Developing guidelines for research institutions

Journey towards research integrity

Krishma Labib

Developing guidelines for research institutions

Journey towards research integrity

Krishma Labib

VRIJE UNIVERSITEIT

Developing guidelines for research institutions
Journey towards research integrity

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op vrijdag 3 februari 2023 om 11.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Krishma Labib

geboren te Kabul, Afghanistan

promotoren: prof.dr. L.M. Bouter
                             prof.dr. G.A.M. Widdershoven

copromotoren: dr. J.K. Tijdink
                             dr. N.C. Evans

promotiecommissie: prof.dr. A.C. Molewijk
                             prof.dr. K. Sijtsma
                             dr.ir. J.T. van der Steen
                             prof.dr. J. Priess-Buchheit
                             prof.dr. K. Varantola
                             drs. J. Hoonhout

To the Tājik lions in Afghanistan resisting, who have shown me that it is possible to stand up for *good*, even when the whole world is against you,

and to the even braver lionesses across Afghanistan marching on the streets to fight for freedom, whose courage, perseverance, and integrity are unmatched.

Table of contents

Summary

Research integrity is concerned with conducting research responsibly, and according to the highest professional and ethical standards. As introduced in **Chapter 1,** in this thesis, I developed guidelines to research institutions on how to foster research integrity. I did this by exploring how research institutions can develop policies to foster, and raise awareness about, research integrity.

In *Section 1* of the thesis, my goal was to set the research agenda by investigating current practices of research integrity promotion at research institutions (the descriptive step), and exploring which topics should be addressed in institutional research integrity policies (the normative step). I addressed the descriptive step of looking at current practices, through a scoping review (**Chapter 2**). In this chapter, I found that while there are already many institutional practices for research integrity promotion globally, most of them focus on researchers', rather than institutions' responsibilities for fostering research integrity. In **Chapter 3**, I tackled the normative question of which topics should be included in institutional research integrity policies. Using a Delphi study that included a number of research policy experts and research leaders, I developed a comprehensive list of 12 topics that research institutions should address to foster research integrity. The comprehensiveness of the list, and order of the topics ranked, reflected experts' preference for a positive, preventative approach towards fostering research integrity. The highest ranked topic in the list in terms of importance was 'Research integrity education and training', followed by 'Responsible supervision'.

Section 2 of the thesis focused on developing guidelines for research institutions on research integrity. I specifically zoomed in further into 'Research integrity education and training' as the topic of the guidelines here, since the topic is considered as one of the most crucial ways that institutions can foster research integrity. The first step to developing the guidelines was to examine researchers' and other research stakeholders' views and preferences regarding how research institutions can develop and implement better research integrity education and training policies (**Chapter 4**). Using focus groups, I found that researchers and other research stakeholders support the provision of continuous research integrity education which targets all researchers (across ranks), and other institutional stakeholders (such as research integrity officers and institutional leaders). Stakeholders highlighted the importance of tailoring the education approach and its goals to the context and target group at hand, as well as making education attractive and motivating trainees sufficiently (for instance by offering target group appropriate incentives and rewards such as digital badges for students). Furthermore, they stressed the importance of taking into account various individual, institutional and system-of-science factors influencing research integrity education when implementing research integrity education and training policies.

After obtaining these insights, I proceeded to the step of co-creating institutional guidelines on research integrity education and training together with users. In **Chapter 5,** I discussed how research integrity guidelines can be jointly developed with users using co-creation methods – methods engaging participants in

interactive exercises aimed at jointly developing user centered outputs. Here, I elaborated on the various steps of co-creation methods including preparation of workshops that are informed by the existing evidence base, sensitization, workshop organization and facilitation, and analysis and guideline drafting and revision. I provided a brief overview of the content addressed by the guidelines to show that co-creation methods allow for the development of research integrity guidelines which are sensitive to stakeholders' needs. Moreover, I highlighted that co-creation methods allow for close stakeholder engagement and inclusion of diverse perspectives, which is important for raising commitment among stakeholders for the implementation of the guidelines. Finally, I argued that co-creation methods are most valuable at the initial phases of guideline development, and that they need to be followed up by further methods to finalize guidelines, such as revision working groups. The resulting RI education and training guidelines are presented in detail **Chapter 6**. The guidelines address the research integrity education of a) bachelor, master and PhD students; b) post-doctorate and senior researchers; c) other research integrity stakeholders; as well as d) continuous research integrity education. In the guidelines, I recommend the implementation of mandatory research integrity training (for all academic ranks); follow-up refresher training; informal discussions about research integrity; appropriate rewards and incentives for active participation in research education; and evaluation of research integrity educational events across target groups.

In *Section 3* of this dissertation, I reflected on an implementation concern regarding the guidelines developed; stakeholders warned that implementing research integrity guidelines might lead to increased bureaucracy within the institution, and that this could be counterproductive, if not accompanied by researcher commitment towards responsible research practices. I therefore explored the question of how research institutions can combine the implementation of research integrity rules with fostering researchers' commitment to engage in responsible research practices (**Chapter 7**). I argued that institutions can use and combine market (governance through incentives), bureaucracy (governance through rules) and network (cooperative governance in a group) mechanisms to foster research integrity. Using Habermas' Theory of Communicative Action, I discussed that institutions can use bureaucratic and market mechanisms to foster research integrity (such as rules and incentives, respectively), as long as these are rooted in network processes (e.g. involvement of stakeholders in the development and improvement of rules or incentives). I delved into the case of the Science Committee at Tilburg University's School of Social and Behavioral Sciences to show how an institution can combine rules about data management with network processes of involving the research community, to create learning and awareness about research integrity.

In **Chapter 8**, I highlighted the main messages that stem from this thesis. These include that: 1) the framing of research integrity matters for institutional policies; 2) research integrity guidelines should be tailored to the local context at hand; 3) it is important to be aware of and countervail the danger of creating a box-checking mentality when implementing institutional research integrity policies; and 4) research integrity is a journey. I concluded by emphasizing that this journey is a continuous one; fostering research integrity requires institutions to work together with researchers and other

stakeholders continuously to develop, revise and update policies in order to truly further the research community's goals of producing good research.

Introduction

Chapter 1

1. General introduction

My journey towards research integrity

When I started university, there seemed nothing more impressive than to be responsible for creating new knowledge and being at the forefront of what humanity knows about themselves and the world. Research seemed like a force for good, as the solution for addressing societal and global problems. I was determined that I would use all the knowledge I gained in my education to contribute towards a better understanding of mental health problems, in order to be able to help in the endeavor of ameliorating or maybe even curing some of them. I believed in the power of science and research – somewhat 'scientistically'[1] as one of my University College philosophy professors once remarked (with me failing to truly understand this statement at that time).

Coming to the realization that the world of research is not just imperfect, but has some serious problems, was a slow and painful process in my academic journey. In my last semester during my bachelor, I took a course on philosophy of science and finally understood what my professor had meant when he called my beliefs about research 'scientistic'. Science[2] is just one approach to looking at the world, and while it can provide us with important insights where other approaches might fail, it has its own limitations; for instance, science cannot be used for confirmation of hypotheses and theories, and even its power in falsification is up for debate (1). Furthermore, there are many important questions about the world, which might not be answerable using what are considered accepted scientific methods (e.g. What is the best way for me to live?). Therefore, scientism, or the belief that the world can be understood solely or mainly through scientific methods is problematic.

Through my neuroscience master, I learned that my initial conception of the role of research in society was blind to the hierarchies present in the world of research, and how certain scholars, methods, and approaches to research are privileged over others. I remember being surprised by claims around me that 'In those social sciences and humanities... all they do is just talk', indicating that experimentation and hypothesis testing are the only legitimate means of producing new knowledge. This attitude of many neuroscience researchers towards different ways of doing research came as a surprise to me, since as a student at a Liberal Sciences and Arts College before that, I had taken it for granted that it is necessary to combine insights from different research approaches in order to understand a phenomenon. In my view, it was necessary to approach mental health – just like many other issues in the world – in a holistic and multidisciplinary way, combining insights from various approaches focused on not only

---

[1] *'Scientism' refers to the strong belief that science is the most valuable approach to understanding the world (49).*
[2] *In this first page of the introduction, I am using the terms 'science' and 'research' somewhat interchangeably, as I had not committed to any specific conceptual distinction during my neuroscience education. In the context of neuroscience, the terms are often used interchangeably. From the next page onwards, I use the term 'research' instead of 'science' to be inclusive to all forms of knowledge production, regardless of whether they are considered as 'sciences' (e.g. the humanities).*

understanding human physiology, behavior, and psychology, but *also* stakeholders' lived experiences, beliefs, and preferences. Yet, I noticed that the collaboration between researchers from different approaches (e.g. philosophers and neuroscientists) was minimal, and thought that this could likely be related to the way that departments and research groups are organized in universities, as well as the kinds of funding that are available to researchers.

Particularly because I had seen the role of research as that of improving the world and contributing to good, I was furthermore unpleasantly shocked to witness that the world of research (seemingly across disciplines) is, just like any other human domain, fraught with integrity and ethics dilemmas. During my neuroscience master, I came to experience – albeit without the ethics vocabulary to be able to express it clearly at the time – that not all decisions that researchers make are based on epistemic or ethical considerations (such as 'What is the best decision here from a methodological viewpoint?' or 'What is the morally right thing to do in this dilemma?'), and that it is sometimes difficult to avoid compromising on epistemic and ethical considerations for purposes such as self-preservation, reputation and career. It became apparent when talking to researchers and teachers that problems in the culture of research on an institutional level (such as excessive work pressure, with sometimes limited support and guidance from supervisors), as well as the overall system of research (such as perverse incentives and financial insecurity), can negatively affect research practices.

What unnerved me most about these realizations was not so much that science is imperfect and has its limitations – in hindsight, since it is a human endeavor, that should have been obvious – but rather that there was not more being done to understand and address the serious problems of research at an institutional and systems level. Discouraged by my realizations of the problems surrounding research, but eager to contribute towards finding answers to some of these, I was lucky to stumble into the field of research integrity through a philosophy master in which I met and was inspired by my current supervisors and collaborators. It was these encounters that opened my eyes to the incredible work being done by research integrity researchers from different disciplinary backgrounds (including biomedicine, ethics, philosophy, and various social sciences) who are passionate about understanding and ameliorating the problems in the world of research that I had myself also witnessed. These experiences lay the groundwork for my ambitions and goals for this PhD project. In the next sections of this introductory chapter, I will elaborate on what research integrity is (including how attention to it arose), the importance of addressing research integrity on an institutional level (including the relevance of education in raising awareness about research integrity), as well as how I[3] have addressed these issues in this PhD trajectory.

---

[3] *In order to not confuse the reader by continuously switching between 'I' to denote personal reflections and views, and 'we' to denote collaborative work, I will use 'I' throughout the introduction chapter. However, it should be noted that the research conducted in this PhD trajectory was a collaborative effort, as is reflected in the content chapters.*

What is research integrity?

To understand what **research integrity (RI)** is (see also Table 1 for an explanation of the concepts in bold), it is important to look at the context in which attention to it arose. Historically, attention to RI has been sparked by scandalous cases of **research misconduct** involving fabrication (manufacturing false data), falsification (manipulating data), and plagiarism (using someone else's ideas without crediting them) (2–4). The notable Andrew Wakefield and Diederik Stapel affairs are just two such examples (5,6). The former involved a fraudulent study linking the mumps, measles and rubella (MMR) vaccine to autism (6,7), likely for personal and financial gains for Wakefield. Even years after the study has been debunked as fraudulent and Wakefield's misconduct has been exposed, the damage of this case on society's trust in research is ongoing, as evidenced by the presence of a strong global anti-vaccination movement (8). While the latter case – the Stapel affair, which involved data fabrication in social psychology research – did not have the same ramifications on public health (because of the topic of the research), Stapel's misconduct case caused shockwaves in his institution, the research community in the Netherlands, social psychology more broadly, and society at large (5). Furthermore, the careers of his PhD students and collaborators – who themselves were shown to be innocent of any misconduct – were harmed (5). These and many similar cases over the past decades have raised attention to the fact that sometimes researchers misbehave and violate the research community's **standards of acceptable research behavior**. In this way, RI initially came to be discussed in relation to misconduct and violations of integrity; RI could then be defined as preventing, and dealing with, these.

More recently, RI is recognized in more positive terms. This is because while cases of serious misconduct are extremely harmful to research and society at large, evidence has repeatedly shown that they are not frequent (9–11). It can be argued that the vast majority of researchers are well-intentioned, and pursue research with the aim of creating new knowledge and contributing positively to society (12). Yet, despite these good intentions, researchers do not always engage in behaviors that align with the highest ethical, professional and methodological standards of doing research, also referred to as **responsible research practices (RRPs)** (13,14). In fact, a troubling proportion of researchers admit to taking part in research practices that – while not considered as outright misconduct – are problematic and fall short of the standards of RRPs (9–11). These practices are referred to as **questionable research practices (QRPs)**. QRPs, such as using inappropriate research methods for the research question at hand or selectively reporting results, undermine the research endeavor by decreasing the trustworthiness, quality and relevance of research (15). It is thought that the high prevalence of QRPs in research is the likely culprit behind the replication crisis that has been demonstrated in several social science and biomedical science research fields (16). The 'replication crisis' refers to the alarming finding that a high proportion of published results in various research fields cannot be replicated by others, and is likely caused by QRPs (such as the ones mentioned in Table 1) (16,17). To truly foster RI, it is not sufficient to merely deal with – what seem to be relatively rare – cases of research misconduct; instead, the research community should also and primarily be focused on promoting RRPs and tackling the causes of QRPs. This more recent shift in attention

from cases of misconduct to QRPs and the replication crisis has reshaped the definition of RI. Instead of referring negatively to misconduct, it is now used positively to indicate efforts to increase the validity and trustworthiness of research (18,19).

The increased attention for the importance of fostering RI to safeguard trustworthy and valid research has given rise to research on RI, herewith referred to as the field of RI. This interdisciplinary field, which is closely related to other fields such as Responsible Research and Innovation (focused mainly on the societal impact of research) and Research Ethics (focused mainly on ethical considerations around research and its impact on research participants) , aims to contribute to a better understanding of researchers' behaviors, problems in the research record (such as biases and flaws in the published literature), and problems in the system and culture of research. During the last decades, many **codes of conduct** – documents containing aspirational principles and rules – on research integrity have been produced on international, national, and local levels, with the aim to guide researchers on how to conduct research responsibly and establish clear standards on RRPs (e.g. 20,21,22). Furthermore, there has been a rise in training programs offered to researchers on research integrity, which aim to increase researchers' awareness and competencies related to research misconduct, QRPs, and RRPs (23).

The importance of addressing RI on an institutional level

Factors influencing RI

Considering the harmful impact of QRPs on research, it might seem perplexing that they are so prevalent – particularly if only viewing RI as determined by the intentions and convictions of individual researchers. RI is, however, thought to be influenced by multiple factors, also including the institutional climate in which researchers conduct their work in, as well as the wider system of research, including funding structures and incentives as shown in Figure 1 (12,24,25). By **institutional climate**, I am referring to the formal and informal rules, infrastructures, social processes, and procedures present in a research institution. The level of support that researchers are provided with by their colleagues, supervisors, and institutions, for instance, has a large influence on whether researchers are able to engage in RRPs and to avoid QRPs (26). Furthermore, institutional rules on various research concerns, such as data management, will influence the kinds of behaviors that researchers engage with. **Policies** are courses of action or decisions which institutions can take to alter or influence the institutional climate, for instance to make the climate more conducive to RRPs.

The wider **system of research** includes extra-researcher and extra-institutional factors influencing research. For instance, incentives and funding structures that reward RRPs as opposed to QRPs or misconduct are also valuable for fostering RI (25,27). It has been highlighted in recent years that several factors in the system of research are likely to contribute towards QRPs and misconduct. An important contributor to this is the intense competition for scarcely available research funding (25,27,28). Another factor is the pressure to publish a high number of articles in high impact journals in order to further one's career (i.e. receive funding, be hired, or promoted, 25,27,28). Furthermore, since there are few permanent job positions

available for researchers, most researchers struggle with unstable job contracts and career insecurity (25,27,28). Given these pressures that researchers are faced with, it is not surprising that they sometimes intentionally or unintentionally resort to QRPs in order to secure their positions and advance their careers (e.g. by only reporting on positive findings to increase the chances of having a publication accepted in a prestigious journal, or lacking the time needed to conduct the most appropriate but time-consuming analysis of their data) (29).

This shows that to foster RI, it is necessary to not only address the behaviors of individual researchers, but also tackle the institutional and wider system of research elements that so strongly influence researchers' behavior (30–32). Current efforts towards promoting RI are mainly focused on the responsibilities of individual researchers (32). Merely focusing RI efforts on individual researchers may place an unfair and tremendous burden on researchers who are already overburdened with many responsibilities. This is especially a risk for junior researchers, whom many current RI initiatives (such as RI training programs) are aimed at (14), since these researchers have less power to make important research decisions and are faced with greater job insecurities than more advanced researchers (28). RI efforts merely focused on the individual level could potentially further increase the pressure on researchers, and might even thereby lead to an increase in QRPs, rather than serving the goal of fostering RI.

Who is responsible for RI?

Multiple stakeholders, including research institutions, funders, publishers, journals, learned societies (also referred to as professional research organizations), and local and national governments are jointly responsible for the institutional and system-of-research elements influencing RI as shown in Figure 1 (31). Of these, research institutions are especially interesting, since researchers are employed by research institutions, and thereby dependent on institutions' infrastructures, policies, and procedures. Research institutions have a direct influence on researchers' behaviors and a key role to play in fostering RI (30). While at least some research institutions have already established procedures for promoting some aspects of RI (e.g. providing RI and RE training and dealing with breaches of RI), there is a large variation with regard to the level of awareness of and emphasis on RI across research institutions (30). Furthermore, even institutions that already have significant attention for RI may lack comprehensive policies that sufficiently address all the necessary support systems, tools and initiatives needed to foster RI and increase awareness about it (30).
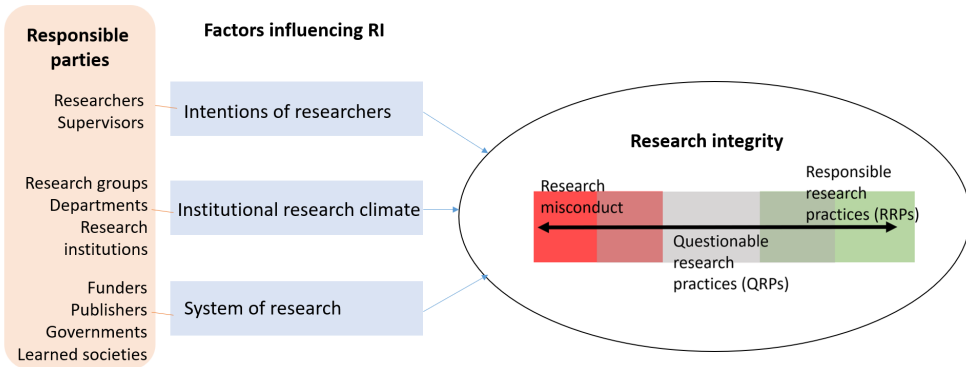
Figure 1 *Research integrity (RI), factors influencing it, and responsible parties. This figure combines insights from (33) and (34) to show the various factors and parties relevant for RI. Please note that while supervisors are also researchers, they have been put as an additional category, since there are certain RI responsibilities which are specific to supervisors compared to non-supervising researchers (e.g. providing sufficient guidance to supervisees).*

RI education

One of the topics that has already received much attention in the field of RI as an intervention to promote RI is the provision of RI education to researchers (e.g. 35). Education is considered crucial for raising awareness about RI and creating a responsible research culture (36). In the past decades, many RI training programs have been developed by institutions as means of raising awareness of RI (14,23,37,38). Often, they target PhD students, and are offered as a stand-alone course (14,23,38). However, focusing on RI training as a way to increase awareness of RI has received criticism for shifting the responsibility of RI on (mostly PhD) researchers, rather than adequately addressing the other parties and factors relevant for RI (e.g. 28). Namely, there is criticism that simply teaching PhD researchers how to behave – and then expecting that they engage in RRP – is unrealistic, and does not sufficiently support researchers to bring the insights they gain in the training into their research practice (36). Therefore, there is a need to look beyond trainings offered to PhD students as the means to raise awareness about RI. To take this into account, I will make a distinction between RI 'education' and 'training' in this thesis. By **'education'**, I refer to all means that can be used to raise awareness about RI (such as role-modeling RRPs, 39), whereas I use **'training'** to specifically discuss formal educational events such as courses and workshops. I see training to be a component of RI education, since it is one way to raise awareness about RI.

The need for guidelines

Despite the importance of broadening the concept of RI education to include more than RI trainings offered to PhD students, there is little information available to research institutions about how to raise awareness about RI (other than providing RI training). Furthermore, guidance on how to provide training which meets stakeholders'

needs is also missing. While international (e.g. 21,22) and national codes of conduct (e.g. 20) provide aspirational guidance to research institutions on how to foster RI more generally, there is little concrete information available to research institutions on how to operationalize these aspirational recommendations into policies which can be used in practice. Therefore, there is a need to develop concrete guidance, herewith referred to as **guidelines**, to research institutions on how to develop policies aimed at fostering, and raising awareness about, RI.

To be useful in supporting researchers to engage in RRP and address the causes of QRPs, this guidelines should be sensitive to researchers' and other research stakeholders' needs and preferences. This is because RI cannot be seen through an 'objective' view that is external to the context at hand. Instead, knowledge about RI should be rooted in the experiences and perspectives of the people involved in the research endeavor. Finding solutions for RI, therefore, requires stakeholders to engage in joint inquiry and come to consensus about RI. It is important to learn about the experiences and needs of the different stakeholders involved in RI, to better understand the challenges and opportunities regarding RI, so as to produce guidelines to research institutions which can adequately address them. Furthermore, in order to root the guidelines in stakeholders' perspectives, it is valuable to create guidelines on RI policy for research institutions which is jointly developed with all relevant stakeholders (e.g. researchers, research administrators, policy makers).

Creating good quality guidelines, requires a rigorous development process (42,43), also for RI. This process can be roughly divided into three separate phases. The first phase, involves studying the situation and evidence currently at hand (a descriptive step), as well as deciding on what questions to focus on when creating the guidelines (a normative step). The second phase consists of the development of recommendations for the guidelines, together with relevant stakeholders. The third and last phase involves exploring the implementation of the guidelines.

Table 1: Explanation of key concepts

Concept	Explanation
Research integrity (RI)	Initially associated with the prevention and sanction of research misconduct. More recently, primarily concerned with promoting RRPs and tackling the causes of QRPs.
Research misconduct	Violating the research community's standards regarding acceptable research behavior. This includes fabrication and falsification of data, as well as plagiarism.
Standards of acceptable research behaviors	Agreements among the research community regarding what is acceptable and unacceptable research practice. Some argue for universal standards that apply across different research contexts (e.g. 40), while others see standards as being dependent on the specific context of the research (e.g. whether the reserach is qualitative or quantitative, 41). While acknowledging the importance of allowing for contextual differences, rather than prescribing a universalist image of research, I would argue that it is uncontroversial to claim that there are some basic professional standards that are universal in research (e.g. providing adequate support to supervisees, reporting data fairly, not fabricating data, and being honest about results).
Responsible research practices (RRPs)	Behaviors that align with the highest ethical, professional and methodological standards of doing research. They constitute research practices that contribute towards trustworthy, high quality, relevant research.
Questionable research practices (QRPs)	Research practices that cannot be classified as outright misconduct, but are still problematic and fall short of responsible research practices. Examples include providing little or poor supervision to supervisees, using insufficient research methodology, and selectively reporting only study results which show flashy results while omitting negative findings.
Codes of conduct on RI	Documents providing aspirational principles and rules that readers can follow to foster RI.
Institutional research climate	The formal and informal rules, infrastructures, social processes, and procedures present in a research institution.
System of research	Extra-researcher and extra-institutional factors influencing researchers' behaviors.
Policies	Courses of action or decisions which institutions can take to alter or influence the institutional climate.
RI education	Approaches used to raise awareness about RI, including training.
RI training	Formal educational events used for RI education, such as courses.
Guidelines	A document containing concrete guidance about how to introduce or adjust institutional RI policies.

<u>The approach in this PhD trajectory</u>

Context of the PhD

The research presented in this PhD thesis is supported by the European Commission HORIZON 2020 framework program for Research and Innovation [Grant Agreement No. 824481], as a part of the Standard Operating Procedures for Research Integrity project (SOPs4RI, www.sops4ri.eu), which aims to develop a toolbox containing guidelines for European research institutions and funders on RI. As such, the work addresses concerns about research and RI that have been identified and prioritized by the European Commission. Furthermore, it reflects the collaboration of many contributors including supervisors, fellow RI researchers, and various stakeholders (e.g. policy makers, researchers, funders, etc.) across Europe and from different disciplinary backgrounds. The chapters in this PhD trajectory have played an important role in informing what to include in SOPs4RI's toolbox, and have contributed to the RI education and training guidelines that are now a part of this toolbox.

Research questions

In my PhD trajectory, I aimed to develop guidelines for research institutions on RI by exploring the research question: ***How can research institutions develop policies to foster, and raise awareness about, RI?*** The ultimate goal was to provide research institutions with guidance on how to support researchers to engage in RRPs. To provide concrete guidelines for research institutions that can be used in practice, I had to focus in on one topic within RI policy in more detail. I chose to address the topic of RI education and training, while other partners from the SOPs4RI consortium developed guidelines for other topics (30,44,45). The main rationale behind studying RI education and training in this PhD trajectory is that this topic has been discussed heavily in the RI literature as important for fostering RI (e.g. 35–37).

I split the main research aim into three sections, each focused on a distinct part of the guidelines development path, and comprised of separate sets of research questions (RQs):

- *Section 1: Setting the agenda*

  - RQ 1: What are current practices of RI promotion in research institutions?
  - RQ 2: Which topics should be addressed in institutional RI policies?

- *Section 2: Developing guidelines*

  - RQ 3: What are researchers' and other stakeholders' views and preferences regarding how research institutions can develop and implement better RI education and training policies?
  - RQ 4: How can RI guidelines be co-created together with lead users?
  - RQ 5: What should be included in the RI education and training policies of research institutions?

- *Section 3: Reflecting on implementation*

  - RQ 6: How can research institutions combine the implementation of RI rules with researcher commitment to foster RI?

Methodology

I used different qualitative research methods throughout this PhD trajectory, including scoping reviews, a Delphi method, focus groups, interviews, co-creation workshops, theoretical reflection, and case study analysis. What united all these approaches was the focus on staying close to research practices on the ground, and engaging with diverse research stakeholders to learn about their perspectives. Since the research was qualitative, I was not interested in conducting studies that produce generalizable findings, but was rather interested in developing insights that can inspire research institutions and other relevant parties, using inclusive samples that provide a diversity of views in terms of countries, disciplinary fields, genders, and stakeholder types in Europe.

I incorporated a research approach that borrows elements from forms of participatory research to involve research stakeholders in the knowledge production process (46). I used an iterative, multi-stage empirical cycle, with each subsequent empirical step consulting a different range of stakeholders and at a different stage of the policy development and implementation process. At different stages of the research, I aimed at different steps of Arnstein's ladder of participatory research (47); in some steps, I merely consulted with a large number of stakeholders to obtain a good understanding of their perspectives, while in others, stakeholders were actively involved in the co-creation of policy guidelines. This was to obtain some triangulation regarding my main research findings by consulting with different stakeholders in various steps of the PhD trajectory, using different methods (48).

Given the qualitative research approach, I did not work with hypotheses in this research, and I was flexible with making adjustments to our data analyses, often in a data driven manner. This was important to allow me to learn from the data obtained, and feed new insights into the next steps of the research process. I was careful to be transparent about various decisions and changes made to our preregistered research protocols, by explicitly adding amendments to all research protocols and providing complete information in publications about the steps I took. Sometimes this information was provided in appendices or hyperlinks in order not to overwhelm the reader.

