. *J Adv Nurs, 62*(1), 107-115. doi:10.1111/j.1365-2648.2007.04569.x

APPENDIX 2

COREQ CHECKLIST

(COnsolidated criteria for REporting Qualitative research)

Topic	Item No.	Guide Questions/Description	Reported on Page No.
Domain 1: Research team and reflexivity			
Personal characteristics			
Interviewer/facilitator	1	Which author/s conducted the interview or focus group?	on page 276
Credentials	2	What were the researcher's credentials? E.g. PhD, MD	on page 276
Occupation	3	What was their occupation at the time of the study?	on page 276
Gender	4	Was the researcher male or female?	on page 276
Experience and training	5	What experience or training did the researcher have?	on page 276
Relationship with participants			
Relationship established	6	Was a relationship established prior to study commencement?	on page 276
Participant knowledge of the interviewer	7	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	on page 276
Interviewer characteristics	8	What characteristics were reported about the inter viewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	on page 277
Domain 2: Study design			
Theoretical framework			
Methodological orientation and theory	9	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	on page 100
Participant selection			
Sampling	10	How were participants selected? e.g. purposive, convenience, consecutive, snowball	on page 98
Method of approach	11	How were participants approached? e.g. face-to-face, telephone, mail, email	on page 277
Sample size	12	How many participants were in the study?	on page 95
Non-participation	13	How many people refused to participate or dropped out? Reasons?	—

Setting			
Setting of data collection	14	Where was the data collected? e.g. home, clinic, workplace	on page 277
Presence of non-participants	15	Was anyone else present besides the participants and researchers?	on page 276
Description of sample	16	What are the important characteristics of the sample? e.g. demographic data, date	on page 95
Data collection			
Interview guide	17	Were questions, prompts, guides provided by the authors? Was it pilot tested?	on page 279
Repeat interviews	18	Were repeat inter views carried out? If yes, how many?	on page 277
Audio/visual recording	19	Did the research use audio or visual recording to collect the data?	on page 277
Field notes	20	Were field notes made during and/or after the inter view or focus group?	on page 277
Duration	21	What was the duration of the inter views or focus group?	on page 277
Data saturation	22	Was data saturation discussed?	—
Transcripts returned	23	Were transcripts returned to participants for comment and/or correction?	on page 277
Domain 3: analysis and findings			
Data analysis			
Number of data coders	24	How many data coders coded the data?	on page 100
Description of the coding tree	25	Did authors provide a description of the coding tree?	on page 96 and 136
Derivation of themes	26	Were themes identified in advance or derived from the data?	on page 100
Software	27	What software, if applicable, was used to manage the data?	on page 100
Participant checking	28	Did participants provide feedback on the findings?	on page 277
Reporting			
Quotations presented	29	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	Throughout
Data and findings consistent	30	Was there consistency between the data presented and the findings?	Throughout
Clarity of major themes	31	Were major themes clearly presented in the findings?	on page 96 and 136
Clarity of minor themes	32	Is there a description of diverse cases or discussion of minor themes?	Throughout

APPENDIX 3

RESEARCH TEAM AND REFLEXIVITY

In accordance with the COnsolidated criteria for REporting Qualitative research checklist (COREQ; Appendix 2), and in respect of transparency, we found important to provide further characteristics about the setting and the interviewer at the time of the study.

Besides one early interview with an institution leader in which WP, assistant professor, attended to provide feedback about the interview, all other interviews and focus groups were conducted by NAB, with no additional non-participant or assistant.

NAB is a female PhD student in the Faculty of Medicine and Life Science of Hasselt University, Belgium, with a background in cognitive neuroscience and bioethics. Coming from Canada, NAB had the advantage of bringing a certain neutrality in the interviews by not being strongly affiliated with one or another Flemish region, and by not corresponding to an established research group.

Before conducting the interviews and focus groups, NAB followed courses about developing interview questions, conducting focus groups, and analysing qualitative data offered from Flemish universities and from the Flanders' Training Network for Methodology and Statistics (FLAMES). In addition, she used the resource books from the Focus Group Kit by Richard A. Krueger and David L. Morgan (Morgan & Krueger, 1998), and discussed with RDV — expert in qualitative inquiries and part of the team that built the original guide upon which we inspired ours — to gain insight on building, conducting, and analysing focus groups and interviews.

Besides a few exceptions, NAB had no prior relationship with most participants, and the first contacts were established with the invitation email. No repeat interviews were carried out. Before the interview, NAB described the project briefly and explained the purpose of the interview informally. On some occasions where interviewees were anxious to know more about the project in advance, NAB would email the main themes targeted, but would not share the interview guide with participants by fear that this may lead to rote answers.

Bias and assumptions

NAB holds the view that research integrity is largely determined by the research system, and the interview guide was necessarily not unbiased to this perspective. Nonetheless, if participants shared a different view (e.g., if they believed that integrity was solely a matter of personality), NAB was careful not to contradict or bias interviewees' ideas towards her perspective. In re-reading quotes with the research team, we were careful for possible misinterpretations, and when quotes were interpreted differently by WP or RDV, we adapted the nodes and interpretations to make sure they fit the words of the participants. Both WP and RDV helped in classifying the main nodes into general categories of *Who*, *What*, *How*, and *Luck*. Initially, we were tempted to classify these four categories in *Products* of success (the What) and *Potential* for achieving success (the Who, How, and luck). However, after several discussions, we realized that doing so may reinforce the perspective that products are the ones which truly indicate success, while potentials are simply increasing the chance of yielding better products. As we describe in our extended findings, many of our interviewees considered the *Who* and especially the *How* to be genuine successes in themselves. In this regard, we intentionally kept the four categories together as each representing successes in themselves.

Study design and interview/focus group setting

Interviews and focus groups were conducted in private meeting rooms or offices or, according to preference, in public spaces (n=2) or through video call (n=3). One of the interview conducted through video call had some sound and connection problems, but the other video calls went very smoothly.

Interviews lasted on average 60 minutes, depending on the time granted by the interviewee (range from 34 to 80 minutes). Focus groups lasted around 120 minutes each and included a five-minute break.

All interviews were audio recorded and transcribed verbatim by the interviewer (NAB) or a university-approved transcription service. Transcripts were not returned to participants except in select cases where participants expressed a wish to monitor their answers, and in cases where the quotes of interest might have jeopardized the confidentiality of participants. No repeat interviews were undertaken. After most interviews, the interviewer filled a self-questionnaire about the interview to note any abnormalities and general feelings of the interview data. We did not use these questionnaires further.

Reference

Morgan, D. L., & Krueger, R. A. (1998). London: Sage Publications.

APPENDIX 4

General interview guide

A part of my research is to explore the views of different actors that contribute to the research system.

To protect your privacy, I want to avoid disclosing your specific job title and to place you in one of bigger category of research actors. I may use a higher level of details to describe the type of participants included in each category, but I won't link direct quotes with company or institution names.

*I placed you in the category ***actor group****. *Does that sound good to you?*

Introduction and information on respondent's career

1. Before anything, I would like you to **describe your work** to me, in your own words.

    *Prompt: In a broader perspective, what would you say is your **role is in the scientific system**?*

2. In this job, you obviously **care for scientific excellence**. How would you say you fulfil this goal in your work?

    We will get back to this a bit later. For now, I will change topic and I want us to talk about success as this is an important topic that we are trying to understand in the project.

Success in science

3. First, try to think about scientists you've known that were very successful. What do you think made these scientists **successful**?

    Prompt: Which characteristics do you think are most important to advance a researcher's career?

4. Do you feel like these characteristics are **captured in current research assessments** and evaluations? In which ways?

5. (If time allows) What do you feel that your actor group **should do to promote successful science**? Do you see that happening?

Tensions or conflict between success and integrity

6. You mentioned that X, Y, Z are criterions that indicate success in research. Do you think that these are **also indicators of quality**? Sound research?

   Prompt: Which criterions do you think ***indicate the quality*** *of the research?*
   Prompt: Which criterions do you feel are ***not suited*** *to indicate the quality of the research? Explain.*

7. Does it happen that you see excellent researchers but for some reason these researchers **don't succeed in getting ahead** with their career? Can you give me some examples?
8. Do you feel that the way in which success is attributed allows to for **emerging scientists** to become successful?

Current problems

Let's change the topic now; leave aside success for a bit and look at when science is not at its best. Like I said, I am not here to denounce or condemn cases, so I will make sure to protect the confidentiality of cases you may discuss.

9. Have you ever had to **deal with science** which you considered was **not really in line with the rules of science**? What happened?
10. Can you give me precise examples of the elements that you consider **signs of bad or sloppy research**? What are **red flags**?

Motives for bad practices

11. **Why** do you think bad research practice happens?
12. Do you think **anyone could** end up in such a situation or only types of people?

Responsibilities towards integrity

We have already discussed how to promote successful science, now I would like to gather your thoughts on how to prevent sloppy research.

13. What do you think should be done to **prevent bad science** from happening?
14. **Who** should take the lead to make these changes happen? Who else should be involved?
15. What do you consider is the **responsibility of your *actor group*** to protect integrity?
16. Where does your responsibility **end**?

One change

Finally, if you could **pick one important change** that needs priority right now in how research works, what would it be? How do you think this change could be done?

(if time allows) Alternatives

17. If there were no rules for evaluating scientists, and you could **start from scratch**, what would you like to look at when assessing scientists?

    *Prompt: What are the characteristics that **YOU think are most important** for researchers to do good research?*

APPENDIX 5

GENERAL FOCUS GROUP GUIDE

Intro:

1. Who you are
2. What is your area of research
3. Describe a typical day of work
4. What's your favourite ice cream flavor

Scientific career

Before starting, I would like to know a little more about your career as a researcher.

18. Specifically, I would like to know what is it that makes your work so great? What do you feel is most satisfying, most rewarding about your career?

    Prompt: When people ask you why you chose to be a researcher, what first comes to mind?

We will get back to this a bit later. But for now I will change topic and I want us to talk about success.

Success in science

19. Think of a person in your field who you think is **very successful**. (It doesn't have to be one person in particular, it can just be some characteristics of many different people, can be yourself in 20 years…) How do researchers **become successful**? What, in your view makes this person a success?

    Prompt: What are the most important factors for advancing in your career?
    Prompt: What are the funders and the employers looking at?

20. Now imagine that I am a newcomer in your field and I ask you what I must do to stay on the top, **what would you tell me**?

So you say that successful scientists are generally scientists who do X, Y, Z.

21. Do these successful scientists **reflect or mirror the kind of scientist you want to be**? Do you have such aspirations for success?

Tensions or conflict between success and integrity

22. As we discussed, you point out that funders and employers look at X, Y, Z… Do you think that **these criterions for success indicate outstanding or excellent research** (e.g., appropriate methods, relevant topic, high quality work)?

    Prompt: Which criterions do you think ***indicate the quality*** *of the research?*
    Prompt: Which criterions do you feel are ***not suited to indicate the quality*** *of the research? Explain.*

23. Now try to think of a **colleague** who, in your opinion, **does good research but cannot reach success in science**?

    Prompt: What in your opinion explains that this researcher cannot reach a successful career?

24. **What would you say to this researcher to help him/her get ahead**?

Current problems

Let's change back the topic now; leave success aside for a bit and discuss what it is like to be a researcher. So you remember that at the beginning of the discussion, I asked you about the aspects of research that make you like your career. Now I want us to talk about the other side of things, about what **frustrates** you as a researcher.

25. So let's say I am a **newcomer** in your field. I just started working in your lab and I am not sure whether I should follow a scientific career. If I asked you what are the **most frustrating things about working in science**, what would you tell me?

    Prompt: what would you tell me are some of the biggest frustration I could encounter?

26. All right, so as we have discussed, being a researcher is not necessarily always easy. It can sometimes happen that things go really wrong. Have

you **ever seen or heard** of a situation in which you thought research was conducted in a way that was **against the 'rules' of science'**? What happened? What did/would you do?

Motives for bad practices

27. **Why** do you think researchers were acting in this way?
28. Do you think **any researcher could end up** in such a situation?

CURRENT VIEWS ON RESPONSIBILITY

29. **What** do you think should be done to **prevent bad science** from happening?
30. **Who should take the lead** to make these changes happen? Who else should be involved?
31. **What can you do**?

Prompt: Do you feel like you miss something to be able to change things yourself?

Solutions

To finish, I would like to ask a more concrete question.

32. Finally, if you could pick **one important change that needs priority** right now in the research system, in how science works, what would it be?

Prompt: *How do you think this change could be done?*

Personal success

Now before we finish, I want you to think back about the discussion we have had on success, and on criterions that are most often used to evaluate a research career. But now, I would like you to think about yourself as a researcher, and to think about your strength, about what makes you feel accomplished in your work. **What do you think is your biggest contribution to your work**, or **something that you think is key to be a good researcher**, regardless of the criterions we have said before. (For example, maybe you think that the fact that you brush your teeth after lunch is key to the success of your research team.)

I will not ask you to discuss it this time, but I would like everyone to take one of these little pieces of paper. On the piece of paper, I would like you to make up 3 to 5 criterions for funders and employers. I want you to think about what you consider your biggest contributions in your work, and to make up criterions you would think, if funders and employers evaluated, you would have better chances to succeed.

Summary

Is there anything you would like to mention that we failed to discuss today?

APPENDIX 6

SAMPLE QUOTES ON PUBLICATIONS

Sample quotes substantiating arguments against and for using publications as a main determinant for scientific success.