As a researcher of RI, I am dedicated to engaging in RRPs myself. Consequently, I have taken several steps in my PhD trajectory. First, already at the start, my supervisors and I set clear expectations and roles regarding collaboration on the research (including issues such as authorship and supervision). There was also a robust data management policy for the entire project, to ensure that our data is responsibly handled. Furthermore, as an advocate for making research more open, I have preregistered all empirical work and taken the steps necessary to make data as openly accessible as is possible without jeopardizing identity protection. I have ensured that all our publications are publicly accessible, as well as made use of preprints. Together with

my supervisors and collaborators, I have also had regular discussions about how to improve the quality of our work, reduce mistakes and address different dilemmas and mistakes that occur in our research.

Thesis outline

Including this general introduction and a general discussion at the end, this thesis is divided into eight chapters. Chapters 2-7 constitute the body of the thesis, and are organized in three separate sections. In **Section 1**, I outline the research I conducted to set the agenda for this PhD thesis. As reported in **Chapter 2**, I started the research by using a scoping review, exploring the published and gray literature, to find existing best practices for the promotion of RI in research institutions *(RQ 1)*. These insights were valuable for identifying the current gaps and lacunas with regard to RI promotion at the institutional level. After exploring the literature, in **Chapter 3**, I identified which topics research institutions should address in their RI policies *(RQ 2)*. For this, I used a Delphi methodology to obtain consensus among research policy experts and leaders across Europe on a set of topics that institutional RI policies should address.

I focused on creating guidelines for research institutions on how to foster RI and raise awareness about it in **Section 2** of the body of the thesis. In **Chapter 4**, I zoomed in on RI education and training. To investigate what research institutions should take into account when developing and implementing RI education and training policies *(RQ 3)*, I conducted 30 focus groups with 147 researchers and other research stakeholders from various disciplines across Europe. To help research institutions develop and implement RI education policies, I next conducted a co-creation project to develop guidelines for RI education together with potential lead users (i.e. RI experts and research administrators). Co-creation methodology is a novel and promising approach in designing guidelines. In **Chapter 5**, I provide advice to researchers and practitioners interested in using co-creation methods for guideline development *(addressing RQ 4)*. I do this by reflecting on the different steps necessary to conduct co-creation research, including preparation, sensitization, workshops, and analysis. I also discuss why and when researchers might want to use co-creation methodology for developing guidelines. In **Chapter 6**, I elaborate on the results of our co-creation work, namely the content of the guidelines on RI education *(RQ 5)*.

In **Section 3** of the thesis, I have a chapter focused on an implementation challenge highlighted in earlier chapters of the thesis. Namely, one of the concerns that was repeatedly raised by various stakeholders in the different empirical steps taken in the PhD trajectory was that RI policies might lead to the creation of many rules, and that these might not receive sufficient commitment from researchers, which might make RI policies ineffective. It was, hence, interesting to explore how research institutions can combine the implementation of RI rules with researcher commitment to foster RI *(RQ 6)*. In **Chapter 7**, I combined a theoretical reflection – using insights from governance theory and Habermas' Theory of Communicative Action – and case analysis – of Tilburg University's 'Science Committee at the School of Social and Behavioral Sciences – to show how RI policy at an institutional level can be a driver of substantial improvements in research practices. In the general discussion – shown in **Chapter 8** – I reflect on how the work done in this PhD trajectory has provided a

concrete and comprehensive overview about topics institutions can address in their RI policies. I discuss the strengths and limitations of my research approach, and also provide directions for future research.

Contributions

KL drafted and revised this chapter. LB, GW, NE and JT reviewed drafts and accepted the final version of the chapter.

References

1. Chalmers AF. What is this thing called science? 4th Ed. Brisbane: University of Queensland Press; 2013. pp. 81-96.
2. Mody CCM, Sibum HO, Roberts LL. Integrating research integrity into the history of science. Vol. 58, History of Science. London: SAGE Publications; 2020. p. 369–85.
3. Titus S, Wells JA, Rhoades LJ. Repairing research integrity. Nature. 2008;453:980–2.
4. Barnes BE, Friedman CP, Rosenberg JL, Russell J, Beedle A, Levine AS. Creating an infrastructure for training in the responsible conduct of research: the University of Pittsburgh's experience. Acad Med. 2006;81(2):119–27.
5. Levelt Committee, Noort Committee, Drenth Committee. Flawed science: The fraudulent research practices of social psychologist Diederik Stapel [Internet]. 2012. Available from: https://www.tilburguniversity.edu/sites/default/files/download/Final report Flawed Science_2.pdf
6. Dyer C. Lancet retracts MMR paper after GMC finds Andrew Wakefield guilty of dishonesty. Br Med J. 2010;340(c696):281.
7. Motta M, Stecula D. Quantifying the effect of Wakefield et al. (1998) on skepticism about MMR vaccine safety in the US. PLoS One. 2021;16(8):e0256395.
8. Opel DJ, Diekema DS, Marcuse EK. Assuring research integrity in the wake of Wakefield. Br Med J. 2011;342(d2):179–80.
9. Gopalakrishna G, Ter Riet G, Vink G, Stoop I, Wicherts JM, Bouter LM. Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. PLoS One. 2022;17(2):e0263023.
10. Xie Y, Wang K, Kong Y. Prevalence of research misconduct and questionable research practices: A systematic review and meta-analysis. Sci Eng Ethics. 2021;27(4):1–28.
11. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS One. 2009;4(5):e5738.
12. Joynson C, Leyser O. The culture of scientific research. F1000Research. 2015;4(46).
13. Gopalakrishna G, Wicherts JM, Vink G, Stoop I, van den Akker OR, ter Riet G, et al. Prevalence of responsible research practices among academics in The Netherlands. F1000Research. 2022;11(471).
14. Boehme O, Föger N, Hiney M, Peatfiled T, Petiet F. Research Integrity Practices in Science Europe Member Organisations. 2016. DOI: 10.5281/zenodo.5060196
15. Bouter L, Tijdink J, Axelsen N, Martinson BC, Ter Riet G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. Res Integr Peer Rev. 2016;1(1):17.
16. Ioannidis JPA. Why most published research findings are false. PLoS Med. 2005;2(8):e124.
17. Cobey K, Fehlmann CA, Franco MC, Ayala AP, Sikora L, Rice DB, et al. Epidemiological characteristics and prevalence rates of research reproducibility across disciplines: A scoping review. OSF Preprints. OSF Preprints; 2022. DOI: 10.31219/osf.io/k6nf4

18. *Research integrity is much more than misconduct. Nature. 2019;570(5). DOI: 10.1038/d41586-019-01727-0*

19. *Downey F, Veitch R. Research integrity—it's about more than misconduct. Forensic Sci Res. 2021;6(4):283–4.*

20. *VSNU. Netherlands Code of Conduct for Research Integrity. 2018. DOI: https://doi.org/10.17026/dans-2cj-nvwu*

21. *ALLEA. European Code of Conduct for Research Integrity [Internet]. Berlin; 2017. Available from: https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf*

22. *Resnik DB, Shamoo AE. The Singapore Statement on Research Integrity. Account Res. 2011;18(2):71–5.*

23. *Fanelli D. MLE on Research Integrity – Training & Education – Thematic Report No 4 [Internet]. 2019. Available from: https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE RI No4 - Training and education.pdf*

24. *Rifai N, Annesley TM, Moore S, Caplan AL, Sweet DJ, Hornung P, et al. Maintaining research and publication integrity. Clin Chem. 2019;65(2):230–5.*

25. *Titus S, Bosch X. Tie funding to research integrity. Nature. 2010;466(7305):436–7.*

26. *Haven T, Pasman HR, Widdershoven G, Bouter L, Tijdink J. Researchers' Perceptions of a Responsible Research Climate: A Multi Focus Group Study. Sci Eng Ethics. 2020;26(6):3017–36.*

27. *Aubert Bonn N, Bouter L. Research assessments should recognize responsible research practices—Narrative review of a lively debate and promising developments. OSF Preprints. 2021. DOI: 10.31222/osf.io/82rmj*

28. *Aubert Bonn N, Pinxten W. Rethinking success, integrity, and culture in research (Part 2)—A multi-actor qualitative study on problems of science. Res Integr Peer Rev. 2021;6(1):1–18.*

29. *Parlangeli O, Guidi S, Marchigiani E, Bracci M, Liston PM. Perceptions of work-related stress and ethical misconduct amongst non-tenured researchers in Italy. Sci Eng Ethics. 2020;26(1):159–81.*

30. *Mejlgaard N, Bouter L, Gaskell G, Kavouras P, Allum N, Bendtsen AK, et al. Research integrity: nine ways to move from talk to walk. Nature. 2020;586:358–60.*

31. *Bouter L. Fostering responsible research practices is a shared responsibility of multiple stakeholders. J Clin Epidemiol. 2018;96:143–6.*

32. *Zwart H, Ter Meulen R. Addressing research integrity challenges: from penalising individual perpetrators to fostering research ecosystem quality care. Life Sci Soc Policy. 2019;15(5).*

33. *Steneck NH. Fostering integrity in research: Definitions, current knowledge, and future directions. Sci Eng Ethics. 2006;12(1):53–74.*

34. *Bouter L, Horn L, Kleinert S. Research integrity and societal trust in research. SA Hear. 2021;18(2):80–1.*

35. *Hermerén G, Zwart H, Marušić A, Fanelli D. MLE on Research Integrity – Final Report [Internet]. 2019. Available from: https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE RI_Final Report_0.pdf DOI: 10.2777/72096*

36. *Kalichman M. Rescuing responsible conduct of research (RCR) education. Account Res. 2014;21(1):68–83.*

37. *Kalichman M. A brief history of RCR education. Account Res. 2013;20(5–6):380–94.*

38. *Abdi S, Pizzolato D, Nemery B, Dierickx K. Educating PhD Students in Research Integrity in Europe. Sci Eng Ethics. 2021;27(1):1–19.*

39. *Maggio LA, Artino AR, Picho K, Driessen EW. Are you sure you want to do that? Fostering the responsible conduct of medical education research. Acad Med. 2018;93(4):544–9.*

40. *Merton RK. Science and technology in a democratic order. J Leg Polit Sociol. 1942;1(1):115–26.*

41. *Burns N. Standards for qualitative research. Nurs Sci Q. 1989;2(1):44–52.*

42. *Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Developing guidelines. BMJ. 1999;318(7183):593–6.*

43. *World Health Organization. WHO handbook for guideline development [Internet]. World Health Organization; 2014. Available from: https://apps.who.int/iris/bitstream/handle/10665/145714/9789241548960_eng.pdf?sequence=1&isAllowed=y*

44. *Pizzolato D, Labib K, Skoulikaris N, Evans N, Roje R, Kavouras P, et al. How can research institutions support responsible supervision and leadership? Account Res. 2022; DOI: https://doi.org/10.1080/08989621.2022.2112033*

45. *Horbach SPJM, Bouter LM, Gaskell G, Hiney M, Kavouras P, Mejlgaard N, et al. Designing and implementing a research integrity promotion plan: Recommendations for research funders. PloS Biol. 2022;20(8):e3001773.*

46. *Bergold J, Thomas S. Participatory research methods: A methodological approach in motion. Hist Soc Res Sozialforsch. 2012;191–222.*

47. *Arnstein S. "A Ladder of Citizen Participation": Journal of the American Institute of Planners (1969). In: The City Reader. Routledge; 2020. p. 290–302.*

48. *Carter N, Bryant-Lukosius D, DiCenso A, BlytheA J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurse Forum. 2014;41(5):545–7.*

49. *Sorell T. Scientism: Philosophy and the infatuation with science. London: Routledge; 2013.*

Section 1: Setting the agenda

Chapter 2

2. Practices for Research Integrity Promotion in Research Performing Organisations and Research Funding Organisations - A Scoping Review

Abstract

Research integrity (RI) is a continuously developing concept, and increasing emphasis is put on creating RI promotion practices. This study aimed to map the existing RI guidance documents at research performing organisations (RPOs) and research funding organisations (RFOs). A search of bibliographic databases and grey literature sources was performed, and retrieved documents were screened for eligibility. The search of bibliographical databases and reference lists of selected articles identified a total of 92 documents while the search of grey literature sources identified 118 documents for analysis. The retrieved documents were analysed based on their geographical origin, research field and organisational origin (RPO or RFO) of RI practices, types of guidance presented in them, and target groups to which RI practices are directed. Most of the identified practices were developed for research in general, and are applicable to all research fields (n = 117) and medical sciences (n = 78). They were mostly written in the form of guidelines (n = 136) and targeted researchers (n = 167). A comprehensive search of the existing RI promotion practices showed that initiatives mostly come from RPOs while only a few RI practices originate from RFOs. This study showed that more RI guidance documents are needed for natural sciences, social sciences, and humanities since only a small number of documents was developed specifically for these research fields. The explored documents and the gaps in knowledge identified in this study can be used for further development of RI promotion practices in RPOs and RFOs.

Keywords: Research integrity · Research ethics · Research integrity practices · Research integrity promotion · Research performing organisations · Research funding organisations

Abbreviations

ALLEA	European Academies
CORDIS	Community Research and Development Information Service
ENRIO	European Network of Research Integrity Offices
EU	European Union
JBI	Joanna Briggs Institute
LERU	League of European Research Universities
MEXT	Ministry of Education, Culture, Sports, Science and Technology, Japan
MLE	Mutual Learning Exercises
NASEM	National Academies of Sciences, Engineering, and Medicine
NESH	The National Committee for Research Ethics in the Social Sciences and the Humanities, Norway

ORI   Office of Research Integrity (United States)
RE   Research Ethics
RI   Research Integrity
RFO   Research Funding Organisation
RPO   Research Performing Organisation
SOP   Standard Operating Procedures
US   United States
WCRI   World Conferences on Research Integrity
WOS   Web of Science

Introduction

The scientific community emphasises the importance of research integrity (RI) because it represents the basis for the advancement of reliable and trustworthy knowledge and scientific endeavours (Aubert Bonn et al. 2017). In some countries, RI is also referred to as responsible conduct of research (RCR) (DuBois 2004; Steneck 2006; Kalichman 2013; Shamoo and Resnik 2015). In addition to providing written guidance for good research and mechanisms to encourage compliance with responsible practices, RI is also an integral part of researchers' moral obligation to be honest and responsible toward the system of science (Institute of Medicine and National Research Council 2002; Kalichman 2013).

Issues related to RI, RCR, research misconduct, and detrimental (questionable) research practices started to get more attention from the scientific community around the 1990s (Resnik and Shamoo 2017). Initiatives to prevent RI breaches started to develop at the same time, including the development of guidance documents, the examination of their quality in helping researchers to tackle these issues (Nobel, 1990), and the establishment of RI bodies, like the Office of Research Integrity in the USA (Steneck 2006). Whereas at the beginning of the development of RI as afield, the focus was on the individual researchers and prevention of misconduct, the promotion of RI and prevention of misconduct are today seen as a mutual responsibility of different organisations and individuals included in research (NASEM 2017, Bouter 2018; Hermerén 2019). Better understanding of RI and its implementation in practice is seen possible only if everyone acts responsibly and accomplishes their tasks related to RI promotion. This includes the responsibility of researchers to conduct research following the good practices and policies provided by research performing organisations (RPOs) and research funding organisations (RFOs). It includes the responsibility of both RPOs and RFOs to implement policies on good research practices, provide education to researchers, and have mechanisms in place that will deal with breaches of RI (Boeheme et al. 2016). Also, the responsibility of journals to prevent poor publication practices that may have detrimental consequences for the scientific community and society in general is important (Marušić et al. 2007; Bouter 2018).

A number of studies have addressed RI issues related to different stakeholders and disciplinary fields. Olesen et al. and Haven et al. explored the research misconduct perceptions of researchers from different disciplinary fields (Olesen et al. 2018; Haven et al. 2019). The effectiveness of existing interventions for RI improvement in different disciplinary fields, such as training and education or implementation of procedures for

handling cases of misconduct have also been explored (Marušić et al. 2016). Recently, emphasis has been put on RPOs' and RFOs' role in RI promotion. RPOs have an important role in the development and implementation of RI policies and compliance mechanisms (Forsberg et al. 2018). Through its organisational directors and boards, they have a role in raising awareness on RI issues, creating the environment of integrity, and changing the research evaluation practices and incentive structures (Hicks et al. 2015, Boeheme et al. 2016; Moher et al. 2019, Zwart and ter Meulen 2019; Bouter 2020). However, RPOs are not the sole actors in this vital mission, since their efforts in RI promotion can be augmented even more with the endeavours of RFOs (Bouter 2018, 2020). By implementing policies for good research practices and emphasising the importance of RI in funded research, such as the Wellcome Trust in their Guidelines for Good Research Practice (Wellcome Trust 2018), funders can impose high RI standards that need to be respected by those who apply for funds (both individual researchers and research organisations). These may include requests for RPOs to have fair procedures for dealing with RI, requests for researchers to provide a clear explanation of the relevancy of their study, and requests for adequate reporting and open access publishing of the study results to achieve reproducibility of its findings (Begley and Ioannidis 2015; Bouter 2016, NASEM 2018).

As the new knowledge on RI responsibilities of researchers and organisations is emerging, new documents are being developed. However, these documents are scattered through the academic literature, official sites of different RPOs and RFOs, and other professional organisations and networks. Also, RI guidance is presented in various types of documents—codes, guidelines, checklists, standard operating procedures, and others. Although there are studies on the existing RI policies in specific disciplinary fields, as well as research on the diversity of existing policies and terminology used across these documents (Godecharle et al. 2014; Komić et al. 2015; Aubert Bonn et al. 2017), there is no systematic effort to synthesise the knowledge of RI promotion practices in RPOs and RFOs. In this scoping review, we provide a broad overview of the RI guidance documents originating from the scientific literature and grey literature sources. In our analysis, we mapped the documents based on their geographical, disciplinary field and organisational origin, as well as based their relevance for different individuals and organisations in the research process. Our analysis also included identification of different RI topics related to different phases of the research process and the analysis of principles of good research declared in the documents. By exploring these guidance documents and the prescriptive and aspirational norms provided in them, we identified gaps in how RPOs and RFOs address RI and issues in this field that require additional attention.

Methods

We used a scoping review methodology (Tricco et al. 2016) following the guidance published in the Joanna Briggs Institute (JBI) Review's Manual (Peters et al. 2015).

Concept and Context

The concept of this review was that there is a wide range of existing practices/guidance documents in RPOs and RFOs with implications on RI promotion and avoiding research misconduct, as well as that these guidance documents may vary in their scope, means of addressing RI issues and stakeholders (e.g. policymakers, researchers, reviewers, students, committees and boards) to which they are directed.

This review examined the practices/guidance documents for RI promotion and avoiding research misconduct related to RPOs, RFOs, and other various stakeholders involved in research (policymakers, researchers, reviewers, students, committees and boards) with the aim of building an overarching view of the current situation regarding RI guidance. Moreover, the review examined RI guidance documents that exist in different research fields and are related to different research phases (research planning, conducting, dissemination and evaluation). It also explored the guiding principles presented across documents, as these principles could serve RPOs and RFOs in creating and preserving the RI environment (NASEM 2017).

Selection Criteria

The main eligibility criterion for the documents from peer-reviewed journals and grey literature was that these documents addressed any aspect of RI in RPOs and/or RFOs. By any aspect we meant RI issues related to different phases of the research process and with the different RI focus. For example, authorship issues, data management issues, investigations of research misconduct, RI education and other.

A description or summary of RI practices had to be provided in these documents in order for them to be included in the analysis. Editorials and commentaries were included as well when they met the above mentioned criterion.

We included all types of guidance documents on RI issues as 'practices'. This included guidance in the form of codes, guidelines, checklists, and standard operating procedures but did not exclude other types of guidance documents. Hence, the list of the different forms in which guidance for RI was presented was updated during the process of document screening and analysis.

Although the majority of documents contained the type of guidance on RI issues in their title or description, for documents that were not defined regarding the type of guidance we used the following criteria:

(a)   Code—a document providing general, rather than detailed guidance on ethical standards, principles, values, and rules of behaviour;

(b)   Guideline—a document more specific than code in providing guidance; a document providing specific instructions for performing a certain task or achieving a certain goal;

(c)   Checklist—a document presented as a clear list of items to be done, checked, or considered in performing a specific task;

(d)   Standard operating procedure (SOP) —a document providing detailed, step-by- step instructions for carrying out routine tasks and aimed at achieving uniformity and efficiency;

(e)   Flowchart—a document presenting guidance in the form of a diagram representing a workflow or process;

(f)   Legal document—a document established by a government or other authority, empowered by law, and outlining legal consequences; and

(g)   Policy—a document established and implemented by an organisation, containing adopted principles, rules, and procedures for conducting certain actions.

Other types of guidance used as a category in this review included reports, statements, declarations, white papers, as they had such a term set out in the title or description of the document.

Since academic integrity comprises fundamental values relevant for researchers and their work (Fishman 2014), documents related to academic integrity were included into our analysis whenever they reflected on research performance or researchers' behaviour, be it professional or unprofessional. Further, documents related to research ethics (RE) were also included if they addressed issues similar to RI, since RE and RI are not always clearly distinguished (Komić et al. 2015).

The search addressed practices relating to different scientific disciplines, categorised in advance as—medical sciences (including biomedicine), natural sciences (including engineering), social sciences, humanities, and 'research in general'. The latter term was used to map the practices that were not developed for RI in a specific field, but rather to be applicable across different scientific fields.

The search of bibliographic databases did not have geographical or language restrictions, while the grey literature search was limited to documents in English because of the possibility of retrieving a large number of documents that would need to be translated in order to be analysed. Since research misconduct emerged as an important problem in the late 1980s and 1990s (Resnik and Shamoo 2017), only the materials dating from 1990 onward were included in the screening process. The reason for this was based on the need for ensuring applicability and contemporaneity of identified practices and exploring currently existing gaps in knowledge.

Search of Bibliographical Databases

The search strategy was developed by three researchers who were assisted by a librarian specialised in systematic review search methodology. The development of the search strategy aimed at high sensitivity and included a broad approach to the field, based on the need for the identification of as many relevant documents as possible. As a starting point in the development of the search strategy, we used terms from the European Code of Conduct for Research Integrity (ALLEA 2017). The search strategy is available in Appendix 1 (Electronic Supplementary Material). We searched Scopus, Web of Science (WOS), Medline and PsycINFO bibliographical database. The search of Medline, WOS, and Scopus was performed on 18 February 2019, while the search of PsycINFO was performed on 12 February 2019. The obtained data were exported to the EndNote™ tool (Clarivate Analytics, Philadelphia, PA, USA).

Search of Grey Literature Sources

The search of grey literature encompassed several different sources: Open Grey database (Open Grey, INIST-CNRS), World Conferences on Research Integrity (WCRI) (The World Conferences on Research Integrity) website, the Community Research and Development Information Service (CORDIS) database (European Commission), Office of Research Integrity (ORI) (The Office of Research Integrity) website, European Network of Research Integrity Offices (ENRIO) (The European Network of Research Integrity Offices) website, the National Academies of Sciences, Engineering, and Medicine (NASEM) publications (The National Academies of Sciences, Engineering and Medicine), Science Europe publications (Science Europe), Mutual Learning Exercises (MLE) on Research Integrity reports (European Com- mission), and the League of European Research Universities (LERU) publication (The League of European Research Universities). Details of the search of grey literature sources are presented in Appendix 2 (Electronic Supplementary Material).

Selection of Documents

For documents that were retrieved by the search of bibliographic databases, duplicates and articles dating before 1990 were first removed and then the screening of the titles and abstracts was performed. The screening was conducted independently by two reviewers. In order to precisely define the criteria and the screening process, as well as to ensure that both reviewers would perform the task in the same manner, the reviewers first performed a pilot screening of the titles and abstracts of 100 records. After the pilot screening, they proceeded with the screening of the titles and abstracts of all the documents, after which they compared and discussed the obtained results in order to decide which documents would be included in the full-text analysis. In cases of disagreement, the final consensus decision was reached after a discussion with the third reviewer. In the following step, the three reviewers performed a full-text assessment of the documents in order to decide whether they were eligible for inclusion into the final analysis. To be included in the final analysis, a consensus had to be reached by at least two reviewers. In cases of major disagreements, the material was discussed with an additional reviewer. Documents that were not written in English were translated using tools such as Google Translate to explore whether they fulfilled the eligibility criteria. Reference lists of the documents included in the final analysis were screened by one reviewer to identify additional documents (sources of practices).

For grey literature sources, one researcher performed the search to identify documents that specifically met the set eligibility criteria. This means that all available documents were not extracted and screened, but rather the full-text screening was performed simultaneously with the search.

Data Extraction Process

For the documents from the bibliographic databases included in the final analysis (Fig. 1), two researchers performed the data extraction. The list of categories to be extracted was defined in advance and was continually updated by each researcher

during the charting process. The list is available in Appendix 3 (Electronic Supplementary Material). The categories were discussed by authors to reach the consensus on the final list. The data extraction of the material obtained from the grey literature search was performed by one researcher.

The following data were extracted: author(s) (for documents from bibliographic databases); title (for documents from bibliographic databases); year of publication; reference type, i.e. journal article, book, book section (for documents from bibliographic databases); journal (for documents from bibliographic databases); country of origin; research fields, i.e. humanities, social sciences, natural sciences (including engineering), medical sciences (including biomedicine), research in general; name of the practice; type of practice (type of guidance on RI issues), i.e. code, guideline, checklist, SOP, legal document, report, declaration, statement, flowchart, white paper, policy; whether the practice was more related to RPOs or RFOs or both; whether the practice was more related to institutions (organisations) or individuals or equally to both; target audience in practice, i.e. researchers, research groups, policymakers, funders, students, men- tors and supervisors, committees and members of committees, RI offices and officers, RI advisors, ombudsman, reviewers, administrators, whistle-blowers; description of the source of practice (for grey literature); principles addressed in practices. Documents were also categorised according to the phase of the research process—planning, conducting, dissemination, evaluation—as well as RI violations and resolutions and RI promotion. Within each research process, several RI topics were identified based on their relatedness to the process. Since the main research processes were defined broadly, the grouping of RI topics which were more related to the specific issue enabled us to capture the most prevalent RI issues addressed across practices. Two researchers independently developed the lists of RI topics during the extraction process. After finalising the analysis, the lists of RI topics were compared to detect any overlaps. The list of topics was finalised through a discussion and consensus between two researchers and in consultation with a third researcher.

Data Synthesis

After the data extraction, all the documents were summarised and analysed based on their geographical origin, the scientific field and organisational (RPO or RFO) origin of the identified practices, the types of practice (the type of guidance), and the target group to which the practices were directed. We also categorised the documents based on the research processes and RI topics addressed in them.

Furthermore, we extracted the guiding RI principles that were explicitly addressed in the documents. This means that RI principles had to be explicitly mentioned and explained in the chapters or parts of the text. The documents just stating RI principles without further elaboration were not included in the analysis. We mapped the extracted principles to the principles presented by the All European Academies in the European Code of Conduct for Research Integrity (ALLEA 2017) and those presented by the US National Academies of Sciences, Engineering, and Medicine in the book Fostering Integrity in Research (NASEM 2017). The aim was to observe the similarity in principles, and terms used to address the guiding principles. We used these

two documents because of their wide recognition and acceptance, as well as their up-to-dateness (both were updated in 2017). The extracted principles were mapped by one researcher and checked with the second researcher, upon which the agreement was reached for the final mapping.