Argument	Sample quote	actor
Arguments against using publications as the main determinant of scientific success		
Reductionist	*It is a very flawed measure of success in a way. I mean… I don't want to give the impression that… of discouraging any of these successes, you know, I mean publishing very important papers in very selective journals is an achievement, that is very clear. But I think that there are other very important contributions to the scientific enterprise which don't necessarily translate into one of these unit of credit of success, which is a first author publication in a very prestigious journal. And I think that currently we collectively, as a community, do not do enough to actually support and reward these kinds of contributions that are very important for the scientific enterprise*	EP
Arbitrary	*Yeah but with publications it's sometimes also just having luck…[…] To me it's not always that you're a good researcher.*	LT
	It is wrong to think that... […] having more publications, it means you're better and better and better, I think it's a very wrong way of thinking.	PMI
	I have less and less confidence in publishing with the fact that 'who is going to be the reviewer?' 'Is he biased?' 'Is it the journal?'	R
	The highest journal [of my field] it's all already fixed before with companies, pharmaceutical companies, who will get published their RCTs, it's already all set in advance...	PhD
Perverse	*They do a lot of experiments just to publish. Just to make an article, because they have to have an article before the four years are done. So they do their experiments in function of an article*	LT
	It's my only drive for some things, that it's just publication.	Res.
Arguments in favour of using publications as the main determinant of scientific success		
Representative	*So people say, you know, publications don't matter, but at the end of the day there clearly is a link. If you end up publishing in a good journal, then you probably started off with a very good research question, and you probably are a very good researcher. They are not 100% linked, but I'm sure there is a link there.*	RIL
Measurable	*"It's the career, it's the way you get the career, it's the number of publications that will count, the number of promotions of PhD theses will count, but for me that's not the most important. I think a researcher who is not... who is publishing (they need to publish of course) but let's say only two A1 publication, or one publication a year, but in the meantime is contaminating other researchers, helping other researchers and is multidisciplinary... That's more valuable for me as a person. But in the academic world, I cannot value that directly. I'm not in a position that I can say "You are the very best researcher, so I promote you to full professor from associate professor". Because there we still have the numbers that count. And ok, that's the way it is, and that's the whole issue nowadays with researchers. They really get troubled with these numbers."*	RIL
	"I think it would also be a bit difficult to really value a PhD or the PhD project without publications. Because how do you determine that someone has done their best, but unfortunately didn't get any publications."	PhD
	"I do believe that you have to have some evidence about the process you have made, and the path that you've walked throughout your doctoral thesis. That's why I find it quite normal that you have to have a certain amount of publications in the procedure… "	RIO
Necessary	*"I think I'm going to be the boring one, but I think it is important to have publications and to also be successful in some research grounds every now and then because I feel like it's my... That's what is expected from me, but that's also how you can make the research... you can keep the research going. I think it's one I see as my duty to publish the results, to share them so that others can build further on them and you yourself can build further on them."*	Res.
	"If you don't have the publications you're not noticed. And if you're not noticed, your research might be extremely interesting, but if it's not read, if it's not noticed, what's the value."	RCC

Note: Researcher is abbreviated to Res. Colours added for polarising arguments.

APPENDIX 7

SAMPLE QUOTES 'WISH FOR CHANGE'

Sample quotes from the 'wish for change' which relate to changes in the ways success is defined and assessed.

actor	Sample quote
CHANGES TO RESEARCH ASSESSMENTS	
Value quality over quantity	
RIO	*I will then insist that the money is spent on projects of high quality. Quality of the research.*
Res.	(immediate response) *Take the output pressure away! So you can have more room for quality.*
PMI	Participant: *My wish is that scientific outcomes, papers, pieces, news, are assessed by their intrinsic value, intrinsic scholarly value and not by indirect measures as it is the case right now. Journal impact factors, citation index, et cetera, these are all proxies.* Interviewer: *What would you say is any intrinsic value?* Participant: *I think open peer review would solve this problem.*
Reduce output pressure and competition	
EP	*I would like to see a world where the pressure is off the researchers, you know, not... there are not pressured, in the world that they can do their research without pressure of publishing in high impact journals, and like to see that there is no impact factor anymore at least not in such a way that there is usually considered today. And that to bring more joy in their life, essentially, because I think that they are so stressed out, and they are always chasing some next step in their career advancement, and they forgot that the science is actually fun thing to do, you know, it can be a way of, you know, living a life, not just working as a hamster in a wheel, you know, just yeah, chasing your own tail or something like that.*
EP	*(Laughs!) It is a really tough one. Because I don't see… Do you know Merton's model? […] OK. There is tension, there is obvious tension between the kudos, and the whole system that has been put in place where it's… you have to be special. It doesn't fit. It doesn't fit with the kudos! It doesn't fit with the universalism, etc. So I think that that's where something is wrong. I don't have the solution, but that's what needs to be addressed! [...] I would try to solve that tension that exists right there, to be able to go back to the other communalism, to the universalism etc. You know, the kudos itself.*
EP	*Change the reward system! (Laughs) Change the reward system. Completely. Because would then allow everyone (A) to publish wherever it's really most relevant, it's not linked to the impact factor any longer… You know if people did that, what I said, and this was not relevant, impact factor was not relevant, and it's really truly about what kind of research career have I had and what research have I done, you know, that is really important, and how does this impact in my field. Then I think everything would change. And, yeah, that would be my biggest wish, and I'm working towards that.*
Broaden and adapt indicators	
PhD	*Maybe the cumulative impact factor that they just need to do it really field per field, and not faculty per faculty*
RIN	*I think it would be broadening of the criteria for recruitment, promotion, funding. I think if we could really get everybody behind that, it would have a huge impact I think.*
PhD	*...maybe looking at PhD as a career. Because now you have only one main outcome, the publications, but in a career you have a lot of competencies that are important.*
EP	*If I have a magic wand, I think I would want to get rid of the Impact Factor in research assessment. And getting rid of… You know changing this problem that we started this conversation with. Which is that it's only publication in a certain amount… in a small number of very selective journals that is considered a measure of success. So, you know, I would want a magic number that represents all these other things and that's probably, that's completely unrealistic, but I would want at a minimum the research assessment framework to change to move away from that single dominant measure that is being used at the moment. To appreciate different kinds of contributions much more effectively.*

actor	Sample quote
CHANGES TO RESEARCH FUNDING	
Fairness in evaluation	
PhD	Participant A*: I think that... I feel that there is a bias that certain groups will always get funding, and smaller universities [...] are really struggling to get like an FWO project funded. So maybe there should be some regulations about it. But I don't know how... maybe restricting the number of projects that the group can apply for. I think they already have some regulations for that...* *Participant B: I think so too.* *[...]* Participant C: *Anonymous selection... (laughs)* Participant A*: That would be maybe the best.* Participant B*: But that is also important, because a lot of professors or researchers knew people that are the judges, and they have like, the privilege, and will get funding. So anonymous would be better.* Participant A*: But then maybe you lose the advantage of collaboration. You can't say "but we'll be collaborating".*
Fairness in distribution	
EP	*Yes, my fairy wish would be a change of the grant system, and I'm saying of that it… in my… so say 80 percent of the money might be divided according to the prevailing system, for the proposals and give 80 percent to the best proposals, and then we have a pile of proposals that are rejected, and make it a lottery, for 20 percent.*
RCC	*Participant: More freedom. Less bullshit. More... and maybe the money should be just divided equally or something like that, which is also not really realistic because then the amount would probably be so small that you still cannot do anything. But at least then everybody can not do anything, instead of being, you know, when you have this big in house thing and here is the people that get a lot of money and get a lot of stuff done, and here's me [laughs]. At least we all will be...* *Interviewer: Everyone would be fair.* *Participant: Yeah, it's just not fair. That's it, it's not fair. And if you... I can completely understand why big science people don't go to [small university] because you kill your career if you [go there].*
Long-term and baseline funding to increase security	
RIL	*I think it's the research funding, but I don't... I just know it should change, but I don't have the answer for you. I think a researcher should not have these short term financing situations. I think that's probably the worst perverse incentive you can give a scientist. I think you should have a Tenure track where you require that a scientist proves him or herself, but once you have an established scientist, they should have some sort of basic funding which could be adjusted based on how they perform, but it should not be this 'yes/no' thing on a four year term which is what most grants are. Because I really need to deliver in four years, and that gives me perverse incentives.*
RCC	*But maybe it might be interesting to give people different kinds of contracts. To don't give always these short-term contracts, but give people longer term contracts. But I know that there's a discussion. I know a lot of people say 'well I give the best of myself because I have a short-term contract and the edge is on… I don't know whether the edge should be so strong. I don't know whether the competition should be so strong. I don't know whether that's really helpful. If you really want to achieve trust and if you really want to achieve openness to mistakes, people should feel secure enough to do it. And I think one of the answers is 'you will not lose your job'. So… Maybe job security might be an answer. (RCC)*
FA	*Hm hm. Well exactly what I said form the start. I think that we should have a very close look at the way we are funding institutions for doing their research. I think this is key […] but there are some elements I recognise, and we recognise, that are worth a good discussion. And one of these elements is that indeed apart from competitive funding, which is important because competition, and what we are doing here can make for good quality research, excellent research, and apart from this competitive funding, you also need some sort of basic funding to give people a chance to start and to launch their career as an academic. Also to do some things that are less fashionable, because also research has its fashions, less fashionable, or less appealing to evaluators at the moment, with which you can prove after a while that there is something in it and then you become stronger to an evaluation panel. So I think that reconsidering the way you are funding research institutions is also letting some pressure, or diminishing some pressure on institutions like us. I think you get better competition, by also making it less stringent. Maybe this sounds as a paradox, but I don't think it is.*
Res	*Start-up money? For creative plans which are not judged from the beginning?*

Note: Researcher is abbreviated to Res.

APPENDIX 8

SELECT QUOTES TO ILLUSTRATE THE PROBLEMS BEHIND MISCONDUCT AND 'BAD PRACTICES'

Sample quote	actor
CHANGING THE COURSE OF SCIENCE	
I suspect that those tiny [fraudulent] publications, they will get falsified, but nobody puts too much attention to that anyway, **they are not moving the real direction of science, you know, this is such a small study somewhere, it doesn't really matter that much** if it turns out to be incorrect. [...] Cutting corners sometimes is not changing the course of science. [...] So it can happen to everybody that you cut corners. It's just, is that a scientific corner or is that an esthetic thing? But everybody will be tempted at some stage to do that so let's just hope that the majority of people will have the scientific integrity not to do that. But... (sighs) I think, certainly with PhD students I only hear transiently... But I don't think you can get a full article based on one little corner that has been cut. I think you really need to cross that line, and then go up and not do that again. I think that's where things go wrong when it becomes a standard, where you've made it once and realised it doesn't make a difference...	RIL
CHANGING CONCLUSIONS	
In my lab if I look, the only misconduct I've picked up was just stupidity. PhD students who scanned a little too short and had to go back to the scanner and thought 'I could just copy-paste the bottom bit because there's nothing on it anyway'. **That's real misconduct, but at the same time, that's not scientific fraud. Well it was, it is scientific fraud, but he was not changing a conclusion**, he was just too lazy to scan a really nice experiment [...] What I consider cheating is that you leave out the data that don't suit your model. Or you make up data to get your model correctly. That is what I call cheating.	RIL
It's difficult to prove intention, and **for us that's not that important. If it's actually a deception, doesn't matter if it's intentional or not**. Then we need to correct the literature if it's published. So you know, again I think a lot of the misconduct investigation that institutions do, they put a lot of emphasis on the intentional bit because that's part of the employer status and so on. Whereas we as journals are not that interested in that part. **We're really interested in 'Is this research trustworthy or not?**	EP
Intent is something that editors are not in a position to properly evaluate. [...] And this is where they have to rely on institutions to determine, to really establish and ascertain whether there is misconduct or not. Where the responsibility of the editor is in **correcting the scientific record**. And that, it doesn't matter whether it's misconduct or not in a way. [...] **if something is wrong**, and you're unsure as an editor **whether it's misconduct or not, it doesn't matter**. It also needs to be corrected in the scientific record.	EP
I think usually that's something that is important in evaluating those cases. It's really, does the act, the problem that you have identified, **does that actually lead to a different... To change the nature of the conclusion?** And really make the data say something different than what it says? And so… You know that tends to be misconduct.	EP

POOR QUALITY OF FINDINGS	
You always have to take into account the rules, the procedures of good research.[...] If you want to talk about society, your reference group must be big enough so it can be a reference, a real reference to the society. So if you don't put up the research in advance in a good way, then **it's also sloppy or bad research. Because all the results that come out of this, even if they're positive, they won't be representative for the reality.** And that can lead to – if it's research in medicine– that can be very dramatic. So that's sloppy research.	PMI
MISUSE OF RESEARCH MONEY	
[During a discussion about copying deliverables for a different research grant] Yeah, well, for us it's fraud directly. Because **you do it in order to win money**.	PMI
INTENTIONAL	
For me a bad scientist is someone who actually intentionally knows that what he or she is doing is wrong and might inflect the public opinion upon a publication. […] So for me, being a bad scientist **actually means intentionally being bad**. You can perform a research without being aware [of the rules] [...] That doesn't make you a bad scientist or a bad researcher.	RIO
If someone warned you that this is not okay, that it is bad science, and **you still continue**, then, yeah, it's also not okay.	RIO
There is big misconduct or minor misconduct, if we can put it like that. It's like when you **consciously know** you are really doing some change in your results to make them look beautiful and then get this publication in nature, for whatever reason, or when you are just tweaking here and there and the supervisors is telling you everybody does that, so you are able to do that [laughs]. There's **different degrees of seriousness** also in here.	PMI
Well if it's **willingly** then it's a… It's a border you don't cross.	PMI
MORAL MISMATCH	
[Sometimes researchers say] 'yeah but it didn't change the main results of my article, so what's the problem?' [...] OK **if the results are being the same, that's not the issue actually**, it should be the process also. And at that point you see that there is this **moral mismatch**.	RIO

Note: Bold added for emphasis

APPENDIX 9

RANGE OF RESPONSES TO THE QUESTION '*WHY MISCONDUCT AND QUESTIONABLE RESEARCH PRACTICES HAPPEN*'

The number of quotes and interview should be considered with caution since they depend on the capture and the coding, yet we found interesting to show it as a rough estimate of the coverage of select themes.