Results

The search of Scopus, WOS, Medline, and PsycINFO retrieved 32,887 documents, 26,805 of which remained after removing the duplicates. The screening of the titles and abstracts left 130 documents for the full-text assessment of eligibility for the final analysis. In the following step, 73 documents were excluded, leaving 57 for the final analysis. The most prevalent reason for exclusion of 73 documents was that the documents did not present actual practices related to research or to RI or RE. Full details on the excluded documents are presented in Fig. 1. Five documents were excluded because we were unable to retrieve them in full text for analysis. The screening of the references from 57 documents included in the final analysis identified additional 35 documents (sources of practices) that were subsequently included in the final analysis and data charting. These additional documents (n = 35) were documents (codes, guidelines, books) provided on the websites of RPOs, RFOs, or other professional organisations. Reference search identified a single additional journal article (a commentary).
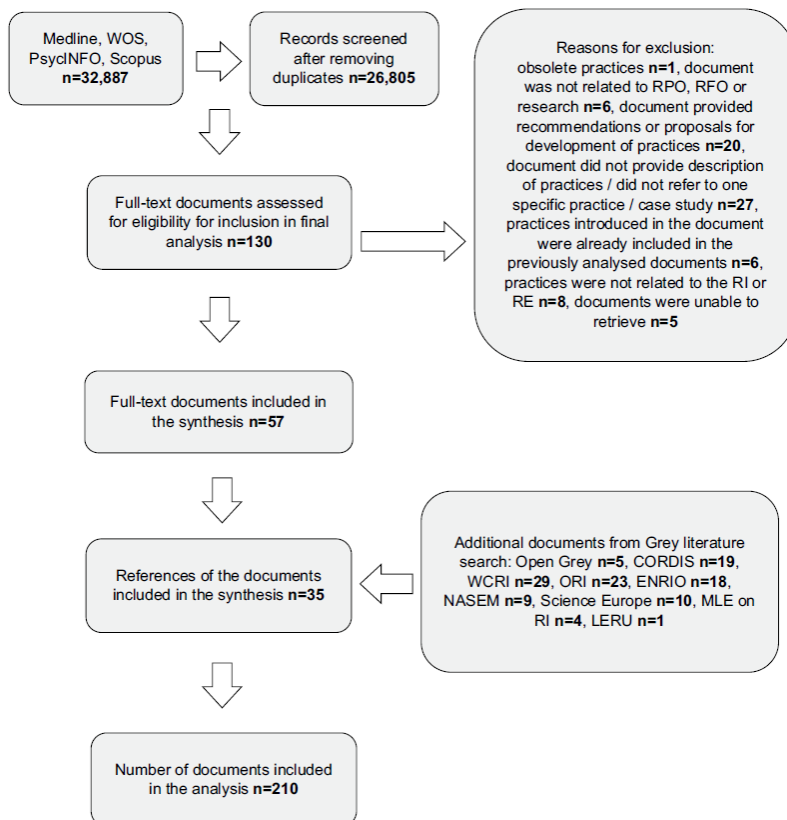
Fig. 1 *PRISMA-ScR flow diagram for the scoping review process. CORDIS Community Research and Development Information Service; ENRIO European Network of Research Integrity Offices; LERU League of European Research Universities; MLE on RI Mutual Learning Exercises on Research Integrity; NASEM National Academies of Sciences, Engineering and Medicine; ORI Office of Research Integrity; RE research ethics; RI research integrity; RFO research funding organisation; RPO research per-forming organisation; WCRI World Conferences on Research Integrity; WOS Web of Science*

Flow diagram contents:
- Medline, WOS, PsycINFO, Scopus **n=32,887**
- Records screened after removing duplicates **n=26,805**
- Reasons for exclusion: obsolete practices **n=1**, document was not related to RPO, RFO or research **n=6**, document provided recommendations or proposals for development of practices **n=20**, document did not provide description of practices / did not refer to one specific practice / case study **n=27**, practices introduced in the document were already included in the previously analysed documents **n=6**, practices were not related to the RI or RE **n=8**, documents were unable to retrieve **n=5**
- Full-text documents assessed for eligibility for inclusion in final analysis **n=130**
- Full-text documents included in the synthesis **n=57**
- References of the documents included in the synthesis **n=35**
- Additional documents from Grey literature search: Open Grey **n=5**, CORDIS **n=19**, WCRI **n=29**, ORI **n=23**, ENRIO **n=18**, NASEM **n=9**, Science Europe **n=10**, MLE on RI **n=4**, LERU **n=1**
- Number of documents included in the analysis **n=210**

The search performed in the Open Grey database, the websites of the World Conferences on Research Integrity, CORDIS, ORI, ENRIO, NASEM, and MLE identified 118 documents that described the practices for the analysis. The total number of all documents included in the final analysis was 210 (Fig. 1).

Origin of RI Practices

The largest number of documents was related to practices from the USA (n = 65), followed by practices that were developed by international organisations or pro- jects and not aimed at or developed by a specific country or countries, but instead could be applicable internationally (n = 50). Some examples of the practices that we mapped as international are Responsible Conduct in the Global Research Enterprise: A Policy Report by Inter Academy Council and the Inter Academy Partners (IAC and IAP 2012), World Health Organisation Guidelines for Good Clinical Practice for Trials on Pharmaceutical Products (WHO 1995) European Science Foundation Good scientific practice in research and scholarship (ESF 2000), and the Hong Kong Principles for Assessing Researchers: Fostering Research Integrity (Moher et al. 2019). Some documents contained the descriptions of practices related to more than one country, i.e. two or more counties were explicitly mentioned. In those cases, we included all the mentioned countries in the analysis. The origin of practices by country and the number of identified sources related to a particular country are presented in Fig. 2.

Fig. 2 Origin of practices by country (without international practices; number of international practices n = 52). The United States of America (n = 65), United Kingdom (n = 27), Japan (n = 9), the Netherlands (n = 9), Australia (n = 7), Norway (n = 6), Canada (n = 5), Austria (n = 4), South Africa (n = 4), Denmark (n = 3), France (n = 3), India (n = 2), Spain (n = 2), Switzerland (n = 2), Brazil (n = 1), Estonia (n = 1), Finland (n = 1), Germany (n = 1), Ireland (n = 1), Lithuania (n = 1), Nepal (n = 1), New Zealand (n = 1), Nigeria (n = 1), Poland (n = 1), Romania (n = 1), Singapore (n = 1). There were 52 documents which were international and could not be located to a single country. Source for the geographical map: https:// commons.wikimedia.org/wiki/File:BlankMap-World.svg (public domain)

In terms of scientific fields, the majority of documents referred to RI issues that are not related to any specific field, i.e. research in general (n = 117), followed by documents that addressed RI in medical research (n = 78). We identified 10 documents for RI

practices in social sciences, 10 for natural sciences (including engineering), and 4 related to RI practices in the field of humanities. Some documents referred to more than one scientific field, and in those cases, we counted each scientific field that was addressed. The most significant number of items included in the final analysis were practices that were more related to RPOs (n = 150). Although some practices related to RPOs were related to RFOs as well, we considered these practices to be primarily intended for RPOs since the guidance addressing the RFOs was only briefly mentioned. Guidance related equally to RPOs and RFOs was identified in 54 documents. Practices related to RFOs were identified in only 6 documents.

Type of Guidance for RI Promotion.

Based on the distinction between the types of guidance on RI issues, we identified 11 types of practices. Among them, guidelines were most prevalent (n = 136). Other identified types of guidance were codes (n = 35), policies (n = 26), legal documents (n = 14), reports (n = 10), checklists (n = 9), statements (n = 6), declarations (n = 4), flowcharts (n = 2), white papers (n = 1), and standard operating procedures (n = 1). Some sources of practices referred to more than one type of guidance, and in these cases we counted and mapped each practice that was mentioned. For this reason, the numbers that are presented are higher than the number of documents included in the final analysis.

We analysed the number of different types of guidance identified in this study over three time periods: 1990–1999, 2000–2009, and 2010–2019 (Fig. 3). Most of the identified practices dated from 2010 onward and the guidelines were mostly represented throughout all the three time periods. For some practices (n = 11), we were not able to define the exact time when they were developed, hence we did not include them in this analysis.
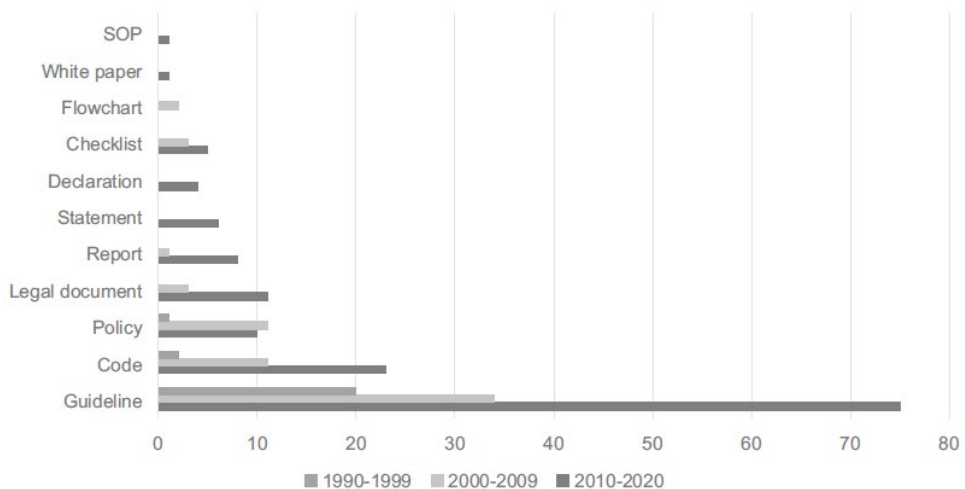
Fig. 3 *The number of practices in different time periods. The x-axis shows the number of practices, and the y-axis lists different types of practices. SOP standard operating procedure*

Target Group to Which Practices were Directed

RI practices addressed different individuals and organisations as target groups (Table 1). We grouped different individuals and organisations in the research chain that practices were aimed at into five primary categories: researchers, RPOs, RFOs, RE or RI bodies and other policymakers respectively. Most of the practices referred to more than one category.

Research Processes and RI Topics Identified in the Documents

We first classified the documents according to the steps of the research process—planning, conducting, dissemination, evaluation—as well as according to RI violations and resolutions and RI promotion. These were then broadened by the list of RI-related topics that were mentioned in the analysed documents. For example, in the category of 'RI violations and resolutions' we put documents that addressed research misconduct investigations, sanctions and other, while in the category of 'RI promotion' we put documents related to the development and implementation of RI practices, implementation of RI training and establishment of RI bodies. Some topics were related to more than one research process. We analysed which of the extracted topics were related to RPOs, RFOs, and/or other policy- makers. These practices reflected on the organisational procedures and measures that could be put into effect for individual researchers and for RI improvement in general. Classification by research processes and RI topics, together with the list of documents aimed at organisational level is available in Appendix 4 (Electronic Supplementary Material). The list of practices aimed at individual researchers only is presented in Appendix 5 (Electronic Supplementary Material).

principles and values that researchers and organisations should follow. We matched these principles to those outlined in two major policy documents: the European Code of Conduct for Research Integrity (the ALLEA code) (2017) and the National Academies of Sciences, Engineering and Medicine (NASEM) book Fostering Integrity in Research (2017). The comparison of fundamental principles is available in Appendix 6 (Electronic Supplementary Material).

Regarding the type of documents in which the principles were addressed, the majority were codes (n = 13), followed by guidelines (n = 9), statements (n = 3), and policies (n = 4). The list and description of all extracted principles are available in Appendix 7 (Electronic Supplementary Material).

Table 1: Individuals and organisations address in identified RI practices

Individuals and organisations by category and sub-category	No. of documents
Researchers (including research groups, students, mentors and supervisors, reviewers, whistle-blowers)	167
RPOs (including administrators)	111
Research integrity and research ethics bodies (REC, RIC, research councils, IRB, RIOs, RIAs, Ombudsman)	51
RFOs	41
Policy makers	41

IRB institutional review board; REC research ethics committee; RI research integrity; RIA research integrity advisor; RIC research integrity committee; RIOs research integrity offices/officers; RFOs research funding organisations; RPOs research performing organisations

Discussion

Our scoping review identified a number of available practices for the improvement of RI and research in general at RPOs and RFOs. Most of these practices were related to RPOs, in the form of guidelines, and addressed the RI topics related to the processes we categorised in our study as 'RI violations and resolutions', as well as 'RI promotion'. The fact that only a small number of identified practices were related to RFOs shows the differences regarding RI in the context of different types of organisations. While the majority of identified RI practices were developed for research in general and could be applicable across different scientific fields, a small number of disciplinary-tailored guidance for fostering RI was identified in the natural sciences (including engineering), social sciences, and humanities. Besides practices that could be applicable across various disciplinary fields, this review showed that a substantial amount of RI practices were explicitly developed for medical sciences (including biomedicine).

While most of the practices were more related to RPOs, the gap in knowledge on RI guidance was noticed in the number of identified practices for RFOs. Some guidance documents, which were mapped as those related more to RPOs, briefly mentioned funders as important stakeholders in the research process. However, there were only a few examples of practices related solely to RFOs and their specific initiatives in fostering RI. In the context of RI, this can be problematic because RFOs, together with RPOs, play an important role in influencing researchers' good or bad scientific behaviour (NASEM 2017). Although researchers build their career within RPOs and their behaviour is often influenced by organisational climate and policies, RFOs can impose additional safeguards if RPOs fail to promote and protect the integrity of their research. Usually, these measures by RFOs are aimed at RPOs rather than individual researchers although in some cases RFOs and researchers have a direct relationship (for example when setting out calls for funding, selecting certain projects to fund and monitoring funded projects). However, by demanding the establishment of RI promotion policies and procedures from RPOs, RFOs indirectly impact also the behaviour of individual researchers (Bouter 2018). Some of the important requests that RFOs may impose to RPOs for safeguarding RI may include a request for implementation of clear procedures for handling research misconduct or request for

compliance with principles of open science and transparency in research publications (Bouter 2016).

Furthermore, the analysis of stakeholders at whom the documents were aimed showed that although a large number of practices addressed RPOs (organisational directors, managers and boards), most practices addressed individual researchers. A small number of guidance documents was directed for RI structures such as RI offices, committees or advisors. This could be because many organisations still do not have specific bodies appointed to deal with RI issues; instead, RI issues are handled by ethics committees (Marušić 2019). Additionally, research processes and RI topics analysis showed that efforts to establish RI bodies are emphasised as an important role of policymakers and organisational management, but mostly in newer documents dating after 2010.

The finding that most practices for RI promotion originated from the United States may be due to our methodology, which included the search of the United States (US) Office of Research Integrity (ORI) website and the publications of the US National Academies of Sciences, Engineering, and Medicine. A large number of identified guidance documents were also from the United Kingdom, which could be because our grey literature search was limited to documents written in English. Moreover, many documents were identified as ones that could be applied widely, i.e. internationally, regardless of the country-specific differences. These documents were mostly developed as efforts of collaborative projects and international organisations that deal with RI issues. To conclude on the usage of these documents, it would be necessary to additionally explore which documents were implemented across organisations in different countries. This, of course, excludes the international legislative documents, for example those of the European Union (EU), which are mandatory for EU member states.

Guidelines were the most common form of RI guidance identified in this study. However, there was considerable variability in the topics covered and the level of elaboration presented in different guidelines. Some guidelines were focused on a single RI issue or specific stakeholders and described the specific procedures in detail, for example on data management (Science Europe 2018) or how to respond to misconduct in research (MEXT 2014). Others presented various RI issues in a more general manner with the addition of specific recommendations and were aimed toward different stakeholders (NESH 2016, NASEM 2017). Only one of the guidance documents in this study was in the form of SOPs (n = 1). Although RPOs and RFOs probably have SOPs for different kinds of administrative issues, in this study we focused on the SOPs for RI, which perhaps not all research organisations have and our study suggests that SOPs focused on RI might be rare. Further, another reason may be that RPOs and RFOs do not publish their internal SOPs which may also include the SOPs related to RI issues. The approach offered by SOPs could be helpful for the initiatives supporting research organisations and researchers on their path to integrity (Bouter 2020). For example, SOPs could be developed for defining responsibilities or describing a procedure that should be performed in the same manner, such as uploading research results to a repository or the registration of research protocol. The same could be applicable for RI bodies when it comes to handling the cases of misconduct to ensure that the same

procedure, from investigation to sanctioning, was followed in each case (Lerouge and Hol 2020).

The analysis of processes and RI topics for RPOs, RFOs, and other policymakers brought up several RI issues that were emphasised across identified practices as responsibilities of those at the organisational level. Most of these practices were related to the processes of 'RI violations and resolutions' and 'RI promotion'. For the 'RI violations and resolutions' most documents were focused on RI topics related to describing processes of investigating and handling misconducts, as well as the importance on providing clear definitions of what constitutes research misconduct. For the 'RI promotion' most documents were focused on the development and implementation of RI policies and establishment of RI bodies. Providing RI training courses and education, as well as developing infrastructure for adequate data management were also mentioned in many documents as an important responsibility of research organisations. All this reflects the organisations' valuable role in creating an environment and organisational culture in which researchers will be motivated to pertain to RI principles and rules in their work (Forsberg et al. 2018; Moher et al. 2019; Lerouge and Hol 2020).

The analysis of guiding principles showed that, although the naming of the principles was not consistent through all documents, the meaning of the principles in RI perspective was mostly the same. For example, the ALLEA code emphasises the principle of 'reliability' as employing a research methodology that will help enhance the quality of research, as well as to help ensure the trust- worthiness of one's work (ALLEA 2017). In NASEM, the same guidance regarding the validity of research was described under the principle of 'accountability' (NASEM 2017). However, 'accountability' is also used to demonstrate the responsibility of researchers toward research organisations and society (NASEM 2017) which corresponds to the principle of 'accountability' as described in the ALLEA code. The principles of 'honesty' in the ALLEA code, is defined as being honest and fair in every step of the research, valuing transparency in reporting research, as well as having an unbiased approach to the research tasks (ALLEA 2017). NASEM explicitly defines two other principles besides honesty—'objectivity' and 'openness'—which emphasise avoiding biases and transferring the real results of research to the community. The principle of 'respect' by the ALLEA code is directed toward different parties involved in research, starting from other researchers and collaborators to the research participants and society. NASEM describes respect toward others involved in research by using the terms 'stewardship' and 'fairness'. The ways of emphasising 'respect' were the most diverse regarding the terms used by different documents in comparison with other main principles. The variety of principles used across documents showed an over- view of what values need to be taken into account when considering RI issues. However, general guidance might not be enough in judging research misbehaviours and principles can be used as a valuable starting point in creating more specific guidance documents.

Study Strengths and Limitations

The main strength of this study is a comprehensive literature search that encompassed both peer-reviewed documents and grey literature from various sources and was

performed according to a rigorous methodology that required documents to be screened by multiple researchers. This comprehensive search allowed us to create a library of documents containing RI practices that could be used by organisations and individual researchers in different scientific fields. It also helped us identify gaps in the currently existing practices for RPOs and RFOs and thus create opportunity for further development of RI practices and RI in general. One of the possible limitations of our study could be that we may have missed important documents during the assessment of titles and abstracts, because the information provided therein was not sufficient for the inclusion in the analysis. Besides that, we were not able to perform the search of documents from every existing RPO and RFO and our grey literature search was limited to documents in the English language only. Therefore, we can assume that there are certainly more good practices that have not been included into this study. However, expanding our search to various RPOs' and RFOs' websites, as well as to include grey literature in languages other than English would raise a question of feasibility. Furthermore, the accessibility of guidance documents on RI may be low, as was shown for 18 universities from 10 European countries (Aubert Bonn et al. 2017), meaning that the search of individual organisations' websites would not provide a comprehensive insight into the totality of the RI guidance at RPOs and RFOs. We were unable to retrieve five documents, but they dated from the nineties, so the guidance presented in them is potentially obsolete or has already been captured in contemporary documents. Since the aim of the study was to map the existing RI practices and gaps in the content of practices, we did not take into account whether there were interventions regarding the effectiveness of the identified practices.

Conclusion

Although practices for RI promotion and initiatives to improve RPOs' and RFOs' effort in fostering RI exist it seems that more initiatives are needed for funders, RI bodies, and in certain disciplinary fields. As far as the form in which RI guidance is presented is concerned, it varies from general guidance outlining the principles and values that stakeholders should follow to more specific guidance for RI issues that are procedural in nature. When dealing with the latter, researchers and other stakeholders could find SOPs, checklists, and flowcharts to be a valuable resource of RI guidance. Through a systematic and thorough literature search, we collected a significant number of documents that RPOs and RFOs could use as guidance on RI issues or as inspiration for the development of new policies. Further research to determine factors, facilitators and barriers that may influence the implementation of RI practices could additionally help RPOs and RFOs in fostering RI.

Author Contributions: RS developed the protocol for the study and performed the data screening, full-text analysis, data charting and synthesis of all data sources, interpreted the results, and wrote and edited the manuscript. KL performed the data screening and

full-text analysis of the documents from the bibliographic databases and contributed to the revision of the manuscript. IB edited the protocol, performed the full-text analysis, data charting and synthesis of the documents from the bibliographic databases, interpreted the results, and contributed to the revision of the manuscript. JT contributed to the analysis of the results and revision of the manuscript. AM developed and edited the protocol; was a third researcher for decisions in the process of screening and analysing the data, interpreted the results, revised the manuscript and supervised the research.

Data Availability: The search strategies used to obtain documents for the analysis are available in Appendices 1 and 2. The list of all documents used in analysis and data extracted from the documents is avail- able in Appendix 3. Research processes and RI topics identified across practices and list of documents in which RI topics were addressed toward organisations and policymakers are available in Appendix 4. List of documents for individual researchers, categorised by research processes and RI topics is available in Appendix 5. Comparison of fundamental principles from the ALLEA code NASEM – Fostering Integrity in Research book and matching principles found in other documents is available in Appendix 6. List of guiding principles extracted from documents is available in Appendix 7.

All appendices are available as electronic supplementary material at the Open Science Framework (OSF). The data are available at the link below: https://osf.io/byw8s/?view_only=33369ee20c0a46cd9702 f6ab8d6d1ad9.

Compliance with Ethical Standards

Conflict of interest: The authors of this study declare that they have no conflict of interest.

Protocol Registration: The protocol was registered at the Open Science Framework (OSF) under the registration of the WP3 component (Systematic review of practices and research cultures) of the SOPs4RI project, on April 11, 2019. The protocol link is available at: https://osf.io/mzgwd/.

References

All European Academies (ALLEA). (2017). European code of conduct for research integrity. https://alleaorg/code-of-conduct/. Accessed 17 June 2020.

Aubert Bonn, N., Godecharle, S., & Dierickx, K. (2017). European Universities' guidance on research integrity and misconduct. Journal of Empirical Research on Human Research Ethics : JERHRE, 12(1), 33–44. https://doi.org/10.1177/1556264616688980.

Begley, C. G., & Ioannidis, J. P. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. Circulation Research, 116(1), 116–126. https://doi.org/10.1161/CIRCR ESAHA.114.303819.

Boeheme, O., Foeger, N., Hiney, M., Peatfield, T., & Petiet, F. (2016). Research integrity practices in science Europe member organisations. Science Europe: Science Europe Working Group on Research Integrity. https://www.scienceeurope.org/our-resources/research-integrity-practices-in-science-europe-member-organisations/. Accessed 15 June 2020.

Bouter, L. (2016). What funding agencies can do to prevent sloppy science. http://www.euroscientist. com/what-fundingagencies-and-journals-can-do-to-prevent-sloppy-science/. Accessed 30 September 2020.

Bouter, L. (2018). Fostering responsible research practices is a shared responsibility of multiple stakeholders. Journal of Clinical Epidemiology, 96, 143–146. https://doi.org/10.1016/j.jclin epi.2017.12.016.

Bouter, L. (2020). What research institutions can do to foster research integrity. Science and Engineering Ethics. https://doi.org/10.1007/s11948-020-00178-5.

DuBois, J. M. (2004). Is Compliance a professional virtue of researchers? Reflections on promoting the responsible conduct of research. Ethics & Behavior, 14(4), 383–395. https://doi.org/10.1207/s1532 7019eb1404_8.

European Commission (EC) [Internet]. Research and Innovation Observatory-Horizon 2020 Policy Sup- port Facility. Mutual Learning Exercise on Research Integrity. https://rio.jrc.ec.europa.eu/polic y-support-facility/mle-research-integrity.

European Commission (EC) [Internet]. The Community Research and Development Information Service (CORDIS). https://cordis.europa.eu/en.

European Network of Research Integrity Offices (ENRIO) [Internet]. http://www.enrio.eu/.

European Science Foundation (ESF). (2000). Good scientific practice in research and scholarship. http://archives.esf.org/fileadmin/Public_documents/Publications/ESPB10.pdf. Accessed 16 June 2020.

Fishman, T. (Ed.). (2014). The fundamental values of academic integrity (2nd ed.). Clemson: International Center for Academic Integrity, Clemson University.

Forsberg, E., Anthun, F. O., Bailey, S., Birchley, G., Bout, H., Casonato, C., et al. (2018). Working with research integrity—Guidance for research performing organisations: The Bonn PRINTEGER statement. Science and Engineering Ethics, 24, 1023–1034. https://doi.org/10.1007/s11948-018-0034-4. 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. https://doi.org/10.1177/1556264614540594

Hermerén, G., Zwart, H., Marušić, A., & Fanelli, D. (2019). European Commission. Research and Innovation Observatory-Horizon 2020 Policy Support Facility. MLE on Research Integrity: Final report. https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE%20RI_Final%20Report_0.pdf. Accessed 29 September 2020.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. Nature, 520(7548), 429–431. https://doi.org/10.1038/520429a.

Institute of Medicine and National Research Council. (2002). Integrity in Scientific Research: Creating an Environment That Promotes Responsible Conduct. Washington, DC: The National Academies Press. https://doi.org/10.17226/10430.

Institute of Scientific and Technical Information—Open Grey [Internet]. http://www.opengrey.eu/. Accessed 16 June 2020.

Inter Academy Council (IAC), & Inter Academy Partners (IAP). (2012). Responsible Conduct in the Global Research Enterprise: A Policy Report. https://www.interacademies.org/publication/responsible-conduct-global-research-enterprise. Accessed 17 June 2020.

Kalichman, M. (2013). A brief history of RCR education. Accountability in Research, 20(5–6), 380–394. https://doi.org/10.1080/08989621.2013.822260.

Komić, D., Marušić, S. L., & Marušić, A. (2015). Research integrity and research ethics in professional codes of ethics: Survey of terminology used by professional organizations across research disciplines. PLoS ONE, 10(7), e0133662. https://doi.org/10.1371/journal.pone.0133662.

League of European Research Universities (LERU) [Internet]. https://www.leru.org/about-leru. Accessed 16 June 2020.

Lerouge, I., & Hol, A. (2020). Towards a research integrity culture at universities: From recommendations to implementation. LERU publications. https://www.leru.org/files/Towards-a-Research-Integ rity-Culture-at-Universities-full-paper.pdf. Accessed 16 June 2020.

Marušić, A. (2019). European Commission. Research and Innovation Observatory-Horizon 2020 Policy Support Facility. MLE on Research Integrity: Thematic Report No 1 - Processes and structures. https://rio.jrc.ec.europa.eu/policy-support-facility/mle-research-integrity. Accessed 16 June 2020.

Marušić, A., Katavić, V., & Marušić, M. (2007). Role of editors and journals in detecting and preventing scientific misconduct: Strengths, weaknesses, opportunities, and threats. Medicine and Law, 26(3), 545–566.

Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT). (2014). Guidelines for Responding to Misconduct in Research. https://www.mext.go.jp/a_menu/jinzai/fusei/1359618.htm. Accessed 17 June 2020.

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V. et al. (2019). The Hong Kong principles for assessing researchers: Fostering research integrity. https://doi.org/10.31219/osf.io/ m9abx. Accessed 17 June 2020.

National Academies of Sciences, Engineering, and Medicine (NASEM) [Internet]. https://www.natio nalacademies.org/home. Accessed 17 June 2020.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2017). Fostering integrity in research. Washington, DC: The National Academies Press. https://doi.org/10.17226/21896. Accessed 17 June 2020.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2018). Open science by design: realizing a vision for 21st century research. Washington, DC: The National Academies Press. https://doi.org/10.17226/25116. Accessed 17 June 2020.

Nobel, J. J. (1990). Comparison of research quality guidelines in academic and nonacademic environments. Journal of the American Medical Association, 263(10), 1435–1437. https://doi.org/10.1001/ jama.1990.03440100159023.

Peters, M. D., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. International journal of evidence-based healthcare, 13(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050.

Resnik, D. B., & Shamoo, A. E. (2017). Fostering research integrity. Accountability in Research, 24(6), 367–372. https://doi.org/10.1080/08989621.2017.1334556.

Science Europe [Internet]. https://www.scienceeurope.org/.

Science Europe. (2018). Practical guide to the international alignment of research data management. https://www.scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research- data-management/. Accessed 17 June 2020.

Shamoo, A. E., & Resnik, D. B. (2015). Responsible conduct of research (3rd ed.). New York, NY: Oxford University Press.

Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. Science and Engineering Ethics, 12(1), 53–74. https://doi.org/10.1007/pl00022268.

The National Committee for Research Ethics in the Social Sciences and the Humanities (NESH). (2016). Guidelines for research ethics in the social sciences, humanities, law and theology (4th ed.). Oslo: The Norwegian National Research Ethics Committees. https://www.etikkom.no/en/ethical-guide lines-for-research/guidelines-for-research-ethics-in-the-social-sciences--humanities-law-and-theology/. Accessed 17 June 2020.

The Office of Research Integrity (ORI) [Internet]. https://ori.hhs.gov/.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Kastner, M., et al. (2016). A scoping review on the conduct and reporting of scoping reviews. BMC Medical Research Methodology, 16, 15. https://doi.org/10.1186/s12874-016-0116-4.

Wellcome Trust. (2018). Good research practice guidelines. https://wellcome.ac.uk/grant-funding/guidance/good-research-practice-guidelines. Accessed 18 June 2020.

World Conferences on Research Integrity [Internet]. https://wcrif.org/.

World Health Organization (WHO). (1995). Guidelines for good clinical practice (GCP) for trials on pharmaceutical products. http://www.femhirb.org/content_pages/files_add/doc_arb/I01_97120 11000.pdf

Zwart, H., & ter Meulen, R. (2019). Addressing research integrity challenges: from penalising individual perpetrators to fostering research ecosystem quality care. Life Sciences Society and Policy, 15, 5. https://doi.org/10.1186/s40504-019-0093-6.

Authors and Affiliations

Rea Ščepanović[1], Krishma Labib[2], Ivan Buljan[1], Joeri Tijdink[2,3], Ana Marušić[1]

1   School of Medicine, Department of Research in Biomedicine and Health, University of Split, Split, Croatia
2   Department of Ethics, Law and Humanities, Amsterdam University Medical Centers, Amsterdam Public Health Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
3   Department of Philosophy, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Chapter 3

3. Important Topics for Fostering Research Integrity by Research Performing and Research Funding Organizations - A Delphi Consensus Study

Abstract

To foster research integrity (RI), it is necessary to address the institutional and system-of-science factors that influence researchers' behavior. Consequently, research performing and research funding organizations (RPOs and RFOs) could develop comprehensive RI policies outlining the concrete steps they will take to foster RI. So far, there is no consensus on which topics are important to address in RI policies. Therefore, we conducted a three round Delphi survey study to explore which RI topics to address in institutional RI policies by seeking consensus from research policy experts and institutional leaders. A total of 68 RPO and 52 RFO experts, representing different disciplines, countries and genders, completed one, two or all rounds of the study. There was consensus among the experts on the importance of 12 RI topics for RPOs and 11 for RFOs. The topics that ranked highest for RPOs concerned education and training, supervision and mentoring, dealing with RI breaches, and sup- porting a responsible research process (e.g. through quality assurance). The highest ranked RFO topics concerned dealing with breaches of RI, conflicts of interest, and setting expectations on RPOs (e.g. about educating researchers about RI). Together with the research policy experts and institutional leaders, we developed a comprehensive overview of topics important for inclusion in the RI policies of RPOs and RFOs. The topics reflect preference for a preventative approach to RI, coupled with procedures for dealing with RI breaches. RPOs and RFOs should address each of these topics in order to support researchers in conducting responsible research.

Abbreviations

RI   Research integrity
RPO   Research performing organizations
RFOs   Research funding organizations
QRPs   Questionable research practices

Background

Fostering research integrity (RI) is important to ensure trustworthy research (Drenth et al., 2010; Horn, 2013). We see RI as a spectrum of research practices, with serious misconduct (e.g. fabrication, falsification, plagiarism) found on one end of the extremes, responsible research practices (i.e. research conducted according to high professional standards (Boehme et al., 2016)) found on the other end, and question-

able research practices (QRPs, e.g. hypothesizing after the results are known) found in between (Steneck, 2006). RI is influenced by multiple factors, including (i) the virtuousness of individual researchers, (ii) the institutional research climate, and (iii) the wider system-of-science (e.g. incentive structures) (All European Academies, 2017; Joynson & Leyser, 2015; Rifai et al., 2019; Titus & Bosch, 2010). To foster RI, it is important to consider each of these factors (All European Academies, 2017; Forsberg et al., 2018; Joynson & Leyser, 2015; Rifai et al., 2019; Titus & Bosch, 2010). While most researchers could be considered well-intentioned, motivated by a desire to improve their understanding of the world, and striving to conduct research with integrity (Joynson & Leyser, 2015), they might be deterred from engaging in responsible research practice when their institutional environment does not sup- port them sufficiently, or when they are faced with perverse incentives, for instance related to the funding system (Joynson & Leyser, 2015; Titus & Bosch, 2010). Currently, many RI initiatives focus on addressing individual researchers' responsibilities in conducting research responsibly (e.g. by setting requirements on individual researchers) (Zwart & ter Meulen, 2019). However, knowledge on the institutional and system-of-science factors influencing RI is still limited (Richman & Richman, 2012), and further research is needed to understand and tackle these factors (Bruton et al., 2020; Council of the European Union, 2015).

Various stakeholders such as research performing organizations (RPOs) (e.g. universities, independent research institutes, contract research organizations, etc.), research funding organizations (RFOs), journals, national policy makers, and publishers influence the institutional and system-of-science factors of RI (Bouter, 2018; Hermeren et al., 2019). Of these, RPOs and RFOs are particularly interesting, because RPOs have a direct impact on the institutional research climate (VSNU, 2018), while RFOs have a direct impact on elements within the system-of-science (e.g. incentive structures) (Titus & Bosch, 2010). RFOs can also have an indirect impact on the institutional research climate, since they have the means to influence institutional policies of RPOs by setting funding requirements (Tereskerz & Mills, 2012). By addressing the responsibilities of RPOs and RFOs regarding RI, it is possible to tackle some of the institutional and system-of-science factors influencing RI. While existing documents provide RPOs and RFOs with aspirational principles to follow to foster RI (e.g. All European Academies, 2017), there is a lack of concrete guidance available on how to implement these principles in practice (Mejlgaard et al., 2020). An RI policy containing a comprehensive set of concrete infrastructures, trainings, and support systems aimed at fostering RI can provide RPOs and RFOs with the means to apply aspirational RI principles to practice (Bouter, 2020; Lerouge & Hol, 2020). While many institutions globally have begun to implement various initiatives and policies on different aspects of RI (Mejlgaard et al., 2020), they often lack a comprehensive plan that addresses RI systematically. This is why the European Union's next Horizon Framework program asks institutions receiving funding to state that they have a comprehensive RI plan (Mejlgaard et al., 2020).

The first step to developing a comprehensive RI policy at RPOs and RFOs is to identify which topics to include. For instance, the Bonn-Printeger statement lists several topics that RPOs should address to foster RI, such as providing RI education, improving the organizational research culture, protecting whistle-blowers, etc.

(Forsberg et al., 2018). Similarly, the International Funders' Collaboration 'Ensuring value in research' highlights several elements related to RI that RFOs should address, such as research design, and reporting (EViR Funders' Forum, 2020). Although this shows that several RI topics have been identified as important in various national or international documents (Bruton et al., 2020), there is currently no European level consensus among research policy experts and institutional leaders about which topics should be included in the RI policies of RPOs and RFOs. In this study, we used a Delphi survey method to fill this gap. Our first objective was to explore what to address in the institutional RI policies of RPOs and RFOs by seeking consensus from research policy experts and institutional leaders on which RI topics are important for RPOs and RFOs. After achieving this objective, we set an additional second objective: to, with the experts' input, rank the RI topics in prior- ity to identify which topics should be included first in RI policies. Since RPOs and RFOs likely influence RI in different ways (Tereskerz & Mills, 2012; Titus & Bosch, 2010; VSNU, 2018), we divided the study into two parts, of which Part 1 focused on RPOs and Part 2 focused on RFOs.

Methods

The key characteristics of a Delphi study include: (1) recruiting an anonymous panel of experts (Diamond et al., 2014), (2) sending multiple rounds of surveys to the panel (Pare et al., 2013), (3) providing feedback to experts in between rounds, based on the results of the previous round (Pare et al., 2013), and (4) seeking out experts' views on a specific topic (Keeney et al., 2006). As such, Delphi studies use a structured and anonymous data collection process on a purposive sample of experts' views over several rounds of questionnaires with the purpose of informing decision making (Brady, 2015). Delphi studies do not aim at creating new knowledge, but rather on organizing and structuring existing knowledge based on experts' views (Powell, 2003). While there are numerous forms of Delphi studies available (Keeney et al., 2001; Powell, 2003), we used the 'modified' Delphi approach, starting with a document search before constructing and sending the first questionnaire round to make use of existing literature and minimize time demands on the recruited experts (Brinkman et al., 2018). We used both qualitative and quantitative measures in the Delphi study to ensure that experts had sufficient room to make suggestions and provide comments, and we strongly relied on the qualitative data to interpret the quantitative results. The study was conducted under the guidance of a Delphi expert (LM). The methods have also been described in the preregistered study proto- col (https://osf.io/ne85b/) and deviations from study protocol have been added later (https://osf.io/bcjyu/).

Document search

National and international European RI policy documents were identified to see which RI issues have already been addressed by RPOs' and RFOs' policies, using the following search terms on Google: '(research integrity OR research ethics) AND ([Country or 'Europe']) AND (guidelines OR codes of conduct)', followed by an exploration of links found on relevant pages. Additionally, using the search terms '(research institution OR university) AND [Country]' (RPOs) and 'research funding AND [country]' (RFOs),

followed by an exploration of links found on relevant pages, institution-specific policy documents (e.g. standard operating procedures, guidelines, codes, policy statements) were searched from 1 RPO and 1 RFO in each country in the European Research Area (proceeding through countries in alphabetical order). RI issues (e.g. 'RI education' and 'data management facilities') were extracted from the identified documents. The search for documents ended once saturation was reached for data extraction (i.e. when the same issues kept repeating in subsequent documents and no new issues arising) (Saunders et al., 2018). Based on the issues extracted, and considering overlap and the relationship between issues, two preliminary lists of topics and subtopics were created for 1) RPOs and 2) RFOs. KL was responsible for the document search, extraction of issues, and creation of the preliminary lists of topics. After discussion of the topic lists among all authors, the lists were refined further (e.g. altering phrasing, adding additional topics, etc.).

Participant selection

The inclusion criteria for study participants (i.e. the 'experts') was: people with (1) experience in research policy, and (2) working at RPOs or RFOs. A purposive sampling technique was used to identify experts and consisted of simultaneously approaching personal contacts followed by snowballing, and performing a web search of contacts at RPOs and RFOs across Europe. For the web search of RPO experts, we looked for RI or research policy contacts by browsing the website of three RPOs in each European country; we selected the three RPOs that appeared first in a Google search of '(research institution OR university) AND [Country]'. For the RFO experts, we searched the website of at least one RFO in each European country by searching for 'research funding AND [country]' on Google. In many countries, only a single RFO could be found; in the case that we were able to identify multiple national RFOs using that search method, we included experts from additional RFOs from the country. We supplemented this search strategy by looking for the contact details of authors of documents we identified in our document search. All experts identified were invited to participate in the study. The experts' identities remained anonymous to all until study completion, except KL and JT, who were responsible for selection and correspondence.

Procedures

Our study consisted of two parts, each with three online Qualtrics survey rounds, as depicted in Fig. 1. In each survey round, experts were provided with an updated description of each topic and subtopic presented, both in a separate PDF (available at https://osf.io/jc6u2/ for RPOs, and https://osf.io/82dwk/ for RFOs), as well as in popup text included in the survey when the topics or subtopics were mentioned. We refined descriptions each round by incorporating the input from the experts from the previous round.
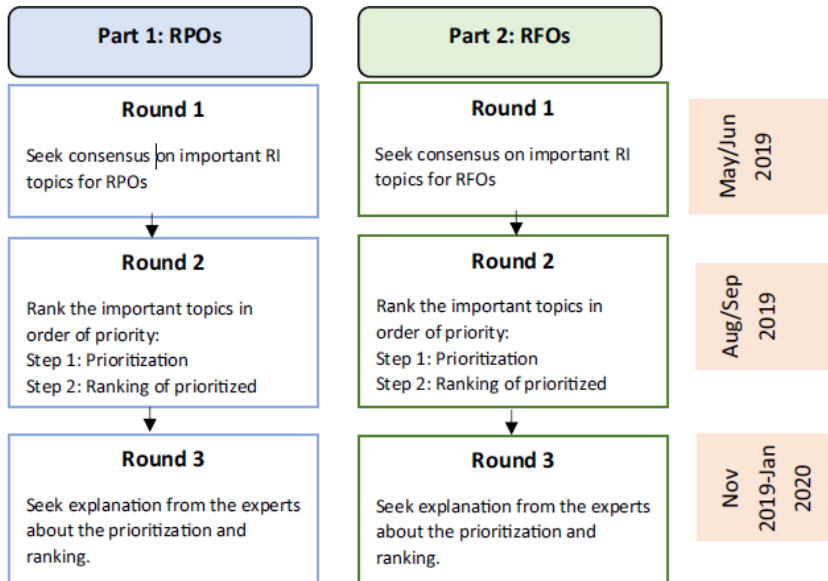
Fig. 1 *The Delphi procedure*

Pilot tests

Before we sent the Round 1 surveys to all experts, we decided to first pilot them with 2 experts from RPOs and 3 experts from RFOs to ensure the clarity of concepts and questions. The pilot experts were personal contacts who met the study inclusion criteria and showed enthusiasm in the study prior to being invited to participate. Based on the feedback of the pilot experts, we made some final adjustments in the survey (e.g. rephrasing some questions) and refined the lists of topics further (e.g. renaming topics). More information on this can be found in Online Resource 1.

Round 1

In Round 1, we asked experts to rate each topic in the preliminary list on a 1–5 Likert rating scale, ranging from the topic is 'not important at all', to it is 'absolutely essential' for institutional efforts in fostering RI. When experts rated a topic 3 or higher, they were also able to rate the subtopics we had identified under that topic, by selecting to 'Include' or 'Do not include' the subtopic. We randomized the order of the topics, to control for any biases that could occur due to the order of the topic presentation. We used a forced response feature on the survey when asking the experts to rate the topics; we did not force the experts to rate the subtopics. Experts were encouraged to provide arguments for their ratings in open answer options, as well as to suggest new topics/subtopics. The questionnaires of Round 1 can be found on the Open Science Framework (OSF) (RPOs: https://osf.io/w78bj/ and RFOs: https://osf.io/gp5jt/).

Round 2

Since consensus on the topics was already reached in Round 1, in Round 2, we asked experts to complete a prioritization and a ranking exercise in two steps. In Step 1, experts were asked to select half the topics that achieved consensus on importance in Round 1, to prioritize. In Step 2, they ranked the prioritized topics in order of priority. Additionally, experts were asked to rate the subtopics which had not achieved consensus in Round 1, as well as newly proposed subtopics. We encouraged the experts to provide comments on their ratings, and prioritization and ranking choices. The links to the RPO and RFO questionnaires of Round 2 are https://osf.io/wtu6r/ and https://osf.io/5j642/, respectively.

Round 3

After analyzing the results of Round 2, we were unsure about whether experts' prioritization and ranking were motivated by considerations of (1) the feasibility of creating institutional policies on each topic; (2) the impact that topics could have on research practice; (3) the need for RPOs and RFOs to address certain topics; or some other rationale. Therefore, we used a feedback round, which can be used as a member check to increase rigor in Delphi studies (Brady, 2015), as a third exploratory round, in which we asked experts to share their thoughts on what considerations might underlie the ranking of each topic. The RPO questionnaire of Round 3 can be found here: https://osf.io/qmw94/, while the RFO questionnaire can be found here: https://osf.io/48n97/.

Data Analysis

Quantitative

Rating of Topics

To analyze the responses of the rating exercise in Round 1, we had originally defined consensus as agreement among 2/3 of the experts (67%) on ratings 3–5 (moderately important, very important, absolutely essential) per topic and ratings of 'Include' per subtopic. However, the threshold for the topics did not allow us to see differences since all topics were deemed at least moderately important by > 80% of the experts. Therefore, we retrospectively raised the threshold for consensus to 67% agreement on ratings 4–5 (very important—absolutely essential). To determine whether to include or exclude a topic or subtopic, we considered whether consensus had been achieved. We excluded responses from experts who completed less than 51% of the topic rating questions from the analysis, as they had not completed an assessment of all topics and subtopics, making it difficult to interpret their results.

Prioritization and Ranking of Topics

To create a ranked list of topics, we analyzed the prioritization and the ranking exercise results from Round 2 in three phases, A, B and C:

A.   We looked at how often each topic was prioritized in Step 1 of the prioritization and ranking exercise.

B.   We calculated the total ranking score per topic. To do this, we had to: (1) assign a ranking number per topic per expert, and then (2) per topic, sum the multiple of each ranking number with the number of experts assigned to it. An example of how the ranking scores were calculated can be found in Online Resource 6 (p. 4). We used the following procedure to assign a ranking number per topic per expert. For the prioritized topics we assigned a number of 1 to the topic ranked lowest in Step 2 of the ranking and prioritization exercise by the expert, with each higher ranked topic receiving a number 1 point above that. We assigned a score of -2,5 for each topic not prioritized by the expert in Step 1 of the ranking and prioritization exercise. Since these non-prioritized topics were not ranked relative to each other, we had to use the same ranking score for all the non-prioritized topics per expect; the ranking score of -2,5 that we used was the mean between the values of 0 and -5 that they would have received had they been ranked relative to each other.

C.   We looked at the results of phase A to place the most frequently prioritized topics (i.e. those selected by 50% or more of the experts) higher in the final ranked list of topics than those prioritized less frequently. To see in which order to place the most frequently prioritized topics relative to each other, we used the results of phase B; topics with a higher total ranking score were placed higher in the ranked list. We also compared the ranking scores of the less frequently prioritized topics to order them relative to each other.

Qualitative

We looked at experts' arguments for and against the importance of each topic/ subtopic. We used the arguments per topic/subtopic to refine the topic/subtopic descriptions and make them more adequate and specific, as well as to improve proposals in the second round (e.g. if after discussion among the authors, we agreed that an argument for excluding a topic was convincing, we proposed to exclude the topic regardless of the quantitative results). Additionally, when making final decisions about the inclusion and exclusion of the topics, we checked the qualitative data from Rounds 1–3 to see whether experts' views aligned with the quantitative results. In addition to analyzing the qualitative data per topic/ subtopic, we noticed some general patterns regarding experts' views on RI policy across topics. Thematic analysis was used to explore these general considerations (by KL) (Brady, 2015). To check the reliability of the codes, 25% of the data was also analyzed by an independent coder (RR). Discrepancies were discussed by the two coders (KL and RR) to come to agreement; in case of disagreements, a third coder was consulted to reach this agreement (JT). Based on the discussions of the discrepancies, the first coder (KL) rechecked and adjusted the codes of the 75% of the data that were not second coded.

Results

Document Search

After searching for national and international policy documents on RI in Europe, 10 documents (i.e. codes of conduct and guidelines) were identified. Addition- ally, 18 RPO and 14 RFO institution-specific policy documents (e.g. policy statements, guidelines, etc.) were analyzed from which to extract issues. The decision that data saturation was reached was made while extracting issues from RPO and RFO documents from Poland; therefore no documents from subsequent countries (i.e. those following Poland in alphabetical order) were searched for. An overview of these documents can be found in Online Resource 2. In total, 164 issues for RPOs and 64 issues for RFOs were extracted from the documents (by KL). By removal of duplicates (126 RPO issues and 28 RFO issues), combining content, assessing overlap and the relationship between topics, the extracted issues were captured in 13 topics and 34 subtopics for RPOs, and 11 topics and 18 subtopics for RFO (by KL). The topic lists were discussed among the authors to further refine them (e.g. rename topics, regroup topics, remove, and add topics), which resulted in 14 topics and 36 subtopics for RPOs, and 11 topics and 27 subtopics for RFOs (Online Resource 3).

Results From the Delphi Rounds

The datasets generated from the Delphi rounds can be found on OSF (https://osf.io/3quj6/).

Response Rate and Respondent Characteristics

A total of 305 RPO experts and 215 RFO experts were invited to participate in Rounds 1 and 2, while only responders of Round 1 or 2 or both could participate in Round 3. The response rate for the RPO study was 17% (51/305), 18% (53/305) and 52% (35/68) in Rounds 1, 2 and 3, respectively. For the RFO study, the response rate was 18% (39/215), 12% (37/215) and 46% (24/52) in Rounds 1, 2 and 3, respectively. About half the respondents (47% for RPOs and 52% for RFOs) were personal contacts, or contacts found through snowballing. The demographic data showed diversity in the respondents in terms of gender, country, and the disciplinary background of their organization (Online Resource 4). About half of the participants identified themselves as female (44% RPO and 52% RFO experts) and were employed in Northwestern Europe/Scandinavia (46% RPO and 42% RFO experts). The mean number of years of experience in research policy was 14 for both the RPO and RFO experts. A large majority (> 96%) considered themselves at least moderately experienced in RI issues.

Important Topics for RI Policy

Of the included 14 topics for RPOs and 11 for RFOs, we achieved consensus on the importance of all but two RPO topics (Fig. 2). Due to a lack of consensus on its importance (59% agreement), and the argument that only RFOs have power to set requirements on RPOs, we excluded the RPO topic 'Relationship between RPOs and

RFOs'. Additionally, we renamed the RFO topic 'Relationship between RPOs and RFOs' to 'Funders' expectations of RPOs' to reflect experts' views that the relationship between RPOs and RFOs is mostly unidirectional. Similarly, we excluded the RPO topic 'Societal involvement in research' based on a lack of consensus on its importance (55% agreement) and concerns that the topic is discipline-specific and controversial with one participant stating "If you write a SOP [i.e. a policy document about this topic] you have to be compliant. Is that what you want?" (Round 1, RPO study) and another mentioning that the topic is "not a 'must' criteria since not all research might be relevant to this target group [i.e. research stakeholders and policy makers/public authorities]" (Round 1, RPO study). More information about the ratings of the topics and subtopics can be found in Online Resource 5.
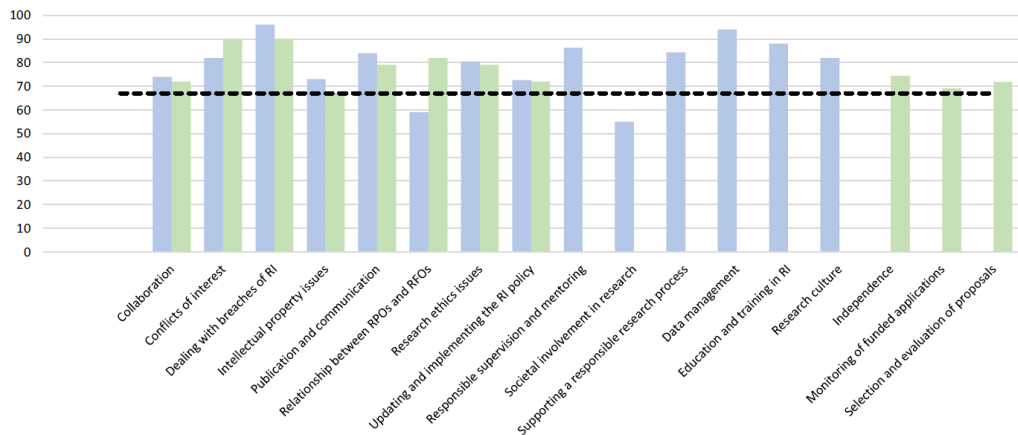
Fig. 2 *Agreement on importance of RI topics. The x-axis represents the topics that were presented to the experts. The bars in blue indicate the RPO topics, whereas the bars in green indicate the RFO topics. The y-axis represents the percentage of experts who rated each topic 4–5 on importance on the 5 point Likert scale (i.e. very important-absolutely essential). The black dotted line shows the consensus cut-off value of 67% agreement*

Prioritization of RI Topics

The final ranked lists of RPO and RFO topics, based on the results of Phase C of the analysis of the prioritization and the ranking exercise in Round 2, can be found in Table 1. The topics' frequency of prioritization (Phase A) and the ranking score per topic (Phase B) can be found in Online Resource 6. The complete ranked list, including the subtopics, can be found in Online Resource 7.

Table 1 Ranked list of RI topics

Rank	Topic	Consensus on importance (% agreement)
For inclusion in the RI policy of RPOs		
1	Education and training in RI	Yes (88%)
2	Responsible supervision and mentoring	Yes (86%)
3	Dealing with breaches of RI	Yes (96%)
4	Supporting a responsible research process	Yes (84%)
5	Research ethics issues	Yes (80%)
6	Data management	Yes (94%)
7	Conflicts of interest	Yes (82%)
8	Research culture	Yes (82%)
9	Publication and communication	Yes (84%)
10	Updating and implementing the RI policy	Yes (73%)
11	Intellectual property issues	Yes (73%)
12	Collaborative research among RPOs	Yes (74%)
–	*Relationship between RPOs and RFOs*	*No (59%)*
–	*Societal involvement in research*	*No (55%)*
For inclusion in the RI policy of RFOs		
1	Dealing with breaches of RI	Yes (90%)
2	Conflicts of interest	Yes (90%)
3	Funders' expectations of RPOs	Yes (82%)
4	Selection & evaluation of proposals	Yes (72%)
5	Research ethics issues	Yes (79%)
6	Collaboration	Yes (72%)
7	Monitoring of funded applications	Yes (69%)
8	Updating and implementing the RI policy	Yes (72%)
9	Independence	Yes (74%)
10	Publication	Yes (79%)
11	Intellectual property issues	Yes (67%)

Qualitative Results

Overall, the quantitative results (ratings and rankings) were reflected in experts' qualitative responses. The qualitative responses from each Delphi round also helped to refine the descriptions of the topics and subtopics presented to the experts in the next round (please see Online Resource 8 to see the evolution of the topic descriptions). These changes were deliberated improvements or additions to the descriptions, rather than major deviations. An example of a definition that changed across rounds was the subtopic 'Secure data collection, storage, retention, archiving and sharing infrastructure' falling under the topic 'Data management' in the RPO part of the study. Initially we had named this subtopic 'Secure data storage infrastructure', but the experts recommended adjusting this name to clarify that the subtopic includes

additional stages of the data management process than just storage. There- fore, we adjusted the name and description of the subtopic.

A summary of experts' comments on the importance of each topic can be found in Online Resource 9, while a summary of the results of Round 3 can be found in Online Resource 10. Furthermore, we noticed that there were patterns in experts' comments about RI policy that cut across topics and could be insightful for the interpretation of the study. We used thematic analysis to explore these patterns further. Seven overarching themes emerged from experts' comments (Table 2).