Topic	# of quotes	# of interviews		actor
Pressure	29	22	...pressure for career, reputation is playing a role, competition is playing a role... that's all external.	RIL
			I can imagine that if you're in a situation where you're forced to have a certain outcome and your reputation depends on it – and I find it very shameful that there are, or there might be situations like that – that as a researcher you try to bend the truth in your favour. [...] I do believe that there can be situations where stress might force you to go into a direction that you wouldn't have walked in normally.	RIO
			The pressure is huge. You know, basically your career depends on it. And so when you are in a system in which basically your next paycheck, your next grant will be dependent on the results, I've never thought about it, but I can imagine that you will have some people think 'Well… Why don't just, you know, make it up?'.	PE
			They need their numbers of publications. [...] Pressure, yeah.	PMI
			…the more mundane reason I think has a lot to do with time and publication pressure and the pressure for funds, fundraising and things like that. Which puts an enormous pressure on people to produce results, to publish results…	RF
Ego and personal morals	23	12	Internal it's idolness, the personal 'I want to be a big researcher', you always should be modest as a researcher I think. So the lack of modesty is the internal factor that makes people just improving their data a little bit., just adding something here and there, deleting something here and there…	RIL
			I think that there is the egos, and the egos in science is still underestimated I think. yeah... [...] if you look at it when professors are doing it [i.e., committing misconduct], then it has a lot more to do with status, image, ego, trying to score, get off easy, yeah... I see that a lot more than it is of the pressure issue. [...] it's about image and scoring.	RIO
			But the ones that really matter, that make it to the press, those are very often the leading universities and I think there the perverse incentives are big egos, big prizes, top publications... There, you know, if you manipulate your data, you become a big hero. And I think that's at the level of the promoter, and that's probably the biggest problem of them all.	RIL
			You know that when you start fabricating papers to have like two papers a year in Nature, you don't do this for policy reasons. [...] You do this because you want to be kind of a king or a god in your discipline, and that's… well maybe narcissistic, maybe psychopathic type of behaviors.	PMI
			I think that depends on the person. (laughs)…	PMI
Normalisation of smaller misbehaviours	11	10	Because they start with QRP, and it gets more and more and more, and then they cannot admit it anymore without seeing the consequences so they make the choice to do even more terrible fraud to cover up the rest.	RIL
			And also the fact that it's a slippery slope, in fact. And that when you start being a little sloppy about certain things, you can actually very easily drift into something that is much worse.	EP
			I think if you get away with small infringements then you get bolder and bolder each time and, you know that happens.	EP
			And when from the moment you do it, you do it again, and when you think your colleague does something you will do it again etc.	RCC

Perverse incentives	8	7	I think we have a lot of perverse incentives. […] But, when you are a student, and your four years are up, your promoter does not have money to pay you any longer, and there's nothing you have to publish, and there's one excel spreadsheet if you change a few numbers, will give you a publication, I think it's very tempting to do that.	RIL
			if you actually do experiment with small numbers of animals, you're going to have a much larger effect. […] These are the kinds of perverse incentives that we have in the system at the moment.	EP
Lack of awareness	7	5	But people come willingly and tell us "It's the first time that I hear about the codes. It's the first time that I hear about those things".	PMI
			I don't believe that there are researchers who intentionally perform bad research, I do believe that there might be some researchers who are not aware of what the better practice or the best practice might be. And if they are informed about the better and the best practices in specific research and they follow these practices and they adjust their work methods, that's not a bad scientist.	RIO
			What I think is important though, is that, again in my generation, we were not made aware enough of that. A bit like gender bias; until you're explained what gender bias is, until you're explained what research integrity is, and what misconduct is, I think you will have more flexibility towards it because you just don't know where the red line is. Once you explain to people where the red line is, they will know when they cross it. And I guess… I think, they will cross it less likely.	EP
Lack of control (low risk high gain)	5	5	People don't see the severity of [misbehaviors like gift authorship]. So then the gain is much bigger than the risk.	RIL
			...you know that indeed if you do that, you will get a top publication which is very beneficial for your career.	RIL
			While big industrial laboratories have standard operating procedures that are very expensive and standardised when it's not easy, in research institutions, we are missing them. [...] Sometimes this is missing. Right? So that can lead to sloppy research.	PMI
Unrealistic demands	4	2	But at the end of the day I think the motivations are fairly similar for sloppy research and misconduct. And I think, in my view, it's related to the first question we discussed, which was the question of incentives, and research assessments. [...] Success is measured by publishing in very selective journals that are looking for very important ground-breaking claims. You are incentivised to find these ground-breaking claims… And so you are incentivised to really get something that is extraordinary, and ground-breaking. And let's face it, all the research in biomedical research, is not ground-breaking and extraordinary. Most of it is not.	PE
			I think there are too many PhD students who are really forced into working 24/7, which definitely cannot be what it should be either…	RIL
Lack of openness to failure and negative findings	3	2	If a researcher does every step of the research process correctly, and after ten years, he only had negative results, will the management of the institution still fund him for the next five years without guarantee that there will be a positive result? Or will they say after ten years 'Now it's enough you stop. You can go and find another job.'? In the latter case if that's the case, then I think every researcher, theoretically, every researcher can be changing its data.	PMI
			And if we do not tolerate failure, and use failure as a motor to drive you to success, you foster misconduct and sloppy research. Most of my researchers would be very upset, but that's how it is!	PMI
Over-specialisation	2	2	...there's also this hyper-specialisation where people always go further and further and become more sophisticated within a paradigm, without even questioning the paradigm anymore.	RF
			In all research misconduct that have been analysed, there are usually three that are present in all. The researcher always knew better. He was under pressure. And he was in a research area that was very difficult to replicate.	PMI
Cultural background	2	2	These were two [foreign] PostDocs who faked some western blot data I think and I think that this was a very ambitious lab with a lot of pressure. [I know people from this nationality] so I dare to say that [they] have a slightly different opinion about rules. So they're more relaxed than [western Europeans] for example, and well... I think they didn't think it's that serious.	RIL
			...he/she doesn't mind plagiarising as long as he/she doesn't get caught. So yeah... That also had to do with his/her nationality. [...] I think [some cultures] have this mentality that it's almost, you honor somebody by plagiarising them. And they just want to get their diploma so they can do a post doc in America. And he already had a PostDoc lined up. So he was really annoyed that he now had to postpone his PostDoc by a couple of months...	RIL

APPENDIX 10

PRINTOUT OF THE SURVEY

Start of Block: Introduction

Welcome to our Survey for the project Re-SInC.

The survey aims to rethink research assessments and research careers. We want to know what **you** think.

You may find the full information for participation here.

To make this short, we thought we would give you three reasons for taking part in this survey!