Table 2: Themes identified based on qualitative data from the Delphi study

Theme	Description	Example quotes
Views on RI policy	Experts were ambiguous about the notion of RI policy. They identified risks and opportunities of developing RI policy, among which are: *Risks* - Difficult to capture topics in institutional policy documents (feasibility) - Institutional policies might not be effective *Opportunities* - Raising awareness about RI - Building a comprehensive RI system	*"One should avoid 'ethics of science in general'. This is a very wide and multifaceted world of issues. Cannot be covered satisfactorily [in research policy]"* (Round 1, RPO survey) *"[Including policies on] all of this is important, because many organisations... do not have a comprehensive approach, they only have bits and pieces... but it needs to be a system..."* (Round 1, RPO survey)
Broadness of definition of RI	Some experts advocated for a broad definition of RI, whereas others encouraged keeping the definition narrow. For instance, there was uncertainty about whether research ethics, human resource management issues (e.g. appointment of researchers) and legal issues (e.g. intellectual property) fall within the scope of RI	*"This is important but also necessary not to make RI a container term overloading it with topics... these issues are HR based but have links to RI so in core they should be handled in HR dept. with the exception of a few clear issues"* (Round 1, RPO survey) *"The narrow definition of research integrity would create an artificially constructed concept which would ignore many of the aspects that constitute integrity (the wholeness of the person, alignment of values and actions etc.)."* (Round 2, RPO survey)
Approach to RI	The experts advocated for certain approaches to RI:	*"I would like to see a preventive approach more than a punitive approach. I think education, information and communication, mentoring, could be an effective approach*

Theme	Description	Example quotes
	- Emphasis should lie on prevention, rather than punishment of breaches of RI - Dealing with breaches is complementary to prevention - It is best to focus on the causes, rather than symptoms, of breaches of RI	*to create an institutional good RI environment" (Round 3, RFO survey)* *"It is entirely possible to deal with RI breaches in a way that supports a positive preventative approach. Indeed a calm and open approach to discussing breaches is a hall mark of a good RI culture. No to witch hunts, yes to open discussion." (Round 3, RPO survey)* *"The two subjects are not mutual exclusive—prevention (by training, strong ethical research cultures, responsive supervision, etc.) is essential to fostering sound (and trustworthy) science. However, when detrimental practices are detected, institutional (and national) bodies need to be able to handle them in a clear and transparent manner"(Round 3, RPO survey)* *"Prevention is important, but ways of dealing with RM [i.e. research misconduct] create and maintain the culture of integrity of an institution." (Round 3, RFO survey)* *"The main issues relating RI and publication to both authorship and open science relate to the reward structures around current publication practices. It is the reward structures that need to change." (Round 3, RPO survey)*
Differences	Experts identified that the development of RI policy will be influenced by differences in: - Institutions - Countries - Disciplines - Time (i.e. new developments)	*"The topic 'Supporting a responsible research process' would seem to apply more to research in the biomedical sciences—less so for qualitative research in the social sciences and humanities" (Round 1, RPO survey)*
Interrelatedness of topics	Experts mentioned that many of the topics are interrelated, and that to develop a comprehensive RI policy, they should all be addressed Experts mentioned that some topics can be	*"Prioritizing does not mean neglecting other aspect of the "package". Properly functioning system should take care of all the aspects." (Round 3, RFO survey)* *"This [i.e. research culture] is foundational—underpinning all else…" (Round 3, RPO survey)* *"Research culture is an overarching concept that is influenced by all other issues…" (Round 3, RPO survey)*

Theme	Description	Example quotes
	addressed indirectly through other topics	
Autonomy	Experts highlighted that RI policy should not unnecessarily interfere with the autonomy of RFOs, RPOs and researchers	*"Many funding agencies [do] NOT check whether the research plan was followed in detail. Quite often unexpected developments force researchers to change their plan to achieve their goal. Funding agencies should not interfere in this process. On the other side: we notice that applicants submit the same proposal in various calls and when the proposal is funded twice, they change the topic of one of the proposals without notifying the funding agency. The latter action is questionable." (Round 1, RFO survey)*
Responsibility	Experts highlighted that research stakeholders have different responsibilities for RI	*"Perhaps [publication was ranked low] because publication is difficult to be dealt in the context of RPOs or RFOs but it needs the involvement of science editors and policy makers." (Round 3, RFO survey)*

Table 2 continued

Discussion

To our knowledge, this is the first consensus study which has identified what topics are considered important for fostering RI at RPOs and RFOs according to research policy experts and institutional leaders. While some of these topics are already broadly addressed in European codes of conduct on RI (e.g. All European Academies, 2017; VSNU, 2018), many of them have not yet been adequately implemented in the RI policies of RPOs and RFOs (Bouter, 2020; Lerouge & Hol, 2020). The qualitative data shows that, although there is some variation in the broadness of the definition of RI that is accepted by experts in different institutions and countries, experts prefer focusing on a positive, preventative approach to RI as opposed to a punitive approach. This preference is in line with the literature, which indicates an increasing acknowledgement that a positive approach focused on helping researchers engage in responsible research practice is more desirable in fostering RI (Research integrity is much more than misconduct, 2019; Zwart & ter Meulen, 2019). This is not surprising, given that most of the recognized problems with RI qualify as QRPs—a gray zone of research behaviors that in most cases would not be considered as research misconduct (Bouter et al., 2016). An approach to RI solely focused on punishment might not be able to tackle QRPs in organizations successfully (Iorns & Chong, 2014), because it would place too much emphasis on the individual researcher, and not enough on the institutional and system-of-science factors that influence RI (Drenth, 2015; Kumar, 2010).

      Despite the preference for a preventative approach, 'Dealing with breaches of RI' was prioritized highly for both RPOs and RFOs. The topic was seen as both feasible to address in institutional policies and urgent (i.e. cases of misconduct must be dealt with once they arise). The experts highlighted that prevention and tackling research misconduct can be complementary, since focusing on prevention does not exclude the

importance of handling misconduct cases appropriately. Prevention is arguably more likely to be effective when aimed at QRPs, than outright misconduct. Organizations need to foster an environment of openness and learning to tackle minor misbehaviors, while still holding individuals accountable for outright misconduct (Boysen, 2013). Furthermore, if a 'systems' approach to tackling misconduct is taken, where not only researchers, but also other stakeholders (including RPOs and RFOs) are held accountable for contributing towards research misconduct, institutional and system-of-science factors can still be addressed (Kumar, 2010). Additionally, as the experts indicated, dealing with research misconduct is necessary to raise awareness about RI and develop a responsible research culture (Stemwedel, 2014).

The experts in this study stressed that fostering a responsible research culture is key to fostering RI. Yet, the topic 'Research culture' was not among the highest ranked topics. However, experts mentioned that it overlapped substantially with other higher ranked topics. In fact, the most highly prioritized topics, 'Education and training in RI' and 'Responsible supervision and mentoring', are both thought to have a direct impact on research culture (e.g. Geller et al., 2010; Kalichman, 2014; Satalkar & Shaw, 2019). It could be that education and supervision are considered concrete ways for affecting research culture, since learning from mistakes is necessary to create a responsible research culture (Boysen, 2013), leading experts to prioritize them instead of the seemingly vaguer topic 'Research culture'.

Furthermore, the fact that (1)'Funders' expectations of RPOs' (i.e. funders' requirements) was ranked highly in the RFO part of the study, and (2) the topic 'Relationship between RPOs and RFOs' did not achieve consensus on importance in the RPO part of the study, shows that RPOs and RFOs have different roles and responsibilities regarding RI. Since researchers are dependent on the infrastructures and policies of RPOs, RPOs are directly responsible for supporting researchers in RI (Youngblut & Brooten, 2002). RFOs rely on RPOs for many aspects of RI promotion (e.g. the provision of appropriate data management infrastructure, training, etc.) (Tereskerz & Mills, 2012). Therefore, the relationship between RPOs and RFOs is mostly unidirectional, with RFOs imposing requirements and RPOs having to meet them. As such, while it may be of value for RPOs to have policies that address the relationship with RFOs, it might be that such policies would not be impactful enough—due to the unidirectional relationship between RPOs and RFOs—to be vital in the inclusion of their RI policy.

The RPO topic 'Societal involvement in research' also did not achieve consensus to be included in RPOs' RI policies. Experts argued that the topic is too discipline specific and controversial. This suggests that while the topic might be relevant for some types of research (e.g. fields in which public engagement and inclusion is relevant), it may not be broadly applicable to most research. As such, it may be that it is not an important enough topic to recommend to include in the RI policies of RPOs across Europe.

Strengths and Limitations

Since the study reached out to two heterogenous expert panels, each consisting of more than 50 participants representing different countries (more than 25 countries per

panel), genders, and disciplines, we were successful in reaching out to a diverse range of institutional research policy experts—at least from a geopolitical perspective. This helped to obtain consensus at the European level on a comprehensive set of important RI issues for RPOs and RFOs. Since we did not explore the racial and ethnic diversity of the experts, it could be that our findings are dominated by a white European perspective.

While only two of the RPO topics and none of the RFO topics initially presented to the experts were excluded from the final list of topics, the Delphi rounds provided us with valuable inputs which helped us to remove, add, as well as refine topics and subtopics. Additionally, the study enabled us to rank the important RI issues in priority based on the needs and gaps seen by experts in various contexts, allowing for results which are relevant for RPOs and RFOs across disciplines and countries. Although, the ranking exercise was perceived as difficult, as many experts found all the topics to be important, it helped to provide general guidance to RPOs and RFOs on where to start when developing RI policy. Additionally, our findings help to shed light on which approaches to institutional policies on RI experts prefer.

Through the Delphi method, we were able to systematically and democratically engage with the experts (Powell, 2003). This helped to obtain both quantitative and qualitative data from the respondents. Since the experts' identities remained anonymous to the other experts and researchers (except for KL and JT), we reduced biases that would occur when participants are well acquainted (e.g. higher status stakeholders dominating the discussion) (Powell, 2003). Of course, by focusing on research policy experts and institutional leaders only, we cannot confirm whether the identified topics are also deemed as important by other research stakeholders. However, other studies we have conducted with various research stakeholders (including senior and junior researchers) validate the topics that emerge from this study (Ščepanović et al., 2019; Sørensen et al., 2020).

Since Delphi studies help with structuring existing knowledge, rather than creating new knowledge, our list of topics likely present consensus about known approaches necessary for fostering RI, and might miss out on novel ways that RI can be addressed (e.g. potentially new approaches to reducing perverse incentives).

However, given that many of the topics have not yet been implemented widely across institutions in Europe, these topics can still be considered novel and promising to address. For instance, if institutions have policies in place on the subtopic 'reducing publication pressure and hyper-competition' (e.g. by rewarding researchers not only on the number of publications but also other outputs and activities), this can help to create a more responsible research culture.

We obtained a sufficient number of responses, but our response rate (12–18%) was lower than reported elsewhere (e.g. 70% in Brinkman et al., 2018; Terwee et al., 2018). This could be because studies with higher response rates base their rates on the number of experts who declare willingness to participate before receiving the study invitation (Boulkedid et al., 2011; De Villiers et al., 2005; Hamilton & Bowers, 2006; Pare et al., 2013). We sent invites to most participants directly, rather than informing them about the study beforehand, which likely explains the lower response rates. This explanation is supported by the fact that two recent Delphi studies using the same means of reporting their response rates had similar figures to ours (Haven et al., 2020;

Mokkink et al., 2020). The possibility that the response rate introduces bias into the study is unlikely to be high, given the diversity in the study participants. Since we approached the Delphi study from a primarily qualitative research methodology where the aim was to explore experts' perceptions and opinions rather than to obtain generalizable knowledge about the mean impression of certain topics, we do not expect the study validity to have been hampered by the response rate (Keeney et al., 2001).

　　　Another methodological concern in this study was the consensus threshold value of 67%, which we set based on the idea that obtaining consensus by 2/3 of the experts would be sufficient to make a well-informed consensus assessment about a topic. However, there are no set standards about how to measure consensus, nor on what threshold value to choose in Delphi studies (von der Gracht, 2012). The threshold value is to some extent arbitrary and serves an instrumental role in helping to explore differences between the importance of items in the Delphi survey. This is why, although we had initially defined consensus as 67% agreement on ratings 3–5, when we saw (based on the results of Round 1) that this consensus definition did not allow us to differentiate between important versus moderately important topics, we modified the consensus definition to 67% agreement on ratings 4–5. Moreover, in addition to relying on the quantitative data, we also examined the qualitative arguments to see whether including or excluding a topic was in line with the experts' views. If the qualitative data did not support the consensus reached, we brought up the topic/subtopic in the next Delphi round to ask experts to rate the topic/subtopic again.

Next Steps

While the ranking provides suggestions on which topics to tackle first, RPOs and RFOs will need to address most of the identified topics to build a comprehensive institutional RI policy. However, more empirical work is needed to guide RPOs and RFOs on how to build effective RI policy on the topics identified. Despite our efforts to differentiate between the topics and subtopics by providing experts with clear descriptions of each, some of the topics remain interrelated and connected to each other (e.g. research culture was thought to be both influenced by and underpinning other topics). This is not surprising considering that RI is a complex phenomenon, consisting of multiple stakeholders and factors, and interventions designed to sup- port it are often intertwined (All European Academies, 2017; Bouter, 2018; Joynson & Leyser, 2015; Rifai et al., 2019; Titus & Bosch, 2010). Therefore, it is necessary to explore the relationship and dependencies between the topics further in order to untangle them in future studies. To tackle RI effectively, it will be necessary to address the causes of breaches, rather than only the consequences (National Academies of Sciences Engineering & Medicine, 2017). Furthermore, future research is needed to explore how to prevent the risk that implementing RI policies will intro- duce unnecessary administrative burdens and a 'check-box mentality' towards RI, where RPO and RFO leaders merely address RI topics to show that they are following necessary RI developments. Such research is necessary to ensure that RI policies are instead sensitive to researchers' needs, and focus on supporting researchers to engage in responsible research practices through a positive and constructive manner. More insight is needed on how country, discipline

and institution-specific differences influence topic-specific institutional policies, especially considering that different experts across Europe view RI in narrower or broader terms. For instance, it might be that different types of institutional policies are needed for different disciplines. This will likely also influence the broadness of the definition of RI that should be applied for developing RI policy in a specific organization. Similarly, differences between countries in the availability of and approach to national RI structures and policies (Godecharle et al., 2013; Hermeren et al., 2019) might influence the way that RPOs and RFOs should address each of the topics identified in this study. For example, RPOs and RFOs in countries with no national RI guidelines or structures might have a greater role to play in building RI policies than those in countries with more established structures, as the latter can partially rely on national structures (Godecharle et al., 2013; Hermeren et al., 2019). Additionally, it is important that RPOs' and RFOs' RI policies are sensitive to the social, cultural, and historical factors present in the local context (e.g. communication styles) to internalize RI successfully in the institutional culture (Hermeren et al., 2019). Moreover, RPOs and RFOs should work together with the other important stakeholders, including researchers and journals, to jointly produce RI policies that (1) accurately reflect different stakeholders' responsibilities and needs, (2) sufficiently take into account country, discipline and institution-specific differences, (3) do not unnecessarily interfere with the autonomy of stakeholders, and 4) adequately promote responsible research practices rather than QRPs or misconduct.

Conclusions

Despite the growing awareness of the importance of RI, there is still little progress on improving the institutional and systemic factors that influence RI. Since RPOs and RFOs have an important role to play here, they should develop, implement and optimize institutional RI policies. This study has made a first step towards changing the landscape by, together with institutional policy experts, exploring which RI topics should be addressed in the RI policies of RPOs and RFOs. Tackling each topic is necessary to effectively support researchers in con- ducting responsible research.

Supplementary Information

The online version contains supplementary files 1-11, available at https://doi.org/10.1007/s11948-021-00322-9.

Acknowledgements

Authors' contributions

KL was responsible for writing and editing the study protocol, selecting and communicating with participants, creating the surveys, collecting and analyzing the data, and writing and editing the manuscript. RR was involved in analyzing the data and editing the manuscript. LB, GW, NE and LM were involved in editing the study protocol, the surveys and manuscript. AM was involved in editing the manuscript. JT was responsible for writing and editing the study protocol, selecting and communicating with participants, editing the surveys and manuscript, and supervising the research. Funding This study was funded by the European Commission HORIZON 2020 framework programme for Research and Innovation under Grant Agreement No 824481. The funder was not involved in the design, conduct, analysis or interpretation of the study.

Availability of data and materials

The datasets generated and analyzed during the current study are avail- able on the Open science framework, https://doi.org/10.17605/OSF.IO/8PZXF. Based on discussions with our privacy officer, we have removed demographic information from the publicly available data set.

Code availability: Not applicable.

Declarations

Conflict of interest: The authors have no relevant financial or non-financial interests to disclose.

Ethical approval: The study protocol was sent to the Institutional Review Board of the Amsterdam University Medical Centers (location VUmc), which decided that the study does not fall under the Dutch Medical Research Involving Human Subjects Act (Wet medische-wetenschappelijke onderzoek met mensen). Additionally, ethics approval

was obtained from the Ethics Review Board of the Vrije Universiteit Faculty of Movement and Behavioral Sciences (approval number: VCWE-2019-029).

Consent to participate: Participants were provided with an information leaflet, including details on our privacy policy, prior to commencing the study. We assumed informed consent when participants took part in the survey.

Consent for publication: Not applicable.

References

All European Academies (2017). *The European code of conduct for research integrity, revised edition. Berlin.*

Boehme, O., Föger, N., Hiney, M., Peatfiled, T., & Petiet, F. (2016). *Research integrity practices in science Europe member organisations. Science Europe, July.*

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011). *Using and reporting the Delphi method for selecting healthcare quality indicators: A systematic review. PloS one, 6(6). https://doi. org/10.1371/journal.pone.0020476.*

Bouter, L. (2020). *What research institutions can do to foster research integrity. Science and Engineering Ethics, 26(4), 2363–2369. https://doi.org/10.1007/s11948-020-00178-5.*

Bouter, L. M. (2018). *Fostering responsible research practices is a shared responsibility of multiple stake- holders. Journal of Clinical Epidemiology, 96, 143–146.*

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(17). https://doi.org/10.1186/ s41073-016-0024-5.*

Boysen, P. G. (2013). *Just culture: A foundation for balanced accountability and patient safety. Ochsner Journal, 13(3), 400–406.*

Brady, S. R. (2015). *Utilizing and adapting the Delphi method for use in qualitative research. International Journal of Qualitative Methods, 14(5). https://doi.org/10.1177/1609406915621381.*

Brinkman, D. J., Tichelaar, J., Mokkink, L. B., Christiaens, T., Likic, R., Maciulaitis, R., et al. (2018). *Key learning outcomes for clinical pharmacology and therapeutics education in Europe: A modified Delphi study. Clinical Pharmacology & Therapeutics, 104(2), 317–325. https://doi.org/10.1002/cpt. 962.*

Bruton, S. V., Medlin, M., Brown, M., & Sacco, D. F. (2020). *Personal motivations and systemic incentives: Scientists on questionable research practices. Science and Engineering Ethics. https://doi.org/ 10.1007/s11948-020-00182-9.*

Council of the European Union (2015). *Draft council conclusions on research integrity. Brussels.*

De Villiers, M. R., De Villiers, P. J., & Kent, A. P. (2005). The Delphi technique in health sciences education research. *Medical Teacher, 27(7)*, 639–643. https://doi.org/10.1080/13611260500069947.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., et al. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology, 67(4)*, 401–409. https://doi.org/10.1016/j.jclinepi.2013. 12.002.

Drenth, P. J. (2015). Institutional dealing with scientific misconduct. *Eruditio, 1(6)*, 136–146.

Drenth, P., Ftacnikova, S., Hiney, M., & Puljak, L. (2010). *Fostering research integrity in Europe.* Strasbourg.

EViR Funders' Forum Guiding Principles. https://sites.google.com/view/evir-funders-forum/guiding- principles. Accessed April 4, 2020.

Forsberg, E.-M., Anthun, F. O., Bailey, S., Birchley, G., Bout, H., Casonato, C., et al. (2018). Working with research integrity—Guidance for research performing organisations: The Bonn PRINT- EGER Statement. *Science and Engineering Ethics, 24*, 1023–1034. https://doi.org/10.1007/ s11948-018-0034-4.

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. https://doi.org/10. 1097/ACM.0b013e3181e5f0e5.

Godecharle, S., Nemery, B., & Dierickx, K. (2013). Guidance on research integrity: No union in Europe. *The Lancet, 381(9872)*, 1097–1098.

Hamilton, R. J., & Bowers, B. J. (2006). Internet recruitment and e-mail interviews in qualitative studies. *Qualitative Health Research, 16(6)*, 821–835. https://doi.org/10.1177/1049732306287599.

Haven, T., Errington, T., Gleditsch, K., van Grootel, L., Jacobs, A. M., & Kern, F., et al. (2020). Preregistering qualitative research: A Delphi study.

Hermeren, G., Zwart, H., Marusic, A., & Fannelli, D. (2019). *Mutual Learning Exercise (MLE) on research integrity - Final report.* European Commission.

Horn, L. (2013). Promoting responsible research conduct in a developing world academic context. *South African Journal of Bioethics and Law, 6(1)*, 21–24.

Iorns, E., & Chong, C. (2014). New forms of checks and balances are needed to improve research integrity. *F1000Research, 3(119)*. https://doi.org/10.12688/f1000research.3714.1.

Joynson, C., & Leyser, O. (2015). The culture of scientific research. *F1000Research, 4(66)*. https://doi. org/10.12688/f1000research.6163.1.

Kalichman, M. (2014). Rescuing responsible conduct of research (RCR) education. *Accountability in Research, 21(1)*, 68–83. https://doi.org/10.1080/08989621.2013.822271.

Keeney, S., Hasson, F., & McKenna, H. P. (2001). A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies, 38(2)*, 195–200. https://doi.org/ 10.1016/S0020-7489(00)00044-4.

Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: Ten lessons from using the Del- phi technique in nursing research. *Journal of Advanced Nursing, 53(2)*, 205–212. https://doi.org/10. 1111/j.1365-2648.2006.03716.x.

Kumar, M. N. (2010). A theoretical comparison of the models of prevention of research misconduct. *Accountability in Research, 17(2)*, 51–66. https://doi.org/10.1080/08989621003641132.

Lerouge, I., & Hol, T. (2020). *Towards a research integrity culture at universities: From recommendations to implementation.* League of European research universities.

Mejlgaard, N., Bouter, L. M., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A.-K., et al. (2020). Research integrity: Nine ways to move from talk to walk. *Nature, 586(7829)*, 358–360. https://doi. org/10.1038/d41586-020-02847-8.

Mokkink, L. B., Boers, M., van der Vleuten, C., Bouter, L., Alonso, J., & Patrick, D. L., et al. (2020). COSMIN risk of bias tool to assess the quality of studies on reliability or measurement error of out- come measurement instruments: A Delphi study.

National Academies of Sciences Engineering and Medicine. (2017). *Fostering integrity in research.* The National Academies Press.

Pare, G., Cameron, A.-F., Poba-Nzaou, P., & Templier, M. (2013). A systematic assessment of rigor in information systems ranking-type Delphi studies. Information & Management, 50(5), 207–217. https://doi.org/10.1016/j.im.2013.03.003.

Powell, C. (2003). The Delphi technique: Myths and realities. Journal of Advanced Nursing, 41(4), 376– 382. https://doi.org/10.1046/j.1365-2648.2003.02537.x.

Research integrity is much more than misconduct (2019). Nature, 570(5). https://doi.org/10.1038/ d41586-019-01727-0.

Richman, V., & Richman, A. (2012). A tale of two perspectives: Regulation versus self-regulation. A financial reporting approach (from Sarbanes–Oxley) for research ethics. Science and Engineering Ethics, 18, 241–246. https://doi.org/10.1007/s11948-011-9260-8.

Rifai, N., Annesley, T. M., Moore, S., Caplan, A. L., Sweet, D. J., Hornung, P., et al. (2019). Maintaining research and publication integrity. Clinical Chemistry, 65(2), 230–235.

Satalkar, P., & Shaw, D. (2019). How do researchers acquire and develop notions of research integrity? A qualitative study among biomedical researchers in Switzerland. BMC Medical Ethics, 20(1), 1–12.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., et al. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. Quality & Quantity, 52, 1893–1907. https://doi.org/10.1007/s11135-017-0574-8.

Ščepanović, R., Tomić, V., Buljan, I., & Marušić, A. (2019). D3. 3: Report on the results of explorative interviews. https://sops4ri.eu/wp-content/uploads/D3.3_Report-on-the-results-of-the-explorative- interviews-1.pdf.

Sørensen, M. P., Ravn, T., Bendtsen, A. K., Reyes-Elizondo, A., Kaltenbrunner, W., Ščepanović, R., et al. (2020). https://sops4ri.eu/wp-content/uploads/D5.2_Report-on-the-Results-of-the-Focus-Group- Interviews-1.pdf.

Stemwedel, J. D. (2014). Life after misconduct: Promoting Rehabilitation while minimizing damage. Journal of Microbiology & Biology Education, 15(2), 177–180. https://doi.org/10.1128/jmbe.v15i2. 827.

Steneck, N. H. (2006). Fostering integrity in research: Definitions current knowledge and future directions. Science and Engineering Ethics, 12(1), 53–74. https://doi.org/10.1007/PL00022268.

Tereskerz, P., & Mills, A. (2012). COI policies: Tax dollars should not be used to fund US institutions not making the grade. Accountability in Research, 19(4), 243–246. https://doi.org/10.1080/08989621. 2012.700884.

Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M., Patrick, D. L., Alonso, J., et al. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. Quality of Life Research, 27, 1159–1170. https://doi.org/10.1007/s11136-018-1829-0. Titus, S., & Bosch, X. (2010). Tie funding to research integrity. Nature, 466, 436–437. https://doi.org/10.1038/466436a

von der Gracht, H. (2012). Consensus measurement in Delphi studies: Review and implications for future quality assurance. Technological Forecasting and Social Change, 79(8), 1525–1536. https://doi.org/ 10.1016/j.techfore.2012.04.013.

VSNU (2018). The Netherlands code of conduct on research integrity.

Youngblut, J. M., & Brooten, D. (2002). Institutional research responsibilities and needed infrastructure. Journal of Nursing Scholarship, 34(2), 159–164. https://doi.org/10.1111/j.1547-5069.2002.00159.x. Zwart, H., & ter Meulen, R. (2019). Addressing research integrity challenges: From penalising individual perpetrators to fostering research ecosystem quality care. Life Sciences, Society & Policy, 15(5). https://doi.org/10.1186/s40504-019-0093-6.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Krishma Labib[1], Rea Roje[2], Lex Bouter[3,4], Guy Widdershoven[1], Natalie Evans[1], Ana Marušić[2], Lidwine Mokkink[3], Joeri Tijdink[1,4]

1   Department of Ethics, Law and Humanities, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Public Health Institute, De Boelelaan 1117, 1081HV Amsterdam, The Netherlands
2   School of Medicine, University of Split, Šoltanskaul. 2, 21000 Split, Croatia
3   Department of Epidemiology and Data Science, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Public Health Institute, De Boelelaan 1117, 1081HV Amsterdam, The Netherlands
4   Department of Philosophy, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Section 2: Developing guidelines

Chapter 4

4. Education and training policies for research integrity – Insights from a focus group study

Abstract

Education is important for fostering research integrity (RI). Although RI training is increasingly provided, there is little knowledge on how research stakeholders view institutional RI education and training policies. Following a constructivist approach, we present insights about research stake- holders' views and experiences regarding how research institutions can develop and implement RI education and training policies. We conducted thirty focus groups, engaging 147 participants in eight European countries. Using a mixed deductive-inductive thematic analysis, we identified five themes: (1) RI education should be available to all; (2) education and training approaches and goals should be tailored; (3) motivating trainees is essential; (4) both formal and informal educational formats are necessary; and (5) institutions should take into account various individual, institutional, and system-of-science factors when implementing RI education. Our findings suggest that institutions should make RI education attractive for all and tailor training to disciplinary-specific contexts.

Key words: research integrity; responsible conduct of research; education; training; research institutions; institutional policies; research stakeholders

1. Introduction

There is a growing awareness of the importance of research integrity (RI) for producing high-quality and relevant research (Nature, 2019). RI can be defined as doing research according to high ethical, professional, and methodological standards (Boehme et al. 2016). RI education is considered to be crucial in fostering RI and in reducing misbehaviors—both serious (e.g. fabrication, falsification, and plagiarism) and minor (e.g. questionable research practices such as hypothesizing after the results are known; Hermeren et al. 2019). It is widely acknowledged that research institutions are responsible for providing good-quality RI education and training (Resnik and Shamoo 2011; Forsberg et al. 2018; Fanelli 2019). Although there are various ways to differentiate between 'education' and 'training' (Masadeh 2012), in this paper, we use the term 'RI education' to refer broadly to all approaches — both formal and informal — used to develop knowledge, skills, moral values, and understanding of RI, whereas we refer to 'RI training' when discussing formal instructional events or programs used for RI education (e.g. courses and workshops). RI training is, thus, a crucial part of RI education.