- We need your input to know what really matters in research.
- It will only take 15-20 minutes of your time (we tested it with a few people, and it never took more than 16 minutes!).

You will help a very grateful PhD student graduate!

Note on ethics and privacy: This project has been approved by the Medical Ethics Committee of Hasselt University, protocol number CME2019/O35. Answers to the survey will be **fully confidential**, and **no identifiable information** (e.g., IP adresses, emails, etc.) is collected in the survey. The dataset, which contains **no identifiable information**, will be made public when findings are published.

Closing date: October 31st 2019
If possible, use a computer to complete this questionnaire.
If you only have a phone at hand, you should use the landscape mode.

End of Block: Introduction

Start of Block: Demographics

Before we start, we would like to get to know you!

You are a...

- o PhD Student
- o PostDoc / Non-tenure-track position
- o Tenure-track researcher / Professor
- o Tenured researcher / Full professor
- o I was a researcher in the past, but moved to another career
- o Other (specify) ____

Display This Question:
If Before we start, we would like to get to know you! You are a... = PhD Student
Or Before we start, we would like to get to know you! You are a... = PostDoc / Non-tenure-track position
Or Before we start, we would like to get to know you! You are a... = Tenure-track researcher / Professor
Or Before we start, we would like to get to know you! You are a... = Tenured researcher / Full professor
Or Before we start, we would like to get to know you! You are a... = Other (specify)

You have been in this position for ...

▼ less than one year ... 9 years or more

Display This Question:
If Before we start, we would like to get to know you! You are a... = I was a researcher in the past, but moved to another career

How long ago did you stop being a researcher?

▼ less than one year ... 9 years or more

Display This Question:
If Before we start, we would like to get to know you! You are a... = I was a researcher in the past, but moved to another career

As someone who left academia your opinion is also very important to us! **Please answer the rest of this survey thinking back at your time as a researcher.**

Example: Read '***Are you** affiliated with a Flemish University or scientific institute?*' as '***Were you** affiliated with a Flemish University or scientific institute?*'

Are you affiliated with a Flemish University or scientific institute?

- o Yes
- o No

Display This Question:
If Are you affiliated with a Flemish University or scientific institute? = Yes

Your main affiliation is at...

- o Hasselt University
- o VU Brussels
- o University of Antwerp
- o Ghent University
- o KU Leuven
- o IMEC
- o Institute of Tropical Medicine Antwerp
- o Other ____

Display This Question:
If Are you affiliated with a Flemish University or scientific institute? = No

Your main affiliation is at:

- Institution ____
- Country ____

You are in the Faculty of...

- Medicine / Medicine and Health Sciences / Medicine and Life Sciences / Medicine and Pharmacy *(or equivalent)*
- Other ____

And you have published ...

▼ ...fewer than 10 peer-reviewed papers ... ...over 210 peer-reviewed papers

Your gender is...

- Male
- Female
- Other
- Prefer not to answer

Are you currently working in your country of origin?

- No
- Yes

Display This Question:
If Are you currently working in your country of origin? = No

What is your country of origin?

▼ Prefer not to say / Afghanistan ... Zimbabwe

Display This Question:
If Are you currently working in your country of origin? = No

Have you worked as a researcher/research student in your country of origin?

- Yes
- No

Have you ever been involved in **evaluating researchers for promotion, tenure, or career advancement**?

- Yes
- No
- Not sure (explain) ____

End of Block: Demographics

Start of Block: Time management

Allright! Now we would like to know how you spend your time as a researcher.

Do you work full time as a researcher/PhD student?

- o Yes
- o No

Display This Question:
If Allright! Now we would like to know how you spend your time as a researcher. Do you work full tim... = No

Where would you situate yourself?

▼ less than 25% research employment ... over 75% but less than 100% research employment

On average, **HOW MANY HOURS** per week **do you work?** (overtime included)

I really work...

0 20 40 60 80

And during your work time, what **PERCENTAGE (%) of your time** do you spend on the following three pillars **in reality**, and **how would you like it to be**?

In reality...

PERCENTAGE (%) of your time

0 20 40 60 80 100

TEACHING
RESEARCH
OTHER

In my ideal world...

PERCENTAGE (%) of your time

0 20 40 60 80 100

TEACHING
RESEARCH
OTHER

And now a little more into the details:

What PERCENTAGE (%) of your time would you say you spend on average on each of the following activities **in reality**, and **how would you like it to be**?

In reality...

PERCENTAGE (%) of your time

	0	20	40	60	80	100
DIRECT STUDENT SUPERVISION						
HANDS-ON RESEARCH WORK (e.g., lab work, data analysis)						
STAYING UP TO DATE (e.g., reading, listening, building skills, etc.)						
WRITING PAPERS						
REVIEWING						
GRANT WRITING						
ANYTHING ELSE (e.g., administration, meetings, etc.)						

In your ideal world...

PERCENTAGE (%) of your time

	0	20	40	60	80	100
DIRECT STUDENT SUPERVISION						
HANDS-ON RESEARCH WORK (e.g., lab work, data analysis)						
STAYING UP TO DATE (e.g., reading, listening, building skills, etc.)						
WRITING PAPERS						
REVIEWING						
GRANT WRITING						
ANYTHING ELSE (e.g., administration, meetings, etc.)						

End of Block: Time management

Start of Block: Statements

Great! Now that we know each other, let's get to business!

In the following questions, we wish to know the impact of typical research activities
(A) on **advancing your career**
(B) on **advancing science**, and
(C) on **your personal satisfaction**

There will be **18 research activities** to rate. We numbered then from 18 to 1 so you know how many you have left!

18. Publishing papers is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

17. Publishing in high impact journals is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

16. Publishing commentaries or editorials is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

15. Publishing more papers than others is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

14. Publishing open access is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

13. Peer reviewing is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

12. Replicating past research is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

11. Publishing findings that did not work (i.e., negative findings) is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

10. Sharing your full data and detailed methods is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

9. Reviewing raw data from students and collaborators is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

8. Conducting innovative research with a high risk of failure is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

7. Connecting with renowned researchers is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

6. Collaborating across borders, disciplines, and sectors is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

5. Getting cited in scientific literature is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

4. Having your papers read and downloaded is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

3. Having public outreach (e.g., social media, news, etc.) is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

2. Having your results used or implemented in practice is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

1. Having luck is...

	essential	important	irrelevant	unfavorable	detrimental
...in advancing my career	○	○	○	○	○
...in advancing science	○	○	○	○	○
...to my personal satisfaction	○	○	○	○	○

Feel free to leave a comment (optional) ____

End of Block: Statements

Start of Block: End

Do you have any other comments or thoughts you would like to share with us? (optional)