In the past decades, many stand-alone RI courses have been developed globally, mostly aimed at PhD students, using a diversity of training designs, approaches, and organization methods (Kalichman 2013; Boehme et al. 2016; Fanelli 2019; Abdi et al. 2021). However, it is increasingly recognized that RI education entails more than isolated training events. It requires addressing various target groups— rather than merely focusing on junior researchers—as well as a continuous effort by

both institutions and individual researchers (Kalichman 2007; Evans et al. 2018; Fanelli 2019). Therefore, a more systematic institutional approach to RI education is needed to ensure the relevance for stakeholders and to develop the evidence base related to different approaches and their effectiveness. In particular, it would be helpful for institutions to develop a comprehensive RI education and training policy outlining the overarching institutional RI education strategy, as well as the concrete plans and procedures needed to implement it. Indeed, RI education has been recently highlighted as one of nine important topics to address in institutional RI policies and—in Europe— is now actively promoted by the European Commission (Mejlgaard et al. 2020; SOPs4RI 2020).

While there are a variety of RI educational tools currently available (Pizzolato et al. 2020), much of the existing literature on RI education is quantitative and focuses on studying the evaluation of single trainings (Anderson et al. 2007; Powell et al. 2007; Godecharle et al. 2014; Marusic et al. 2016; Watts et al. 2017; Antes et al. 2018, 2009; Fanelli 2019; Mejlgaard et al. 2020). Based on the little evidence available, much of which is considered to be of poor quality (Marusic et al. 2016), it is not clear whether, and if so which, RI training approaches are valuable in changing stakeholders' behaviors related to RI (Powell et al. 2007; Antes et al. 2009; Marusic et al. 2016; Science Europe 2017; Fanelli 2019). A possible explanation for the lack of good-quality available evidence is that there is no consensus about which learning aims RI education should approach, and hence, which outcomes evaluations should measure (Fanelli 2019). Several educational approaches have been discussed and problematized as potentially suitable such as reducing misconduct, improving RI knowledge and skills, internalizing RI values, and navigating the gray areas of research (DuBois 2004; Powell et al. 2007; Geller et al. 2010; Kalichman 2014; Fanelli 2019). It is thought that RI education likely needs to incorporate both formal training and informal educational approaches (e.g. through responsible mentorship; Kalichman 2007; OECD Global Science Forum 2007; Alfredo and Hart 2011; Satalkar and Shaw 2018), although there is little insight available about how education should be organized to optimize the benefits of different educational approaches. For instance, it remains unclear whether mandatory RI educational events are desirable in ensuring participation in RI education for various target groups or whether other—and if so, which—incentives are needed (Fanelli 2019).

Although education is seen as crucial in fostering RI, the literature suggests that training by itself will likely be insufficient in influencing researchers' behaviors (Aubert Bonn and Pinxten 2021). This is because researchers' behaviors are not only influenced by their awareness and attitudes regarding RI, but also the research system in which they operate, as well as their local research culture (Titus and Bosch 2010; Joynson and Leyser 2015; Forsberg et al. 2018; Rifai et al. 2019). In fact, some have even criticized the emphasis on education and training in the field of RI, arguing that a stronger emphasis is needed on the underlying factors and different stakeholders that influence researchers' behaviors rather than focusing on individual researchers (e.g. Aubert Bonn and Pinxten 2021). There is little discussion about how institutions can successfully develop and implement institutional policies on RI education, which adequately take into account these considerations. Furthermore, there are few qualitative studies on how researchers and other stakeholders perceive and engage

with RI education, even though such insights could provide needed guidance for institutions on how to develop engaging RI education and training, and incentivize participation— particularly considering the weak evidence base related to RI education. Such a constructivist approach is important for generating relevant insights with stakeholders for developing and implementing meaningful institutional policies on RI.

A set of American focus groups with medical postdoctoral researchers from 2008 suggests that researchers prefer the use of case study discussions in training to address their day-to-day RI dilemmas (Alexander and Williams 2008). Two recent European focus group projects showed support among diverse research stakeholders for continuous RI education across academic ranks and for everyone involved in the research process (Evans et al. 2018; Pizzolato and Dierickx 2021). Another European stakeholder consultation with RI experts resulted in the recommendation that RI education should consist of a core curriculum on RI basics (e.g. on what are questionable research practices) and a follow-up advanced curriculum to further deepen RI insights (e.g. how to prevent questionable research practices in their daily research; ENERI 2017).

While these reports provide important insights about various research stakeholders' perceptions of RI education and training, it is not clear to what extent the findings apply across different countries and disciplinary fields. Further- more, the perspective of researchers from fields other than RI is particularly unexplored. Yet, following a constructivist approach, understanding stakeholders' diverse research perspectives, views, needs, and preferences about realities on the ground is crucial for developing and implementing meaningful institutional policies on RI. Therefore, in this article, we report on the results of a European focus group study with research stakeholders from various disciplines, to delve deeper into the research question: What are researchers' and other stakeholders' views and preferences regarding how research institutions can develop and implement better education and training policies?

2. Methods

This article is based on data from a focus group study that has been preregistered on the Open Science Framework (OSF): https://osf.io/kdtnx/ (Sørensen et al. 2021). The full data analysis protocol can be found here: https://osf.io/jntck/. The focus group study is part of the Standard Operating Procedures for Research Integrity (SOPs4RI) project (www.sops4ri.eu), which aims to develop guidance for research institutions and funders on how to foster RI. Since the insights from the focus groups were intended to inform the guidance being produced in the SOPs4RI project, we intended to obtain practical information, based on stakeholders' views and experiences, that would be valuable for institutions in developing policies on RI education. The research was approved by the Research Ethics Committee of Aarhus University (https://osf.io/kdtnx/). Information about the goals and procedure of the study were sent to participants at least a week before the focus groups took place, and informed consent was obtained before or at the start of each focus group.

2.1   Study design

We conducted thirty focus groups in various parts of Europe with research stakeholders from the medical, natural, and social sciences, and the humanities (Table 1). The focus groups aimed to provide interpretations and viewpoints of stake- holders, including perceived effects of training, based on experience, as input for guidance on RI education and training for research institutions. Each focus group started with an open discussion about RI policies, followed by a more in- depth discussion of two or three RI topics, and ended with a sorting exercise. The interviews lasted 1.5–2 hours. The topic RI education and training was addressed in depth in eight focus groups, in which we asked participants to share their views on what measures institutions can take to develop RI education and training policies. For a full list of topics dis- cussed in all the focus groups, as well as the topic guide, please see https://osf.io/kdtnx/. All focus groups were con- ducted in English by one, two, or three facilitators (please see Table A.1 for more details) and were audio-recorded and transcribed. Participants received a comprehensive report of the focus group results (Sørensen et al. 2020).

2.2   Participant selection and recruitment

We used a purposive sampling strategy to identify and recruit participants from all main research areas to take part in the focus groups. We included researchers from specific disciplines and of various ranks (i.e. PhD students, post-doctorate researchers, and assistant, associate, and full professors). This was to allow for comparisons of views across disciplines and academic ranks, as well as to ensure that the view of stakeholders from various groups across Europe would be represented in the guidance developed by the SOPs4RI consortium. When recruiting junior researchers (e.g. PhD students), we only recruited those who had at least a few years of experience with research to ensure familiarity with the focus group topic and to minimize potential power dynamics. For half the focus groups, we also attempted to include at least two other research stakeholders with disciplinary- specific knowledge (e.g. confidential RI counselors, funders, editors, research ethics (RE) committee members, and RI trainers). We used two strategies simultaneously to identify and recruit participants: (1) we approached contacts from our networks via email, followed by snowballing, and we (2) looked on research institutions' websites to identify and invite suitable candidates (i.e. researchers of various ranks or other stakeholders with disciplinary-specific knowledge). The composition of the focus groups can be found in Table 1. Information on dropouts from the focus groups can be found in Table A.2.

Table 1. Composition of focus groups.

Participants → Focus groups↓		Total number of participants (% female)	Researchers included		Other stakeholders included [b]	
			# Junior	# Senior[a]	# Included	Type of stakeholder
Discipline Type of focus group	Country					
Humanities	*Total*	34 (53%)	8	23	8	
Researchers only	Netherlands[c]	7 (57%)	3	4	–	–
	Denmark	3 (33%)	1	2	–	–
	Croatia	6 (33%)	2	4	–	–
Mixed stakeholders	Netherlands[c]	5 (60%)	1	4	2	Confidential counselor; RI committee member
	Spain	4 (75%)	0	3	2	RI & RE committee member
	Germany	5 (80%)	1	3	2	Funder; Editorial director;
	Greece	4 (25%)	0	3	2	Diversity expert; Funder; RI committee member
Social sciences	*Total*	32 (63%)	9	20	8	
Researchers only	Germany[c]	5 (60%)	4	1	1	PhD association board member
	Spain	4 (75%)	2	2	–	–
	Netherlands	6 (33%)	2	4	–	–
Mixed stakeholders	Netherlands[c]	6 (67%)	1	5	3	RI teacher; Research director; RE committee member;
	Denmark	4 (50%)	0	2	2	Journal editor RI officer; RE coordinator
	Croatia	5 (80%)	0	4	2	Officer for science; Vice-Dean; Former journaleditor
	Greece	2 (100%)	0	2	0	–
Natural sciences	*Total*	42 (40%)	4	28	13	
Researchers only	Croatia[c]	6 (83%)	2	4	–	–
	Spain	2 (50%)	0	2	–	–
	Denmark	6 (33%)	2	4	–	–
	Belgium	3 (0%)	0	3	–	–
Mixed stakeholders	Netherlands[c]	8 (50%)	0	2	6	Medical research coordinator
Denmark	Denmark	5 (20%)	0	3	2	RI officer; Researcher union
Croatia	Croatia	6 (50%)	0	5	3	R&D employee; Policy maker;

Participants → Focus groups↓		Total number of participants (% female)	Researchers included		Other stakeholders included [b]	
			# Junior	# Senior[a]	# Included	Type of stakeholder
Discipline Type of focus group	Country					
	Greece	6 (17%)	0	5	2	..Vice-dean; Funders
Medical sciences	*Total*	39 (59%)	4	39	12	
Researchers only	Croatia[c]	7 (57%)	1	6	–	–
	Denmark	3 (100%)	0	3	–	–
	Greece	3 (33%)	0	3	–	–
	Netherlands	6 (67%)	1	5	1	Department chair
Mixed stakeholders	Belgium[c]	5 (40%)	0	3	3	Funder; Valorization officer; Research director
	Denmark	4 (75%)	0	2	2	Funder; Research support officer
	Spain	4 (50%)	1	1	2	RI officer; Scientific coordinator
	Italy	7 (57%)	1	6	4	RE & RI review board members
Total		147 (53%)	25	110	41	

Table 1 continued

a. We considered a researcher as being 'senior' if they held a tenure position (which we assumed when participants had a position as an 'assistant, associate or full professor'), or—in case we did not have this information—if participants had 5 or more years of research experience post obtaining a doctorate. We considered a researcher as being 'junior' if they did not yet hold a tenure position (which we assumed when they had a position as a 'PhD student', 'junior researcher', or 'postdoctoral researcher'), or—in case we did not have this information, if they had less than 5 years of research experience post obtaining a doctorate.
b. Please note that some participants represented more than one role (i.e. represented both a researcher and other type of stakeholder, and/or represented multiple other types of stakeholders). These participants are only counted once under each of the columns 'Total # of participants', 'Researchers included' (if the stakeholder is a researcher), and 'Other stakeholders included'.
c. Indicates focus groups in which the topic of RI education and training was explicitly discussed in depth.

2.3   Data analysis

We used a hybrid deductive–inductive thematic analysis approach to analyze the data as in Fereday and Muir-Cochrane (2006) and Swain (2018), in an iterative manner. Data analysis was conducted using the program MAXQDA 2018. In consultation with N.E. and J.T., K.L. developed a preliminary deductive code book based on a categorization of concepts related to RI or RE education contained in two scoping reviews about RI policies (Gaskell et al. 2019; Ščepanovic et al. 2021). With this step, we intended to find views that can serve as a basis for guidance on RI education. The code book was subsequently added to and adjusted over five iterations collaboratively by K.L., N.E., R.R., P.K., and J.T. The development of the code book can be found here: https://osf.io/dyta8/.

The full transcripts of the eight focus groups with in-depth discussions on RI education and training were analyzed; these were supplemented by an analysis of parts of the remaining twenty-two focus group transcripts that contained data relevant for RI education and training (consisting of forty-nine pages of transcripts in total). A first round of coding was done per unit of meaning (i.e. sentence, phrase, or paragraph that refers to one code) by labeling text representing the codes from the preliminary code book. Emerging inductive codes were added to the code book and used for analysis, either as separate from the deductive codes or expanding on the deductive codes. To increase reliability, all the data were also analyzed by an independent coder in a second round of coding using the updated code book. Any discrepancies between the first coder (K.L.) and second coder (N.E., P.K., and R.R.) were discussed, in order to make a final decision and make any necessary alterations to the code book. A third coder (J.T.) was consulted in case of disagreements. No new codes emerged when coding the last focus group transcripts, indicating that data saturation was reached.

After coding was complete, we clustered the codes into overarching themes and subthemes (Fereday and Muir-Cochrane 2006). The data were initially analyzed using an across-case analysis strategy, where we looked at commonalities in the data across cases (i.e. disciplines, stakeholder types, and ranks; Ayres et al. 2003). Next, using a within- case analysis strategy, we re-examined the earlier stages of the data analysis and scrutinized the alignment between the overarching themes/subthemes with quotes within disciplines, stakeholder types (i.e. researchers versus other stakeholders), and academic ranks (i.e. junior versus senior researchers), to ensure that the interpretation of the data applies well across all cases (Ayres et al. 2003; Fereday and Muir-Cochrane 2006).

3. Results

Based on our analysis, we identified five overarching themes arising from the data: (1) education should be available to all involved in research; (2) the education and training approach, goals, and content should be tailored; (3) institutions should actively motivate trainees to engage in RI education and training; (4) both formal and informal educational formats should be included in the RI education and training policy; and (5) institutions should take into account various individual, institutional, and system-of-science factors when implementing RI education. Figure 1 provides a summary of the themes and subthemes. We elaborate further on each theme in the sections below. Our analysis showed few differences in perspectives among participants from different disciplines and ranks in the study; any differences are highlighted in the explanation of results for each theme. More information on the quotes found in the text, as well as additional illustrative quotes per theme, are presented in Tables A.3–A.7.

3.1 Training availability

Participants across disciplines stressed that RI education should be available for all researchers (Table A.3). Many mentioned that their institutions have RI training available at least for PhD students. However, some of these trainings seemed to be recently established, with multiple junior researchers remarking that they had never

received RI training. This might especially be relevant for junior postdoctoral researchers, as highlighted by one postdoc's concern that 'there are so many gaps, relating specifically to the post- doc experience' (junior researcher, humanities, Netherlands). Some participants remarked that starting RI education at the PhD level is too late and that bachelor and master students also need to be targeted.

Many participants across disciplines emphasized that in order to ensure adequate support and supervision to junior researchers, senior researchers should also be trained. Additionally, they highlighted that senior researchers also need to be trained for their own learning process, i.e. not only for the sake of supporting others. Only in one humanities focus group did participants mention that professors in their institution receive RI training. Additionally, a few researchers argued that RI training should also target other research stakeholders in the institution, including staff, managers, rectors, and deans, since they have an important responsibility regarding RI, with one participant even exclaiming that 'they are ruling' research (senior researcher, natural sciences, Croatia).

3.2   Education and training approaches, goals, and content

While one researcher wished that RI trainings would cover all aspects of RI, others argued that it is not feasible to teach everything and stressed that training needs to be tailored to the disciplinary field and target group. A range of—sometimes conflicting—goals, approaches, and content were described as appropriate for RI education and training (Table A.4). First, there was at least some mention of addressing research misbehaviors (such as plagiarism, fraud, and data manipulation) and even sexual misconduct during RI training in most focus groups. One researcher explained that research misbehaviors are often unintentional and 'many students commit plagiarism without realizing it' (senior researcher, medical sciences, Greece), so it is important to raise awareness about them by showing trainees the harmful consequences of serious and minor research misbehaviors. However, a few participants were concerned that merely telling trainees what not to do may not have sufficient impact, with one participant stating that telling researchers 'Okay, this is bad okay' will not change 'anything that drives people towards that behavior' (senior researcher, social sciences, Netherlands).
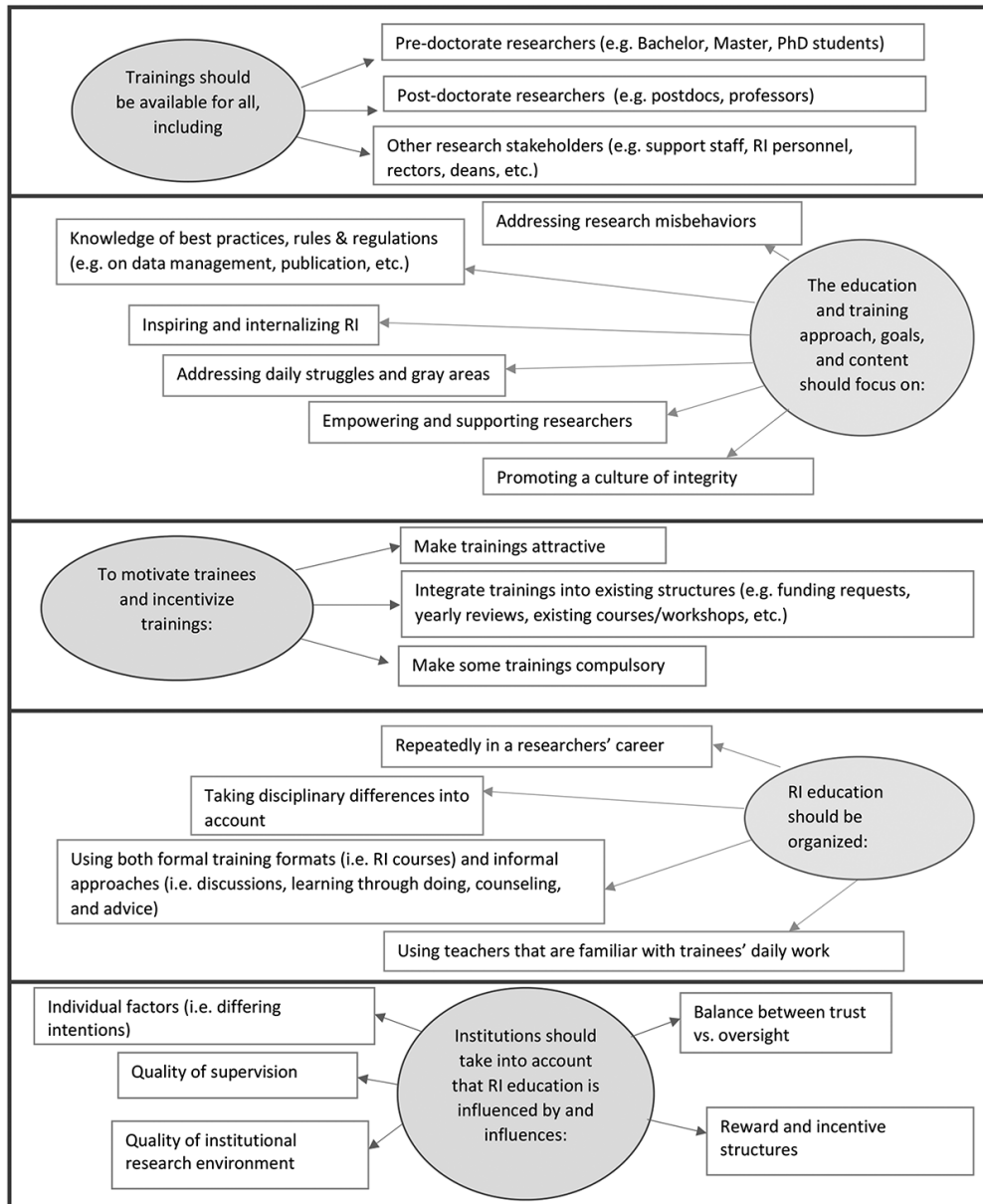
Figure 1. *Main findings on participants' perceptions of RI education and training policies.*

Secondly, across disciplines, there was much discussion about making participants aware of best practices, rules, and regulations relating to RI, since it is important to place—in the words of one participant—'the full responsibility […] on the institution […] to make sure that all its research performing employees are actually aware of current rules or changes in rules' (funder, medical sciences, Denmark). Topics that fell under these discussions varied between disciplines and included 'ethics' (humanities);

open science (humanities and social sciences); conflicts of interest (humanities and medical sciences); authorship, data management, and analysis (social, medical, and natural sciences); peer review (social and medical sciences); citations (social sciences); good clinical practice, research with animals, and confidentiality, (medical sciences); and lab work (natural sciences).

The third approach to RI education discussed across disciplines was about inspiring trainees and helping them to internalize RI by engaging them in reflections on why they should engage in good research practices, as stipulated in RI guidelines. One participant was concerned that discussions about RI came about because researchers' 'professional[ism] is under threat and the public imagery is turning on its head' rather than because 'from the inside [researchers] believe in these things' (senior researcher/RI committee member, humanities, Netherlands). Many acknowledged that trainings should help trainees to reflect on their intrinsic motivations for RI. Ideas on how to do this included discussing virtues, testimonials, experiences, and quotes from inspirational figures and examples.

Fourthly, there was widespread agreement that practical considerations in research can sometimes contradict ethical and legal ones, leading to the presence of so-called 'gray areas' where the right course of action is not crystal clear. For instance, participants in several focus groups mentioned that it can be challenging to abide by best practices regarding safe and secure data management in cases where avail- able secure data management infrastructure is not of high quality or convenient to use. It was stressed that RI education should address the real daily struggles and gray areas that researchers encounter, rather than only—as one participant remarked—'having courses that are idealizing the situation' (senior researcher, social sciences, Netherlands). Many also advocated for training that aims to empower and support researchers to deal with power dynamics, as an additional approach to RI education. One participant explained that PhD students are left with doing many tasks 'on their own time, really,[…] become tennis balls' (senior researcher, humanities, Netherlands) and need support to speak up in research, while another mentioned that 'just knowing what rights you could have if it was an ideal world' would be helpful (junior researcher, social sciences, Germany). Another participant even remarked that 'we can also train our students [… to] stand up for their thing', since 'without that you cannot become a good researcher' (senior researcher, social sciences, Netherlands).

Building an RI culture was considered the ultimate, overarching goal of RI education according to multiple participants from various disciplines. For instance, one participant mentioned that RI education is 'helpful […to] change the culture, the mentality' (senior researcher, natural sciences, Croatia), while another explained that the ultimate goal of not only RI education but institutional RI policies in general is 'that there's a culture of research integrity and there are all kinds of instruments that you can think about to promote this culture', including education (senior researcher/RI committee member, humanities, Netherlands). Participants advocated for creating a research culture consisting of various features such as the presence of common standards and expectations; an error- accepting culture that makes it possible for researchers to be open about mistakes and doubt; a culture of collaboration and data sharing; and a space for open, joint reflections. A few participants were hesitant about the feasibility of building an RI culture through training, with one remarking somewhat

cynically 'I am super curious how you are going to change an existing culture with a training session' (senior researcher, natural sciences, Netherlands).

3.3   Motivations and incentives to participate in RI training

Resistance to RI education was one of the main challenges identified by most study participants, particularly when it comes to senior researchers (Table A.5). For instance, one participant was concerned that 'the thing with [RI] courses is that the people who need it won't do it' (senior researcher, humanities, Netherlands). Another mentioned that 'research integrity is not a favorite' even among PhD students (senior researcher/research director, medical sciences, Belgium). The main reasons discussed for why researchers might not be motivated to participate in RI training were the presence of competing priorities and trainings; the fact that many researchers believe that RI education is not helpful for them since they are already well-intentioned; and the perception that trainings are often not tailored enough to the specific disciplinary needs of trainees. Solutions on how to address resistance revolved around making trainings more attractive using various strategies such as 'dropping the [absolute] term' RI as it is 'religious' (senior researcher, social sciences, Netherlands) and replacing it with a more neutral one such as 'scientific conduct' (senior researcher/RI teacher, social sciences, Netherlands); tailoring training programs to the needs of the trainees; updating train- ing to address new issues in research; focusing on specific RI topics such as data management, rather than general RI consideration; showing that trainings are effective; and making trainees curious by discussing case studies 'maybe a little bit like gossip' (privacy officer, natural sciences, Netherlands).

　　　Participants across disciplines also advocated for integrating RI education into bachelor and master courses related to research, and courses targeted at researchers on other topics, such as academic writing, methodology, and personal development and leadership. Furthermore, there were discussions on tying RI education and training to funding and assessments in order to—as put by one participant—'use that competitiveness' of research to push for RI education (RE committee secretary, natural sciences, Netherlands). In general, there was agreement among the medical and natural science participants that making RI training mandatory could be beneficial, with some recommending it not only for PhD students but also for senior researchers and policymakers. However, a few participants raised concerns that mandating RI training 'will surely create a lot of resistance' (RI officer, natural sciences, Denmark) and could even become an 'obstacle' to training (senior researcher, social sciences, Denmark). Within the humanities and social sciences, it was suggested that the decision on whether to make training mandatory depends on the type of training provided, with one participant proposing that mandatory courses are suitable when addressing 'a certain code of conduct', but courses on issues such as referencing do not 'necessarily' have 'to be mandatory' (senior researcher, humanities, Netherlands).

3.4   Organization of RI education and design of trainings

Participants discussed a range of issues concerning the organization of RI education and design of trainings (Table A.6). First, there was agreement across disciplines that RI

education should be a continuous, even 'lifelong' (senior researcher, medical sciences, Denmark), learning process. Therefore, it was recommended to repeat training for all academic ranks every few years as illustrated by the following remark from a participant: '[Continuous training is] important to make sure that you calibrate your compass each time to make sure you can make decisions [about RI]' (senior researcher/RI teacher, social sciences, Netherlands). Secondly, there were various ideas about the format of RI education with references being made to both explicit training (i.e. 'formal' courses) and making use of implicit opportunities to educate about RI (i.e. via 'informal' educational approaches). While there was consensus among participants in the medical and natural sciences about the usefulness of formal RI courses, there was dis- agreement about this in the humanities and social sciences. For instance, one humanities researcher explained that for- mal 'training doesn't make sense' unless there is a need to address a 'legalistic framework' (senior researcher, humanities, Netherlands), while a social scientist remarked that 'not every aspect of research integrity can be trained explicitly' in a course (senior researcher, social sciences, Germany). Different types of methods were suggested as appropriate for formal courses, including interactive workshops, games, small tutorials, online training, case study discussions, peer interviews about the research process, and open discussions.

Participants in the social and natural sciences, and the humanities, also advocated for using informal discussions outside of formal courses to educate about RI, with one explaining that much can be learned about RI during conversations 'over coffee time' (senior researcher, RI committee member, humanities, Netherlands). There were also ideas about specific situations that can facilitate such informal discussion opportunities, including institute and team meetings, hackathons, mistake cakes (i.e. events where researchers openly discuss mistakes they have done in research over cake), and journal clubs. Similarly, it was highlighted that much of RI education happens implicitly while doing research, applying for ethics review, receiving supervision, and being socialized in a particular research environment; one participant referred to this as 'learning by doing' (senior researcher, social sciences, Germany). It was thought that RI education policies should exploit this implicit type of informal learning. Participants mentioned that as particular questions about RI come up in practice, researchers need access and referral to resources and guidance documents on good research practice to consult in order to address their questions and further their RI education. Besides this, the need for researchers to have access to people that can provide face-to-face advice on doubts that come up in research was also discussed. Both specialized advisors such as librarians, privacy officers, and ethics committee members, and general RI advisors were referred to. However, there was some concern raised that RI advisors are often perceived as too formal and related to procedures of reporting misconduct to be deemed approach- able by researchers for day-to-day questions on RI, with one participant asking 'So why would you take that step [of approaching one]?' (senior researcher/RI teacher, social sciences, Netherlands).

There was a diversity of opinions offered on the issue of who is suitable to provide RI trainings. One humanities participant suggested asking existing experts to allocate time for providing RI training rather than hiring new people, while another mentioned that teachers should have been themselves taught by 'some sort of authority' on how to teach RI (junior researcher, humanities, Croatia). Others

emphasized the need to hire RI trainers with discipline-specific knowledge, with one participant mentioning that that is necessary to ensure that the trainer 'really is aware of what they're talking about' (senior researcher, humanities, Netherlands). While it was thought that trainers should have experience with research, there was disagreement among participants about whether researchers would be motivated to provide RI trainings. How- ever, most agreed that trainings should be provided in a discipline-specific context (e.g. by doctoral schools or faculties). It was mentioned that informal educational events (e.g. RI discussions during journal clubs) could address com- mon RI issues faced by multiple disciplines. Furthermore, there were different ideas on whether trainings should be provided in full courses of smaller workshops or integrated in other courses, with a few suggestions across disciplines that decisions on this should be based on the target group of the training, and the institutional resources and capacities available.

3.5   Factors influencing the implementation of RI education and training policies

The focus group participants identified five factors that might influence the implementation of institutional RI education and training policies: (1) individual factors of trainees, (2) super- vision, (3) institutional research environment, (4) trust versus oversight, and (5) reward and incentive structures (Table A.7). Various facilitators and barriers were mentioned for each factor, and these are further elaborated on in this section. Please also take into consideration that these factors are interrelated and likely influence one another. To begin with, one participant expressed that 'the individual is also important and that's sometimes very difficult to change' via training (senior researcher/research director, medical sciences, Belgium). This view was mirrored by others who warned about a small minority of researchers who are ill-intentioned and will engage in misconduct to cut corners; trainings will likely not be able to change these attitudes. Others highlighted that cultural differences among individual researchers can lead to a difference in understanding of research norms and values.

Regarding mentorship, participants across disciplines high- lighted that unless trainees' supervisors support them in engaging in responsible research practice, RI training might be futile. For instance, one remarked that 'we train the PhD students in these courses for half a day or one day during four years, when they get trained by that professor […] for four years' (senior researcher, natural sciences, Netherlands), while another explained that 'If you're supervised in a correct and responsible way you will behave automatically as you should' (junior researcher, social sciences, Germany). How- ever, it was acknowledged that the relation between RI training and supervision is mutual, and training targeted at supervisors could improve supervision.

Similarly, across disciplines, it was mentioned that the institutional research environment and RI education have a mutual influence on each other. On the one hand, when aimed at creating a responsible research culture, RI education can support the institutional research environment. On the other hand, the research environment influences whether values and practices addressed in RI education and training are actually internalized by trainees. For instance, one participant explained that researchers are mainly taught by being socialized in their research environment since

'it's labs that train the students' (senior researcher, natural sciences, Spain). Characteristics of the research environment that could negatively influence RI education mentioned included the 'vanity' of the 'discourse of excellence' present in academia, i.e. the urge to present oneself as 'excellent' (senior researcher, humanities, Netherlands); too much competition; a rigid hierarchy; a culture of blaming; and a lack of institutional commitment to RI, e.g. in terms of data management support and bylaws.

Maintaining a balance between trust in researchers and oversight of RI in the institution was another factor that was emphasized as important across focus groups. A few preferred to have some oversight on RI in the form of rules and documents, with one researcher remarking that 'although it looks like a bureaucracy, the key is in the documents. If you don't have documents you have nothing' (senior researcher, medical sciences, Croatia). Alternatively, many others explained that there is currently too much emphasis on oversight and bureaucracy in research, which can be counterproductive. For instance, one participant mentioned that '[increased over-sight] creates more distrust towards[...]researchers instead of the thing you want to achieve[which] is to have responsible researchers' (privacy officer, natural sciences, Netherlands). Participants thought that RI education can be a means to reduce oversight; as institutions train researchers to be responsible, trust in researchers can increase, leading to reduced monitoring and oversight. However, a few humanities and social science researchers were concerned that—as put by one participant—'requiring courses is part of a general suspicion against people working here' (senior researcher, humanities, Spain), and also constitutes oversight and added bureaucracy. Finally, there was agreement across disciplines that a major hurdle, which institutions face when providing RI education and training, is the current structure of rewards and incentives in science. One participant argued that '[if] the structures are not good then you can train as long as you like, but you're never going to change what is going to happen really, in the grassroots' (senior researcher/confidential counselor, humanities, Netherlands). Another participant explained that providing RI education and training does not address the root cause of RI problems by saying that 'it's not a lack [of awareness] of guidelines, but the way[...]that researchers are rewarded or not rewarded by funding[...]that causes most of the problems' (senior researcher, medical sciences, Belgium). Particular elements of the current reward and incentive structures that were identified as a barrier for RI included the culture of deliverables, publication pressure, precarious working conditions, and project-based funding. Interestingly, two participants (a senior researcher and funder) remarked that the competitive nature of existing reward and incentive structures 'is not a barrier' for RI training (natural sciences, Netherlands), since training can be embedded into these structures (e.g. if seen as a competitive advantage in grant applications). Furthermore, there was agreement that junior researchers are particularly vulnerable to these structures due to their dependence on seniors for funding.

4. Discussion

In this study, we assume that in order to foster RI through education, institutions need to develop and implement an overarching strategy for RI education and training, which

takes into account research stakeholders' preferences and perspectives. Our results highlight that RI education is unlikely to be successful if implemented without sufficient attention to other institutional RI responsibilities (e.g. ensuring responsible supervision and a culture conducive to RI). Furthermore, our results show wide stakeholder support for RI education targeting all research stakeholders (e.g. researchers and deans; Evans et al. 2018; Fanelli 2019). The RI education and training goals, approaches, and content discussed by the participants in our study reflect the multitude of existing approaches used in training programs (Fanelli 2019; Abdi et al. 2021), suggesting that it might not be necessary or desirable to agree on the same learning goal for all educational programs. Instead, a tailored approach to RI education might be needed to sufficiently take into account disciplinary and country differences. Additionally, our results emphasize the importance of institutional efforts in motivating participants to actively take part in RI education. This is particularly important, considering the results indicating that RI education requires a continuous learning process consisting of formal and informal formats (e.g. both formal RI courses and informal discussions about RI).

A critical finding of our study is that the implementation of successful RI education is highly dependent on various individual, institutional, and system-of-science factors. These results confirm previous studies that suggest that cultural differences among individuals (Antes et al. 2018), supervision (Alfredo and Hart 2011; McGee et al. 2014), the institutional research environment (Kalichman 2007; Alfredo and Hart 2011; National Academies of Sciences, Engineering, and Medicine 2017; Satalkar and Shaw 2018), and reward and incentive structures in the system-of-science (Marie-Claude 2007; Asai et al. 2016) play a crucial role in the uptake and delivery of RI education. This highlights that addressing RI education policies in a vacuum—i.e. solely implementing RI education without addressing other institutional responsibilities for RI—will not be a sufficient strategy for institutions to foster RI, as they will need to develop a comprehensive institutional RI plan including policies on multiple RI topics, including supervision and the research environment, simultaneously (Mejlgaard et al. 2020). This recommendation is in line with literature suggesting that RI education should not be the mere means of fostering RI, as researchers' behavior is highly dependent on other RI factors such as promotions and evaluations (Aubert Bonn and Bouter 2021; Aubert Bonn and Pinxten 2021). Instead, institutions should create policies focusing on creating a collaborative, error-accepting, and open research environment, as well as on reducing unnecessary bureaucracies and potentially corrupting influences of hyper-competition in research. When integrated into the research endeavor—including the socialization into the research process—sufficiently, RI education could also be used as a tool to improving the research environment, e.g. by increasing awareness and reflection on RI in a research group (National Academies of Sciences, Engineering, and Medicine 2017; Labib et al. 2021).

As the results show, a variety of goals can be used for RI education and training; this highlights the difficulty of standardizing RI education and training approaches across institutions and disciplines. The majority of the goals and approaches discussed in our focus groups, including focusing on misconduct, improving knowledge of RI, internalizing RI values, addressing gray areas of research, and culture building, have also been mentioned in previous literature (DuBois 2004; Powell et al. 2007; Geller et al. 2010; Zeng and Resnik 2010; Kalichman 2014; Fanelli

2019; Valkenburg et al. 2021). An additional goal that our participants high- lighted is empowering researchers, i.e. equipping researchers to deal with power dynamics and the rigid hierarchies present in academia (e.g. by giving them relevant tools and knowledge necessary to speak up about RI to their supervisors). Although the literature suggests moving away from educational approaches that solely focus on cases of research misconduct and knowledge of RI concepts toward more aspirational and cultural approaches (DuBois 2004; Kalichman 2014), there was no clear preference for a specific approach across our focus groups. It was suggested that different approaches are suitable depending on the context and tar- get group of the training provided. For instance, the goal of empowering trainees could be considered more suitable for training targeted at junior, rather than senior researchers. Consequently, it might be most appropriate to tailor RI education and training approaches, goals, and content to the specific context and target group at hand (Watts et al. 2017). This view contrasts with pleas to agree on a fixed set of RI educational goals in order to allow for evaluations of training effectiveness (Kalichman 2007) and might explain why such agreement has not already been achieved (Fanelli 2019).

Our finding that RI training appears to be unappealing to researchers stresses the need for RI education and training policies to address training attractiveness and incentives. While multiple ways to incentivize the participation of junior researchers have previously been identified (e.g. digital badges and free meals; Fanelli 2019), our results show that it is senior researchers that need strong incentives to participate. Although it has been recognized that incentivizing senior researchers remains a challenge (Fanelli 2019), the participants in our study provided some concrete suggestions on how to address this, including ensuring that the training itself appears attractive (e.g. by addressing the real needs of the trainees), as well as integrating RI education in existing events, funding schemes, and assessments. Importantly, our results suggest that efforts to make RI education more appealing to researchers will also need to address researchers' resistance to the potential increased oversight that RI policies might introduce (Sørensen et al. 2020). What could help to pre- vent researchers from perceiving RI education as an additional bureaucratic hurdle is to regularly use implicit educational formats (e.g. open discussions at department meetings) to help researchers reflect on RI informally, rather than always relying on formal courses (Kalichman 2007; OECD Global Science Forum 2007; Alfredo and Hart 2011; Satalkar and Shaw 2018).

In line with existing recommendations (Forsberg et al. 2018; Ravn and Sørensen 2021), our results show that RI trainings should be provided in a disciplinary-specific context to ensure sensitivity to the real needs of trainees. Beyond this, we did not find many disciplinary differences in participants' views on important considerations for RI education and training policies. One potentially notable difference observed was that unlike participants in the medical and natural science focus groups, some humanities and social science researchers were skeptical about making RI training events compulsory. Furthermore, there was more hesitation among humanities and social science researchers than those in the medical and natural sciences, about the need for explicit and formal RI training events. These differences might be explained by the lower level of existing research regulation in the humanities and some social science fields, compared to research performed in the medical and

natural sciences (Wells et al. 2014; Haven et al. 2019). The extent to which researchers are already accustomed to having their research regulated could potentially influence their acceptance of new mandatory formal courses.

4.1 Strengths and limitations of the study

Although there are some existing insights about research stakeholders' perceptions of RI education (ENERI 2017; Evans et al. 2018), to our knowledge, this is the largest study providing an in-depth and explicit analysis of researchers' and other relevant stakeholders' preferences regarding the development and implementation of RI education in research institutions across Europe. The results of the eight focus groups explicitly addressing the topic of RI education and training in depth were further validated based on the data from the remaining twenty-two focus groups. The large amount of data collected allowed us to make comparisons across disciplines and ranks. The data also made it clear that there are few differences across disciplines and ranks. Since not all the focus groups explicitly addressed the topic RI education and training, we did not have sufficient data to be able to make cross-country comparisons, however.

Our analysis method allowed us to link insights gained in this study to existing knowledge about RI education. Since multiple topics, including RI education, were discussed throughout the thirty focus groups, we obtained a wide breadth of data on RI during the study, but that required some sacrifices regarding depth and richness of data on specific issues. Our qualitative approach enabled us to provide an additional angle to discussions about RI education, which are currently dominated by considerations of training effectiveness (Fanelli 2019). Our choice to only include junior researchers with a number of years of research experience led to the inclusion of a much smaller number of junior researchers as compared to senior researchers. This means that the perspective of junior researchers, who represent a significant proportion of the research workforce, might not be sufficiently explored. Considering that previous studies have shown differences in junior and senior perceptions of RI (e.g. juniors perceive their existing research climate more negatively; Haven et al. 2019, 2020), it would be valuable to validate our findings with other junior researchers.

5. Conclusions

In this study, we present insights about research stakeholders' views and experiences regarding RI education, which are relevant for developing and implementing RI education and training policies. We show that researchers and other research stakeholders across disciplines in Europe recommend RI education and training policies that (1) provide RI education to all involved in research; (2) use training approaches, goals, and content tailored to the target group; (3) focus on motivating trainees; (4) include formal and informal training formats; and (5) take into account various individual, institutional, and system-of-science factors when implementing RI education. These results confirm the need for research institutions to develop a comprehensive RI plan that integrates RI education into the research endeavor

(National Academies of Sciences, Engineering, and Medicine 2017; Mejlgaard et al. 2020), includes an overarching strategy on how to develop and implement RI education that is engaging for all involved in research, and tailors training to disciplinary-specific contexts.

Declarations

Funding

Conflict of interest statement.

None declared.

Acknowledgements

Authors' contributions

K.L. contributed to the design of the focus group study and was responsible for conducting focus groups, drafting and revising the data analysis protocol, analyzing the data, and drafting and revising the manuscript. N.E. contributed to revising the data analysis protocol, analyzing the data, drafting and revising the manuscript, and supervision. A.R.E., W.K., R.R., I.B., and P.K. conducted focus groups and contributed to revising the manuscript; R.R. and P.K. were also involved in the data analysis. T.R. and M.P.S. were responsible for the design of the focus group study, conducted focus groups, and contributed to revising the manuscript. G.W. and L.B. contributed to revising the data analysis protocol and the manuscript, and G.W. was also involved in conducting the focus groups. C.C. contributed to revising the manuscript. J.T. was involved in conducting and designing the focus groups, revising the data analysis protocol, drafting and revising the manuscript, and supervision.

References

(2019) 'Research Integrity is Much More than Misconduct', *Nature*, 570: 7.
Abdi, S., Pizzolato, D., Nemery, B., et al. (2021) 'Educating PhD Students in Research Integrity in Europe', *Science and Engineering Ethics*, 27: 1–19.

Alexander, M. and Williams, W. (2008) A Guidebook for Teaching Selected Responsible Conduct of Research Topics to A Culturally Diverse Trainee Group., 1st edn. Rockville: The Children's Hospital of Philadelphia.

Alfredo, K. and Hart, H. (2011) 'The University and the Responsible Conduct of Research: Who Is Responsible for What?' Science and Engineering Ethics, 17: 447–57.

Anderson, M. S., Horn, A. S., Risbey, K. R., et al. (2007) 'What Do Mentoring and Training in the Responsible Conduct of Research Have to Do with Scientists' Misbehavior? Findings from a National Survey of NIH-funded Scientists', Academic Medicine, 82: 853–60. Antes, A. L., English, T., Baldwin, K. A., et al. (2018) 'The Role of Culture and Acculturation in Researchers' Perceptions of Rules in Science', Science and Engineering Ethics, 24: 361–91.

Antes, A. L., Murphy, S. T., Waples, E. P., et al. (2009) 'A Meta- analysis of Ethics Instruction Effectiveness in the Sciences', Ethics & Behavior, 19: 379–402.

Asai, A., Okita, T., and Enzo, A. (2016) 'Conflicting Messages Concerning Current Strategies against Research Misconduct in Japan: A Call for Ethical Spontaneity', Journal of Medical Ethics, 42: 524–7.

Aubert Bonn, N. and Bouter, L. (2021) 'Research Assessments Should Recognize Responsible Research Practices—Narrative Review of a Lively Debate and Promising Developments'.10.31222/osf.io/82rmj.

Aubert Bonn, N. and Pinxten, W. (2021) 'Rethinking Success, Integrity, and Culture in Research (Part 2)—A Multi-actor Qualitative Study on Problems of Science', Research Integrity and Peer Review, 6: 1–18.

Ayres, L., Kavanaugh, K., and Knafl, K. A. (2003) 'Within-case and Across-case Approaches to Qualitative Data Analysis', Qualitative Health Research, 13: 871–83.

Boehme, O., Foeger, N., Hiney, M., et al. (2016) 'Research Integrity Practices in Science Europe Member Organisations', Science Europe, 1–52.

DuBois, J. M. (2004) 'Is Compliance a Professional Virtue of Researchers? Reflections on Promoting the Responsible Conduct of Research', Ethics & Behavior, 14: 383–95.

ENERI. (2017), Report of Stakeholder/Focus Group Workshop in Athens. <https://eneri.eu/wp-content/uploads/2018/06/Report- Stakeholder-Workshop-Athens.pdf> accessed 13 Oct 2021.

Evans, N., Veldkamp, C., Valentini, E., et al. (2018), EnTIRE Stakeholder Consultation Report. <https://ec.europa.eu/research/ participants/documents/downloadPublic?documentIds=080166e5 be1fb1a0&appId=PPGMS> accessed 13 Oct 2021.

Fanelli, D. (2019), MLE on Research Integrity – Training & Education – Thematic Report No 4. <https://rio.jrc.ec.europa.eu/sites/ default/files/report/MLE%20RI%20No4%20-%20Training%20 and%20education.pdf> accessed 13 Oct 2021.

Fereday, J. and Muir-Cochrane, E. (2006) 'Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development', International Journal of Qualitative Methods, 5: 80–92.

Forsberg, E.-M., Anthun, F. O., Bailey, S., et al. (2018) 'Working with Research integrity—Guidance for Research Performing Organisations: The Bonn PRINTEGER Statement', Science and Engineering Ethics, 24: 1023–34.

Gaskell, G., Šˇcepanoviˊc, R., Buljan, R., et al. (2019), D3.2: Scoping Reviews Including Multi-level Model of Research Cultures and Research Conduct. <https://www.sops4ri.eu/wp-content/uploads/ D3.2_Scoping-reviews-including-multi-level-model-of-research-cult ures-and-research-conduct.pdf> accessed 13 Oct 2021.

Geller, G., Boyce, A., Ford, D. E., et al. (2010) 'Beyond "Compliance": The Role of Institutional Culture in Promoting Research Integrity', Academic Medicine, 85: 1296–302.

Godecharle, S., Nemery, B., and Dierickx, K. (2014) 'Heterogeneity in European Research Integrity Guidance: Relying on Values or Norms?' Journal of Empirical Research on Human Research Ethics, 9: 79–90.

Haven, T., Pasman, H. R., Widdershoven, G., et al. (2020) 'Researchers' Perceptions of A Responsible Research Climate: A Multi Focus Group Study', Science and Engineering Ethics, 26: 3017–36.

Haven, T., Tijdink, J. K., Martinson, B. C., et al. (2019) 'Perceptions of Research Integrity Climate Differ between Academic Ranks and Disciplinary Fields: Results from a Survey among Academic Researchers in Amsterdam', PLoS One, 14: e0210599.

Hermer´en, G., Zwart, H., Marušic, A., et al. (2019), MLE on Research Integrity – Final Report. <https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE%20RI_Final%20Report_0.pdf> accessed 13 Oct 2021.

Joynson, C. and Leyser, O. (2015) 'The Culture of Scientific Research', F1000Research, 4: 1–11.

Kalichman, M. (2007) 'Responding to Challenges in Educating for the Responsible Conduct of Research', Academic Medicine, 82: 870–5.

—— (2013) 'A Brief History of RCR Education', Accountability in Research, 20: 380–94.

—— (2014) 'Rescuing Responsible Conduct of Research (RCR) Education', Accountability in Research, 21: 68–83.

Labib, K., Roje, R., Bouter, L., et al. (2021) 'Important Topics for Fostering Research Integrity by Research Performing and Research Funding Organizations: A Delphi Consensus Study', Science and Engineering Ethics, 27: 1–22.

Marie-Claude, R. (2007) 'Who Is Responsible? Supervisors and Institutions Need to Focus on Training in the Responsible Conduct of Research and Change the Culture in the Laboratory', EMBO Reports, 8: 706.

Marusic, A., Wager, E., Utrobicic, A., et al. (2016) 'Interventions to Prevent Misconduct and Promote Integrity in Research and Publication', Cochrane Database of Systematic Reviews, 4: MR000038.

Masadeh, M. (2012) 'Training, Education, Development and Learning: What Is the Difference?' European Scientific Journal, 8: 62–8.

McGee, R., Schneeweis, D. M., Hitsman, B., et al. (2014) 'Beyond Responsible Conduct: Taking Responsibility for the Responsible Conduct of Others', Journal of Microbiology & Biology Education, 15: 96.

Mejlgaard, N., Bouter, L. M., Gaskell, G., et al. (2020) 'Research Integrity: Nine Ways to Move from Talk to Walk', Nature, 586: 358–60.

National Academies of Sciences, Engineering, and Medicine. (2017) 'Education for the Responsible Conduct of Research'. In National Academies of Sciences, Engineering, and Medicine (ed.): Fostering Integrity in Research, pp. 195–204. Washington, DC: National Academies Press.

OECD Global Science Forum. (2007), Best Practices for Ensuring Scientific Integrity and Preventing Misconduct. <http://www. oecd.org/science/inno/40188303.pdf> accessed 13 Oct 2021.

Pizzolato, D., Abdi, S., and Dierickx, K. (2020) 'Collecting and Characterizing Existing and Freely Accessible Research Integrity Educational Resources', Accountability in Research, 27: 195–211.

Pizzolato, D. and Dierickx, K. (2021) 'Stakeholders' Perspectives on Research Integrity Training Practices: A Qualitative Study', BMC Medical Ethics, 22: 1–13.

Powell, S. T., Allison, M. A., and Kalichman, M. W. (2007) 'Effective- ness of A Responsible Conduct of Research Course: A Preliminary Study', Science and Engineering Ethics, 13: 249–64.

Ravn, T. and Sørensen, M. P. (2021) 'Exploring the Gray Area: Similarities and Differences in Questionable Research Practices (Qrps) across Main Areas of Research', Science and Engineering Ethics, 27: 1–33.

Resnik, D. B. and Shamoo, A. E. (2011) 'The Singapore Statement on Research Integrity', Accountability in Research, 18: 71–5.

Rifai, N., Annesley, T. M., Moore, S., et al. (2019) 'Maintaining Research and Publication Integrity', Clinical Chemistry, 65: 230–5. Satalkar, P. and Shaw, D. (2018) 'Is Failure to Raise Concerns about Misconduct a Breach of Integrity? Researchers' Reflections on Reporting Misconduct, Accountability in Research, 25:311–39.

Šcepanovic, R., Labib, K., Buljan, I., et al. (2021) 'Practices for Research Integrity Promotion in Research Performing Organisations and Research Funding Organisations: A Scoping Review', Science and Engineering Ethics, 27: 1–20.

Science Europe. (2017), Advancing Research Integrity Practices and Policies: From Recommendation to Implementation. Brussels.

*<https://www.scienceeurope.org/media/lfdme4ev/ws_report_integri
ty_practices_policies.pdf> accessed 13 Oct 2021.*

*SOPs4RI. (2020), Guideline for Promoting Research Integrity in Research Performing
Organisations. <https://ec.europa.eu/ info/funding-
tenders/opportunities/docs/2021-2027/horizon/guidance/guideline-for-promoting-
research-integrity-in-research-perform ing-organisations_horizon_en.pdf> accessed
13 Oct 2021.*

*Sørensen, M., Ravn, T., Bendtsen, A., et al. (2020), D5.2: Report on the Results of the Focus
Group Interviews. <https://sops4ri.eu/wp-content/uploads/D5.2_Report-on-the-
Results-of-the-Focus-Group-Inte rviews-1.pdf> accessed 13 Oct 2021.*

*Sørensen, M. P., Ravn, T., Marušic, A., et al. (2021) 'Strengthening Research Integrity: Which
Topic Areas Should Organisations Focus On?' Humanities and Social Sciences
Communications, 8: 1–15.*

*Swain, J. (2018) A Hybrid Approach to Thematic Analysis in Qualitative Research: Using A
Practical Example. London: Sage Publications Ltd.*

Titus, S. and Bosch, X. (2010) 'Tie Funding to Research Integrity', Nature, 466: 436–7.