____
____
____

End of Block: End

APPENDIX 11

ANALYSIS RESULTS FOR REPORTED TIME ALLOCATION IN REALITY AND IN AN IDEAL WORLD

	Area	Reality	Ideal world	Statistical results
General pillars	Teaching	Mean 11.9 Median 10	Mean 15.3 Median 15	**Paired t(92) 3.35 95% CI 0.10, 0.40 $p<0.01$**
	Research	Mean 63.5 Median 60	Mean 72.6 Median 75	**Paired t(92) 4.28 95% CI 0.04, 0.12 $p<0.001$**
	Other	Mean 24.6 Median 20	Mean 12.1 Median 10	**Paired t(92) –4.99 95% CI –0.51, –0.22 $p<0.001$**
Detailed activities	Supervision	Mean 11.5 Median 10	Mean 11.9 Median 10	Paired t(92) 1.14 95% CI –0.06, 0.20 $p=0.26$
	Hands on research work	Mean 28.7 Median 26	Mean 36.1 Median 35	**Paired t(92) 3.94 95% CI 0.07, 0.21 $p<0.001$**
	Staying up to date	Mean 13.5 Median 10	Mean 16.8 Median 15	**Paired t(92) 4.12 95% CI 0.09, 0.26 $p<0.001$**
	Writing papers	Mean 12.9 Median 10	Mean 15.8 Median 15	**Paired t(92) 2.93 95% CI 0.04, 0.22 $p<0.01$**
	Reviewing	Mean 6.0 Median 5	Mean 6.2 Median 5	Paired t(92) 0.57 95% CI –0.07, 0.12 $p=0.57$
	Grant writing	Mean 9.4 Median 6	Mean 6.0 Median 5	Paired t(92) –1.56 95% CI –0.21, 0.03 $p=0.12$
	Anything else	Mean 18.0 Median 15	Mean 7.2 Median 5	**Paired t(92) –8.10 95% CI –0.53, –0.32 $p<0.001$**

APPENDIX 12

ANALYSIS RESULTS FOR RATINGS OF THE DIMENSIONAL IMPORTANCE OF THE 18 SUCCESS INDICATORS

Statement	Mean rating for each dimension	Result*	Bonferroni post hoc pairwise comparisons		
			Career vs Science	Career vs Satisfaction	Science vs Satisfaction
Publishing papers is…	Career: 4.52 Science: 4.37 Satisfaction: 3.90	(SA) F(2, 250) = 38.188 p < 0.001	CI(0.001, 0.301) p = 0.048 md = 0.151	**CI(0.468, 0.786) p < 0.001 md = 0.627**	**CI(0.341, 0.611) p < 0.001 md = 0.476**
Publishing in high impact journals is…	Career: 4.31 Science: 3.73 Satisfaction: 3.67	(SA) F(2, 250) = 36.701 p < 0.001	**CI(0.418, 0.741) p < 0.001 md = 0.579**	**CI(0.477, 0.809) p < 0.001 md = 0.643**	CI(−0.100, 0.227) p = 0.444 md = 0.063
Publishing commentaries or editorials is…	Career: 3.48 Science: 3.70 Satisfaction: 3.34	(SA) F(2, 250) = 14.538 p < 0.001	**CI(−0.360, −0.085) p = 0.002 md = −0.222**	CI(0.000, 0.269) p = 0.049 md = 0.135	**CI(0.232, 0.482) p < 0.001 md = 0.357**
Publishing more papers than others is…	Career: 3.83 Science: 2.89 Satisfaction: 3.01	(GG) F(1.779, 222.388) = 70.233 p < 0.001	**CI(0.761, 1.127) p < 0.001 md = 0.944**	**CI(0.636, 1.015) p < 0.001 md = 0.825**	CI(−0.258, 0.020) p = 0.092 md = −0.119
Publishing open access is…	Career: 3.48 Science: 4.35 Satisfaction: 3.67	(SA) F(2, 250) = 62.624 p < 0.001	**CI(−1.034, −0.696) p < 0.001 md = −0.865**	**CI(−0.357, −0.024) p = 0.025 md = −0.190**	**CI(0.529, 0.821) p < 0.001 md = 0.675**
Peer reviewing is…	Career: 3.39 Science: 4.43 Satisfaction: 3.47	(SA) F(2, 250) = 81.399 p < 0.001	**CI(−1.221, −0.858) p < 0.001 md = −1.040**	CI(−0.260, 0.101) p = 0.386 md = −0.079	**CI(0.784, 1.136) p < 0.001 md = 0.960**
Replicating past research is…	Career: 2.83 Science: 3.98 Satisfaction: 3.09	(GG) F(1.847, 230.832) = 81.530 p < 0.001	**CI(−1.360, −0.942) p < 0.001 md = −1.151**	**CI(−0.442, −0.066) p = 0.008 md = −0.254**	**CI(0.735, 1.059) p < 0.001 md = 0.897**
Publishing findings that did not work (i.e., negative findings) is…	Career: 2.84 Science: 4.48 Satisfaction: 3.61	(GG) F(1.820, 227.487) = 187.113 p < 0.001	**CI(−1.823, −1.462) p < 0.001 md = −1.643**	**CI(−0.951, −0.588) p < 0.001 md = −0.770**	**CI(0.734, 1.012) p < 0.001 md = 0.873**

Sharing your full data and detailed methods is...	Career: 3.29 Science: 4.40 Satisfaction: 3.67	(GG) F(1.906, 238.310) = 106.656 p < 0.001	**CI(−1.278, −0.960) p < 0.001 md = −1.119**	**CI(−0.554, −0.224) p < 0.001 md = −0.389**	**CI(0.594, 0.867) p < 0.001 md = 0.730**
Reviewing raw data from students and collaborators is...	Career: 3.37 Science: 4.20 Satisfaction: 3.64	(SA) F(2, 250) = 50.707 p < 0.001	**CI(−0.992, −0.658) p < 0.001 md = −0.825**	**CI(−0.446, −0.094) p = 0.003 md = −0.270**	**CI(0.403, 0.708) p < 0.001 md = 0.556**
Conducting innovative research with a high risks of failure is...	Career: 3.29 Science: 4.47 Satisfaction: 3.91	(GG) F(1.614, 201.766) = 66.452 p < 0.001	**CI(−1.406, −0.943) p < 0.001 md = −1.175**	**CI(−0.836, −0.402) p < 0.001 md = −0.619**	**CI(0.410, 0.701) p < 0.001 md = 0.556**
Connecting with renowned researchers is...	Career: 4.35 Science: 3.91 Satisfaction: 3.98	(SA) F(2, 250) = 24.566 p < 0.001	**CI(0.295, 0.578) p < 0.001 md = 0.437**	**CI(0.238, 0.508) p < 0.001 md = 0.373**	CI(−0.185, 0.058) p = 0.304 md = −0.063
Collaborating across borders, disciplines, and sectors is...	Career: 4.25 Science: 4.64 Satisfaction: 4.36	(GG) F(1.558, 194.781) = 14.565 p < 0.001	**CI(−0.550, −0.228) p < 0.001 md = −0.389**	CI(−0.274, 0.068) p = 0.235 md = −0.103	**CI(0.184, 0.388) p < 0.001 md = 0.286**
Getting cited in scientific literature is...	Career: 4.46 Science: 3.66 Satisfaction: 3.98	(GG) F(1.890, 236.280) = 55.630 p < 0.001	**CI(0.633, 0.970) p < 0.001 md = 0.802**	**CI(0.345, 0.623) p < 0.001 md = 0.484**	**CI(−0.463, −0.172) p < 0.001 md = −0.317**
Having your papers read and downloaded is...	Career: 3.90 Science: 3.90 Satisfaction: 4.10	(SA) F(2, 250) = 4.873 p = 0.008	CI(−0.158, 0.158) p = 1.000 md = 0.000	**CI(−0.365, −0.048) p = 0.011 md = −0.206**	**CI(−0.343, −0.070) p = 0.003 md = −0.206**
Having public outreach (e.g., social media, news, etc.) is...	Career: 3.84 Science: 3.77 Satisfaction: 3.72	(SA) F(2, 250) = 1.251 p = 0.288	———	———	———
Having your results used or implemented in practice is...	Career: 4.02 Science: 4.26 Satisfaction: 4.37	(SA) F(2, 250) = 12.875 p < 0.001	**CI(−0.380, −0.096) p = 0.001 md = −0.238**	**CI(−0.479, −0.203) p < 0.001 md = −0.341**	CI(−0.233, 0.027) p = 0.118 md = −0.103
Having luck is...	Career: 4.27 Science: 4.02 Satisfaction: 3.89	(SA) F(2, 250) = 14.229 p < 0.001	**CI(0.115, 0.377) p < 0.001 md = 0.246**	**CI(0.240, 0.522) p < 0.001 md = 0.381**	CI(−0.022, 0.292) p = 0.091 md = 0.135

* We report Greenhouse-Geisser (GG) tests when the results of Mauchly's test of sphericity could not confirm the sphericity of the data. Otherwise, Sphericity Assumed (SA) tests are reported. md = mean difference, CI = confidence intervals

APPENDIX 13

CONTRIBUTOR ROLE TAXONOMY (CREDIT) DEFINITIONS

The following definitions are taken from the CRediT website at https://casrai.org/credit/

Conceptualization: Ideas; formulation or evolution of overarching research goals and aims.

Funding acquisition: Acquisition of the financial support for the project leading to this publication.

Project administration: Management and coordination responsibility for the research activity planning and execution.

Methodology: Development or design of methodology; creation of models.

Resources: Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.

Software: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.

Investigation: Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection.

Data curation: Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.

Formal analysis: Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data.

Visualization: Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.

Validation: Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.

Supervision: Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.

Writing – original draft: Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).

Writing – review & editing: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages.

Acknowledgements

Before addressing individuals who supported me in completing this PhD thesis, I would like to thank everyone who participated in our interviews, focus groups, and survey. In the end, it is their ideas and reflections that made this thesis happen. I sincerely thank each of them for sharing their thoughts, knowledge, and time with me.

Now from those who can be named for their contributions, I would like first and foremost to thank my supervisor Wim Pinxten. Wim did not only find the resources for this project to happen, but also invested himself entirely through his genuine interest and enthusiasm in the project. He made me realise how important it is for PhD students to have someone who believes in their work, but also someone who believes in them, who gives them the strength to go on, and who encourages them to be open and genuine. Wim showed me that looking beyond the box can lead to great things, and that everything can (and maybe should) be questioned, even concepts that have been passively accepted for decades. I thank him for transmitting his hope for change and for allowing me to continue to think and grow. Thank you Wim.

Second, I would like to thank Raymond De Vries. Ray did not only assist me intellectually throughout this project, but he and his wife Charlotte also became family whenever I visited them or met them in conferences around the world. Ray taught me the value of listening, a value which we often forget in today's academic chaos. His generosity of time, attention, and efforts inspired me and shaped the type or scientist I hope I could become if I continue in academia.

I also want to thank Ivo Lambrichts, Geert Molenberghs, and Sven Hendrix, who served on my doctoral committee and with whom I had stimulating discussions about research integrity. It is through these discussions that I decided to focus my thesis on success, a decision which shaped my entire PhD. I appreciate how open for discussion they were, and am thankful to have had such a dedicated committee to guide me through the PhD. I extend additional thanks to all contributors mentioned in the chapters of this thesis, and to external jury members Anna Hatch, Brian C. Martinson, Frank Miedema, and Ann Peters whose contributions were instrumental in strengthening this thesis.

I also wish to thank the colleagues who have helped me adjust to the University and to the new life of a PhD student. Although I would be unable to name them all, I want to extend special thanks to Veronique Pousset who has been a guardian angel in the last months of the thesis and to Kathleen Ungricht, Stefanie Kerkhofs, Ann Peters, Nele Nivelle, Marleen Missotten, Jessica Boes, Pieter Lernout, Roeland Buckinx, Ilse Van Damme, Kim Pannemans, Stefanie Leunen, and many, many more who guided me in the university and patiently answered my never-ending questions.

Completing a PhD is not only an intellectual accomplishment however, but also a very personal one. I wish to thank the friends who were around to support me, those who provided their intellectual help but also their time, presence, and friendship. It would be impossible to name them all, but among those, a few special mentions: Silka Dawn Freiesleben for the incredible person that she is and for always reviving my flame to think, Cristina Murano and Manya Hendriks for their emotional support and empowering ideas, Patrick Miqueu for bringing me back to the here-and-now and making me believe that my cause is worthwhile, Geert Craenen for his inexplicable power to uplift my mood on demand, Marcello Ienca, for his reflections and his contagious way of enjoying the moment, Sheida Rabipour for always motivating me to go further, Carmen Peuters for reminding me to breathe, Heidi Howard, for making this PhD happen and for providing big-sister talks along the way, and so many more who accompanied me in some way or another. The PhD also created new and long-lasting friendships that helped me along the way: Katrien De Smedt-Jans, who became a pillar in my PhD journey and will be one of the greatest reasons to make me come back to Belgium, Dana Hawwash, my PhD-sister whose reflections and enthusiasm continue to inspire me, Bojoura Schouten, the super woman who showed me that with hard work, everything is possible, and Daniela Canu, who brought her sunshine to a very cold winter and taught me what resilience can accomplish.

I also want to thank my family, those whose love and understanding supported me along the way. Adriano e Luisa, che mi hanno offerto una famiglia qui in Europa, Peugeot, et par-dessus tout papa et maman qui m’ont donné ma curiosité, mon courage, et mon désir d'apprendre et qui m'ont donné mes ailes malgré tout l'effort émotionnel que ça imposait.

And finally, I want to thank Paolo Corsico. Paolo saw it all and lived it all with me. He was with me through the confusion, the obsession, the frustrations, the fear, the despair, the tears, the madness, the joys, and the thrills. His patience, his support, and his presence were essential to completing this journey.

Stichting transnationale Universiteit Limburg (tUL) is a cooperation between Hasselt University (Belgium) and Maastricht University (the Netherlands) and can be considered as one university with a home base in each country.

UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Hasselt University
Martelarenlaan 42 |BE-3500 Hasselt