*Valkenburg, G., Dix, G., Tijdink, J., et al. (2021) 'Expanding Research Integrity: A Cultural-
Practice Perspective', Science and Engineering Ethics, 27: 1–23.*

*Watts, L. L., Medeiros, K. E., Mulhearn, T. J., et al. (2017) 'Are Ethics Training Programs
Improving? A Meta-analytic Review of Past and Present Ethics Instruction in the
Sciences', Ethics & Behavior, 27: 351–84.*

*Wells, J. A., Thrush, C. R., Martinson, B. C., et al. (2014) 'Survey of Organizational Research
Climates in Three Research Intensive, Doctoral Granting Universities', Journal of
Empirical Research on Human Research Ethics, 9: 72–88.*

*Zeng, W. and Resnik, D. (2010) 'Research Integrity in China: Problems and Prospects',
Developing World Bioethics, 10: 164–71.*

Supplementary materials

Table A.1: *Characteristics of focus group facilitators and observers*

Researchers → **Focus groups↓** *Discipline,* Type of focus group		Initials, Credentials (Gender)	**Facilitators** **Occupation; experience and training**
Humanities			
Researchers only	Netherlands*	WK, PhD (M)	Postdoctoral researcher, trained in qualitative research methods & focus groups
		ARE, MA (F)	PhD student; trained in qualitative research methods & focus groups
		KL, MA/MSc (F)	PhD student; trained in qualitative research methods & focus groups
	Denmark	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Croatia	RR, Mag. iur./ML(F)	PhD student; experience in conducting focus groups and interviews
		IB, PhD (M)	Post-doctoral researcher, trained in qualitative research methods; experience in conducting interviews and focus groups
Mixed	Netherlands*	KL, MA/MSc (F)	PhD student; trained in qualitative research methods & focus groups

Researchers → **Focus groups**↓ *Discipline,* Type of focus group		Initials, Credentials (Gender)	**Facilitators** Occupation; experience and training
	Spain	GW, PhD (M)	Professor; experienced facilitator
		TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Germany	ARE, MA (F)	PhD student; trained in qualitative research methods & focus groups
		JB, MSc (F)	Project coordinator; experience in qualitative research and trained in focus groups
	Greece	ES, PhD (F)	Postdoctoral researcher; Philosopher with knowledge in qualitative methods
		VM, MSc, (M)	Junior researcher with experience in participating in workshops
Social sciences			
Researchers only	Germany*	ARE, MA (F)	PhD student; trained in qualitative research methods & focus groups
		JB, MSc (F)	Project coordinator; experience in qualitative research and trained in focus groups
	Spain	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Netherlands	WK, PhD (M)	Postdoctoral researcher, trained in qualitative research methods & focus groups
		ARE, MA (F)	PhD student; trained in qualitative research methods & focus groups
Mixed	Netherlands*	KL, MA/MSc (F)	PhD student; trained in qualitative research methods & focus groups
		JT, PhD (M)	Assistant professor, experienced facilitator
	Denmark	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Croatia	RR, Mag. iur./ML (F)	PhD student; experience in conducting focus groups and interviews
		IB, PhD (M)	Post-doctoral researcher, trained in qualitative research methods; experience in conducting interviews and focus groups
	Greece	ES, PhD, (F)	Researcher; Philosopher with knowledge in qualitative methods
		VM, MSc, (M)	Junior researcher with experience in participating in workshops

Table A1 continued

Researchers → **Focus groups**↓ *Discipline,* Type of focus group		Initials, Credentials (Gender)	**Facilitators** **Occupation; experience and training**
Natural sciences			
Researchers only	Croatia*	RR, Mag. iur./ML (F)	PhD student; experience in conducting focus groups and interviews
		IB, PhD (M)	Post-doctoral researcher, trained in qualitative research methods; experience in conducting interviews and focus groups
	Spain	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Denmark	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Belgium	WK, PhD (M)	Postdoctoral researcher, trained in qualitative research methods & focus groups
		ARE, MA (F)	PhD student; trained in qualitative research methods & focus groups
Mixed	Netherlands*	KL, MA/MSc (F)	PhD student; trained in qualitative research methods & focus groups
		JT, PhD (M)	Assistant professor; experienced facilitator
	Denmark	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Croatia	RR, Mag. iur./ML (F)	PhD student; experience in conducting focus groups and interviews
		IB, PhD (M)	Post-doctoral researcher, trained in qualitative research methods; experience in conducting interviews and focus groups
	Greece	PK, MSc, PhD (M)	Senior researcher, experience in facilitating focus groups
		VM, MSc (M)	Junior researcher with experience in participating in workshops

Table A1 continued

Researchers → **Focus groups↓** *Discipline,* Type of focus group		**Initials, Credentials (Gender)**	**Facilitators** **Occupation; experience and training**
Medical sciences			
Researchers only	Croatia*	RR, Mag. iur./ML (F)	PhD student; experience in conducting focus groups and interviews
		VT, MSc (M)	PhD student; trained in qualitative research methods; experience in conducting interviews and focus groups
	Denmark	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Greece	ES, PhD, (F)	Researcher; Philosopher with knowledge in qualitative methods
		VM, MSc, (M)	Junior researcher with experience in participating in workshops
	Netherlands	KL, MA/MSc (F)	PhD student; trained in qualitative research methods & focus groups
		JT, PhD (M)	Assistant professor; experienced facilitator
Mixed	Belgium*	WK, PhD (M)	Postdoctoral researcher, trained in qualitative research methods & focus groups
		ARE, MA (F)	PhD student; trained in qualitative research methods & focus groups
	Denmark	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Spain	TR, PhD (F)	Assistant professor; extensive experience in qualitative research
		MPS, PhD (M)	Senior researcher; extensive experience in qualitative research
	Italy	GV, PhD (M)	Professor; experienced facilitator

Table A1 continued

**Indicates focus groups in which the topic of RI education and training was explicitly discussed in depth.*

Table A.2: Information on dropouts from focus groups

Focus groups↓ *Discipline,* Type of focus group		**Number of dropouts**	**Junior/senior researcher or other stakeholder?**	**Reasons for dropout**
Humanities				
Researchers only	Netherlands*	0	-	-
	Denmark	2	Senior researcher Senior researcher	Personal circumstances
	Croatia	0		
Mixed	Netherlands*	1	Senior researcher	Personal circumstances

Focus groups↓ *Discipline,* Type of focus group		Number of dropouts	Junior/senior researcher or other stakeholder?	Reasons for dropout
	Spain	1	Senior researcher	Personal circumstances
	Germany	1	Other stakeholder	Personal circumstances
	Greece	1	Senior researcher	Competing obligation
Social sciences				
Researchers only	Germany*	0	-	-
	Spain	2	Senior researchers	Urgent competing obligations Change of mind
	Netherlands	0	-	-
Mixed	Netherlands*	2**	Other stakeholder Senior researcher	Competing obligation; Unknown (no show)
	Denmark	2	Senior researcher Other stakeholder	Change of mind Competing obligations
	Croatia	2	Junior researcher Other stakeholder	Other obligations
	Greece (conducted online)	3***	Senior researcher	Competing obligation; Competing obligation; Competing obligation;
Natural sciences				
Researchers only	Croatia*	0		
	Spain (conducted online)	4	Senior researcher Junior researchers Junior researchers Junior researchers	Covid-19 lockdown caused inability to participate for the four drop-outs
	Denmark	0	-	-
	Belgium (conducted online)	0	-	-
Mixed	Netherlands*	0	-	-
	Denmark	1	Other stakeholder	Personal circumstances
	Croatia (conducted online)	0		
	Greece	0	-	-

Table A.2 continued

Focus groups↓ *Discipline,* Type of focus group		Number of dropouts	Junior/senior researcher or other stakeholder?	Reasons for dropout
Medical sciences				
Researchers only	Croatia*	2	Senior researcher Senior researcher	Date of the focus group not suitable Other obligations
	Denmark (conducted online)	3	Other stakeholder Senior researcher Senior researcher	Covid-19 lockdown caused inability to participate for the three drop-outs
	Greece	2	Senior researcher Senior researcher	Unknown; Competing obligation
	Netherlands	0	-	-
Mixed	Belgium* (conducted online)	1	Senior researcher	Competing obligations
	Denmark (conducted online)	2	Senior researcher Other stakeholder	Unknown (no show) Personal circumstances
	Spain (conducted online)	2	Junior researcher Other stakeholder	Covid-19 lockdown
	Italy (conducted online)	0	-	- Table A.2 continued

*Indicates focus groups in which the topic of RI education and training was explicitly discussed in depth.
**We had originally invited 4 additional people to this focus group (3 senior researchers, 2 junior researchers), but due to the high number of participants, we had to cancel their participation.
*** The original focus group took place with physical presence with 5 interviewees. However the recording was stolen, together with the hard disk of the secure server of the laboratory, where the files were kept. NTUA notified the WP5 leaders, according to the privacy policy of the focus group study. The focus group was repeated via online means, but only 2 out of the original 5 interviewees participated.

Table A.3: Example quotes about the theme 'Training availability'

Theme/subtheme	Quotes
Training availability	"And what's important is that absolutely everyone is to take that training, it's not just the junior." – Senior researcher, humanities, Netherlands
	"But if I teach PhD students and I tell them this is how to do things right, this is how to do things in [a responsible] way, then they go back to the labs where the norms are different and where they have to sort of fight against supervisors who are doing things the way they've always done them, which is fine because they just learnt it that way. But now there are new insights, new cultures to create. But then the PhD students have to do all the work. So I think it's very important when you think about education to make sure you include all levels." – Senior researcher/RI teacher, social sciences, Netherlands
At the pre-doctorate level	"I think it's important that mainly at the start of that career as a scientist, it's important to have such a training. Of course also later on it's important, but I think

Theme/subtheme	Quotes
	doctoral school is a good place to do that." – Senior researcher/Research director, medical sciences, Belgium
	"And I think it's, it would be good and helpful to start with that even on graduate or undergraduate level. Maybe not just like full course, but maybe part of some lectures. And then on the PhD level and, it would be definitely, it should be whole course, yeah." – Junior researcher, natural sciences, Croatia
	"There were no systematic trainings on good, scientific practices when I started. Up to now, there was no training in our research institute, also the university I studied my PhD." – Junior researcher, social sciences, Croatia
	"So I did [an RI] course last year. It's actually made, for I think first year PhD students but it wasn't designed when I was a first year PhD student." – Junior researcher, social sciences, Netherlands
At the post-doctorate level	"There are so many gaps, relating specifically to the postdoc experience because…that I go to staff meetings but 85% of what is said, is not something I can do anything about or with… I don't go to the PhD meetings because there all about the PhD experience, which I've already gone through…If anybody is coming to do research in any capacity they should be folded into this training perspective." - Junior researcher, humanities, Netherlands
	"<u>Senior researcher 1</u>: I think earlier studies indicated that the most difficult part, or where do things go most easily wrong have to do with PhD supervisee relationship, so some sort of training on the part of the university to train university professors on PhD supervision. <u>Senior researcher 2</u>: With this, I have that actually, once a year I think, or twice a year they have that now." – Humanities, Netherlands
	"But it's [i.e. RI education is] also something that even senior people should occasionally have. Have a chance to go and listen to some type of education. Because these things change and the requirements that you are facing are changing, and fields are changing. So maybe, for example, later when we come to this open science and things, maybe, senior people may not know this and it can also be a problem for them so I think they should get the opportunity to get educated." - Senior researcher, natural sciences, Croatia
	"Even though you're a researcher and you have a PhD degree, you follow the, the steps of advancement, it's lifelong learning. And you get PhD students who are experts in their little field and you're a supervisor, but you may not know all about what they're doing. So, I think it's a matter of acknowledging that we also have[…]competence [to] develop, just as everyone else." – Senior researcher, medical sciences, Denmark
	"it's really, really hard to train the PIs and that is what we are facing now at the ethical review board that a lot of the resistance to these new kinds of procedures is with the more experienced researchers, but, it is really hard to get them trained." – Secretary of RE committee, natural sciences, Netherlands
For other stakeholders	"And also for the, for this decision makers or the deans, for the rectors, there should be also written something because they are ruling the, they have a large responsibility. If they fail then they can create larger problem." – Senior researcher, natural sciences, Croatia

Table A.3 continued

Table A.4: *Example quotes about the theme 'Education and training approaches, goals and content'*

Theme/*subtheme*	Quotes
Training goals approach and content	"I think it's important to cover every aspect of research integrity. At least just to mention it so the, the students, or the, the senior professor know what research integrity is and to bear in their minds what, what all the topics of, of interest there are." – Junior researcher, natural sciences, Croatia
	"And when we start with students, we must start on basic things. And we think it is not enough. Because they don't have enough time and place for everything." – Senior researcher, medical sciences, Croatia
	"And of course the content of such course depends on the field of research." - Funder, medical sciences, Belgium
	"Then align this line [i.e. RI education goals and approach] so at different stages of career and different subjects." - Junior researcher, social sciences, Germany
Addressing research misbehaviors	"And I am not sure if when we talk about fraud being committed is fraud being committed on purpose, which is a criminal issue, or whether it has to do with ignorance. For example, many students commit plagiarism without realizing it[...]For me what we need is education" – Senior researcher, medical science, Greece
	"We have these workshops on scientific integrity and it's about fraud cases and everyone, at least young PhD students, and they get scared and they think 'That's really bad, I would never do anything like that'". – RE committee secretary, natural sciences, Netherlands
	"The other side of the coin, it's exactly the same talking about misconduct or bad practices because when I plan for what is good practice it means whatever isn't bad." – Junior researcher, social sciences, Germany
	"This is to me saying like: 'Okay, this is bad okay'. But not really providing the opportunities or not changing anything that drives people towards that behavior". – Senior researcher, social sciences, Netherlands
Focusing on knowledge of best practices, rules and regulations	"I think that's very important[...]and expecting these things to be run by host institutions[...]so not putting the full responsibility for this on individual researchers but on the institution, and having the institution run different procedures...training, audits or whatever, to make sure that all its research performing employees are actually aware of current rules or changes in rules." – Funder, medical sciences, Denmark
	"make sure that everyone is aware of what the rules about authorships are." – Senior researcher, medical sciences, Croatia
	"We are higher education researchers and from my everyday research practices there's always the question of citation practices: How to do citations correctly? A very big issue is when you publish a paper and working with quantitative data, quantitative analyses. There are probably guidelines regarding transparency and data, depending on the journal, but they are very different." – Junior researcher, social sciences, Germany
	"often the rules are conducted or taught to students in a such way that they are ambiguous and not detailed enough. So when you actually try to apply them or go deeper and study them, then you, at least in my experience, what you see in studying them is that they were presented to you in mostly the wrong and oversimplistic way." – Junior researcher, medical sciences, Croatia

Theme/subtheme	Quotes
	"I think it's true that now it's flipped over too much on the forms and bureaucracy and rules. But I think there's also been decades where there was too little emphasis on the ruler." – Junior researcher, social sciences, Netherlands
	"and also it doesn't help to be too prescriptive because I have little bit of fear of the situation that is present at some foreign universities and maybe especially in USA or in some, let's say countries in which law is very strict and in details[…]I think it is not good situation that you have to go to committees before you are allowed to do your research. It would be too stifling for creativity; very, very bad. And our present situation, in my opinion, is not very bad regarding this framework because we have some ethical guidelines and other literature and maybe this kind of education, discussion is something that could be, that could advice more inform people about that topics." – Senior researcher, humanities, Croatia
Inspiring and internalizing RI	"now through research integrity, the whole discussion comes from another side. Its entering the university, not because from the inside we believe in these things but well, you know, our professional[ism] is under threat and the public imagery is turning on its head and I think when it comes to norms and values we do need bigger stories, stories which anchor these values" – Senior researcher/RI committee member, humanities, Netherlands
	"you learn[…]that [rule following] is making intrinsic, originally intrinsic motivation extrinsic, which means that you no longer want this, you don't understand this, you're no longer embodying whatever we're talking about" – Senior researcher, social sciences, Netherlands
	"if there is any cognitive effect on moral behavior, it should have to do with seeing, really seeing that some things are valuable in themselves, or at some end, worthwhile to pursue, not because they're instrumental for another reason. So this requires some sort of, really in depth moral reflection" – Senior researcher/RI committee member, humanities, Netherlands
	Table A.4 continued
Addressing daily struggles and gray areas of research	"But there's a whole range of gray area where as a researcher, especially young researcher, you sometimes don't really know what is right or what is wrong. So I would be very happy to be additionally educated better, what are the rules with outliers, what is the best way, what are the options?" – Junior researcher, social sciences, Netherlands
	"This could be a real problem because sometimes when you have project, day-to-day work, practice issues can contradict legal, ethical or whatever issues." – Junior researcher, social sciences, Germany
	"We get many contacts from researchers that are in doubt, 'Is this correct to write like this?'" – Research integrity officer, natural sciences, Denmark
	"Because also sometimes it doesn't only depend, of course, on disciplines, but also on journals you want to publish in and so on and so on, so rules are not fixed, so it's not always easy to give researchers the correct information about the questions they have about it." – Valorization officer, medical sciences, Belgium
	"But if you focus more on specific topics where they are really, you know, in their daily practice they have issues with it, I think it's easier to attract [researchers]." – Senior researcher/Research director, medical sciences, Belgium
	Table A.4 continued

***Theme**/subtheme*	**Quotes**
	"[…]on the one hand there's nothing wrong with having courses that are idealizing the situation […] That's okay if you have both sides […] I'm not against my students having these courses, but I do want to tell them that now they need to kind of meet reality. That needs to enter the teaching, or education." – Senior researcher, social sciences, Netherlands
Empowering and supporting researchers	"something that isn't mentioned in this ethics game, like how do you deal with the fact that you're actually, as a PhD student, subject to certain power dynamics that you really have very little say in[…]So it's just students doing it on their own time, really, they're a kind of… become tennis balls." – Senior researcher, humanities, Netherlands
	"It has to do with raising awareness, too. Just knowing what rights you could have if it was an ideal world. That helps as well." – Junior researcher, social sciences, Germany
	"We can also train our students, to have, sorry for the gender, for the lack of a better gender, to have balls, female or male to stand up for their thing. And without that you cannot become a good researcher." – Senior researcher, social sciences, Netherlands
	"But maybe just to comment on the discussion about power dynamics: I think it would be good then just to make PhD students aware really from the very start what the guidelines are and what their options are, kind of, to empower them from the beginning of their… before they start publishing, so maybe some education or training in that regard would help." – Junior researcher, social sciences, Netherlands
Building a culture of integrity	"The only thing that should be mandatory is that there's a culture of research integrity and there are all kinds of instruments that you can think about to promote this culture, and so at [institution name], that's why we made it [i.e. RI training] mandatory for PhD students and PhD professors, but in an ideal world you would not have any ethics training but the culture of integrity, that you have disciplined." – Senior researcher/RI committee member, humanities, Netherlands
	"I think it's more beneficial to try to create a common culture that people are socialized into and understand what the expected standards are than to try and monitor at every step, which is what I think is happening now, whether they are applying, whatever rules exist as the should be, in that sense, I do find that more productive. Make sure people have the same standards, same expectations." – Senior researcher, humanities, Netherlands
	"You need an error culture in order to get better. Meaning what, meaning that you are free and open to admit errors and talk about them because not everyone who makes mistakes is evil." – Senior researcher, social sciences, Netherlands
	"I mean still, I mean maybe it [i.e. RI education] will be helpful anyway at least for the young people, for the, change the culture, the mentality." – Senior researcher, natural sciences, Croatia
	"I am super curious how you are going to change an existing culture with a training session". – Senior researcher, natural sciences, Netherlands

Table A.4 continued

Table A.5: *Example quotes about the theme 'Motivations and incentives for participation in RI training'*

Theme/subtheme	Quotes
Motivations and incentives	-
Attractiveness of training programs	"The thing with courses is that the people who need it won't do it. It's always the case. The bad supervisors who are not interesting in supervising, they will not go to a 'how do I supervise PhDs' kind of training. They won't." – Senior researcher, humanities, Netherlands "Research integrity is not what people want to put a lot of time in. And even with doctoral PhD students it's very hard to get them to courses of research integrity, they are more interested in other courses about a subject or other transferable skills courses, but research integrity is not a favorite. So, we have to force them, let's say, a little bit, to follow these courses." – Senior researcher/Research director, medical sciences, Belgium "Here it's where it lies on their priority list, right? It's just not, like you say, they; everyone, everyone thinks that they are, that they have good integrity...'I don't have three hours of afternoon to dedicate to this; I have to fly to this conference, I have to do this, I have to.' So, maybe it's just not priority…"- Research support staff, natural sciences, Netherlands "Senior researcher 1 : Make it easier by dropping the term. Because this is exactly the religious idea now: Oh there are people who are good, and there are people who are bad, and we don't want to talk about this. So, learn the right thing, and then if you follow, if you're a follower then you're a good person, then you will get to heaven[…] Senior researcher 2/RI teacher: You want to facilitate those courses, I think there was a good comment about maybe not naming them some black and white term, so maybe just scientific conduct' instead of 'integrity course' or 'moral thing' or- It's doing research, that's what it is." - Social sciences, Netherlands "Senior researcher: Yeah, so make it very recent. So, maybe make a promise like in the training will be the, I don't know, the last five years grey-issues that came up, something like that. RE committee secretary: People want it like case studies. Privacy officer: Which is maybe a little bit like gossip, you know, people are interested in gossip. So, what came up in that company? Ok. RE committee secretary: I recognize that, sort of like the case studies that makes it really alive, and then they can also relate it a little bit easier to their own work." – Natural sciences, Netherlands
Integrating RI education into existing structures	"I've been thinking about a couple of things here, but mainly I think it should just be in the existing things that are in place, like existing forms, the whole fabric of the science that we do. We have a yearly evaluation talk with our supervisor, everybody has that no matter your level, right? Also the dean. If that form for instance would say 'What did you do this year to put quality over quantity?' for instance, and you can give any example there, just make sure that you have some quality mark. Maybe you joined a course on integrity…" – Senior researcher/RI teacher, social sciences, Netherlands "And you can also use that competitiveness [in research]. So, the reason that it is competitive is because there is a reward for doing it, like that. You can do a similar thing implementing around the ethics part. Make that also part of; so that you can actually get a reward out of it." – RE committee secretary, natural sciences, Netherlands

Theme/subtheme	Quotes
	"It will surely create a lot of resistance [if funders require that the PI on a project should go through some sort of research integrity course]" – RI officer, natural sciences, Denmark "Yeah, but is it important [to require from RPOs that these things are in place before giving out funding]? I'm not sure whether it's important, I would see it as an incentive to take a course in research integrity and so on, but I could see it as an obstacle actually." – Senior researcher, social sciences, Denmark
Making RI training obligatory	"I can tell you in our school, before you become assistant professor or docent one of the prerequisites is that you do a so called teaching practice course that goes on for three days. Because, you know, when you become assistant professor you'll be giving a lectures and this is some sort of preparation for that. Why shouldn't t we introduce a course for the research integrity, AS WELL to that. And make that [a] prerequisite. Because, you're expected to be a supervisor of either, you know, the final, the thesis or PhD programs. When you become full professor, in order to get, to get that title well you have to the management course. Because once you're full professor you have a potential to become a head of the department, or even associate professor. So in that academic promotion, in those steps, well let's introduce obligatory courses that people have to, that people have to do it. There's no negotiation about that. You have to tick that. I've done it and I've got a document saying that I've done it." – Junior researcher, medical sciences, Croatia "[Whether training should be made mandatory] depends on the kind of training and the kind of review, so again, if you're dealing with projects where you have to work along through a certain code of conduct, then obviously that has to be reviewed and that has to be mandatory, but, you know, if you want to do a yearly review of referencing, then I don't think that necessarily has to be mandatory, referencing and plagiarism." – Senior researcher, humanities, Netherlands

<div align="right">*Table A.5 continued*</div>

Table A.6: Example quotes about the theme 'Organization of RI education and design of trainings'

Theme/subtheme	Quotes
Organization and design of RI training	-
Continuity of training	"I think it's very important to have such a reoccurring event because I noticed that the new PhD students, so also the PhD students I am advising, when you are talking about stuff, they were like 'I would never do that'. And for me, my fifth year, I was like well that might have sometimes occurred. So I think some things are also, what you mentioned, we know the rules but sometimes it's a gray area. And I think if you're starting your PhD, there are some things which I would say that would never happen to me and then think along the way." – Junior researcher, social sciences, Netherlands "So I wanted to say that I think that one class or one course is not enough." – Senior researcher, medical sciences, Croatia "So that in a reoccurring course throughout all levels, I think it's important to make sure that you calibrate your compass each time to make sure you can make decisions. Because there's going to be new decisions, things you haven't thought about." – Senior researcher/RI teacher, social sciences, Netherlands "even though you're a researcher and you have a PhD degree, you follow the, the steps of advancement, it's lifelong learning. And you get PhD students who are experts in their little field and you're a supervisor, but you may not know all about

	what they're doing. So, I think it's a matter of acknowledging that [...]we also have[...]competence [to] develop, just as everyone else" – Senior researcher, medical sciences, Denmark
Format of training	"[Formal training] only makes sense if there are very specific guidelines that you want people to know about and those guidelines change. If we are not speaking about that kind of framework then I also don't think there is a need for explicit training." – Senior researcher, humanities, Netherlands

"not every part of research integrity can be trained in a training module or course but has to be internalized by doing. We already discussed aspects of research integrity that relate to rules and regulations that can be learned…but not every aspect of research integrity can be trained explicitly." – Senior researcher, social sciences, Germany

"look at these walls here, it's not really inspiring, and a picture of a person you think, well that was really a good scientist, a good academic, you can start talking about her or about him, share the stories. This is what… many people have said beautiful things, give a nice quotation, that's really meaningful, deep, just I mean, this is a bit of a littered environment, right? It isn't easily inspiring but we have been in the midst, we are the heirs of really inspiring people who founded the university, who had big ideas… well, a quotation here and there, a good picture, conversations about these people over coffee time, that's really interesting, I mean that's what it's…" – Senior researcher/RI committee member, humanities, Netherlands

"I think there's no way, I mean you have to do things in order to understand what that means. Just to further simulate, give examples is not enough. Because, dealing with real things, nothing is as, as, impressive as dealing with real data. It's your data and you are committed and you want something or not or whatever" – Senior researcher, social sciences, Netherlands

"I think a number of people now mention specific people or they are part of particular groups like descriptive linguistics or experimental linguistics and within their fields they kind of have the person they go to or they know the form. I mean I'm pretty much working by myself and I'm not part of any particular group so for me to find out what these kind of conventional forms are I would have to kind of hear it through the grapevine or kind of find out who uses what. So it would be good to have kind of a single source as well as the data privacy officer and so on. It would be good to have like a website and just know… or maybe I didn't do enough research but something like, quite sensible to everyone." – Junior researcher, humanities, Netherlands

"No, I was just thinking, if it should be a course then it should definitely not be like an entire online course, because then it will just be some passive compliance. Then it should be something in kind of a physical space that could facilitate some reflections on a given topic, relevant." - RI officer, social sciences, Denmark

"I have [the topic] responsible supervision and mentoring [to rank] which I will put here because it's learning by doing. If you're supervised in a correct and responsible way you will behave automatically as you should, so it promotes research integrity indirectly. It's important. [The topic] education and training and research integrity is second for the same reason because it's better to learn rules by doing and really live them than just learn the rules by rote. On top, not every part of research integrity can be trained in a training module or course but has to be internalized by doing." – Senior researcher, social sciences, Germany |

Table A.6 continued

	"I think it's very important to have more like counselling types of people, who, you know approaching them, like they are very accessible and like it's the normal thing to go, you know. People, I hear a lot of stories of students mainly coming back with things that are happening, but they don't know where to go. It says on the website there's a committee for things, but that all sounds very formal. So why would you take that step?" – Senior researcher/RI teacher, social sciences, Netherlands
	"that kind of training [i.e. formal course] doesn't make sense if the legalistic framework isn't there. So that was a little bit, like proviso, but…because there are very specific things you want people to know about here, everything is a little bit documented, so you can't easily build that kind of training either, I think. I don't know, that's my feeling and I'm not sure…" – Senior researcher, humanities, Netherlands
RI trainers	"I think that these courses should be taught by people who have some research experience, because they have-, they should have been in these gray areas to say, to have real examples of 'Okay, this was a problem. I dealt with this, but I don't know if I did correctly. What do you think?'"- Senior researcher, social sciences, Netherlands
	"It's going to sound incredibly cynical, but usually there's just such a gap between what they [i.e. trainers] know and what you need to know, because often, as you we're saying before, it requires such medium specificity, or at least, you know, disciplinary specificity. So I think, yes all supervisees should be trained but you then have to, kind of, come up with a trainer, or a training body, that really is aware of what they're talking about, you know, not sort of post-its on a pyramid in a board or whatever." – Senior researcher, humanities, Netherlands
	"we have divided it [i.e. RI teaching] over the people from part of the committee for scientific integrity[…]and we divide these courses over these people[…]because we feel that the people who… well, have to solve the problem, so to say, also talk to the people who may cause them, but there's many other ways to think about." – Senior researcher/RI committee member, humanities, Netherlands
	"But that would mean then, each and every teacher or course leader should be aware of that. And some sort of authority should be able to educate them so they actually raise their awareness of that some compliance in their curriculum. That would barely vary from subject to subject"– Junior researcher, humanities, Croatia
Context of training	"The engineering department is completely different than social sciences or humanities. From one supervisor to another everything can change. So, these are things I think we need to take into account when talking about research integrity" – Junior researcher, social sciences, Germany
	"just to stress that in case this would be promoted it [i.e. RI training] should be very multidisciplinary again. That it should be broad." – Senior researcher, natural sciences, Spain
	"I think, so some of people here are also part of the open science community[…] What we aim to do is also have these kinds of discussions with people to make sure that everybody is involved[…]I'll give one short example: we are asked to store our data in an archive in [city] and anthropologists are also asked to do that, but they have a very different relationship with their participants and it's not like they can just share these things" – RI teacher/SR, social sciences, Netherlands

Table A.6 continued

"Of course, of course in Croatia th[ese] issues are related to the capacities basically. Also not just for what we want to be there. Because of the capacity you know. If you want to fulfil all of this things, you need to have some, something in your institute and we have nothing basically. It's not necessary[…]that we should have as institution, but it should be done on, on regional level like other things, for example, for leading projects whatever you know, offices, PR, whatever. It's not necessary that every institution, small institution should have this. But then you should organize the systems somehow different." – Senior researcher, medical sciences, Croatia

Table A.6 continued

Table A.7: *Example quotes about the theme 'Factors influencing the implementation of RI education and training policies'*

Theme/subtheme	Quotes
Implementation factors	-
Individual factors	"And maybe you can learn all the rules about scientific integrity, but the individual is also important and that's sometimes very difficult to change. The individual characteristics of the people are also important and play an important role in scientific integrity, I think." – Senior researcher/research director, medical science, Belgium "But, I think you can train people and having a code is very good to remind people to be, yeah, true etcetera, etcetera. But, in the end if I am a clever researcher and I want to fraud, play; make fraud, I think I can get away with it to be honest. I think, researchers are clever enough to do it. And, I think whatever you will come up with, they will find a way. So, if someone is really maleficent than I think you will get away with it." – Privacy officer, natural sciences, Netherlands "I have a colleague[…]and he is Australian and it seems like he has this quite legalistic approach as well, to resolving things and also behaviors within the university environment" – Junior researcher, humanities, Netherlands
Supervision	"We train the PhD students in these courses for half a day or one day during four years when they get trained by that professor or PI for four years, and I am pretty sure that there is quite often a PhD student that comes up here with some issue and the professor: 'Ah, don't worry about it.'" – Senior researcher, natural sciences, Netherlands "[T]here's very interesting research done by[…]and one of[…]findings was that the stronger the norms were vocally enforced the more unethical the behavior became because people were afraid to report mistake etc. etc. and so the behavior of the manager just above you was key to promoting integrity because you would copy that behavior. But, of course, you have to have the knowledge of, well, what the norms are, so I think we wouldn't have to overstate the importance of these trainings but we have to have them." – Senior researcher, humanities, Netherlands "I have responsible supervision and mentoring which I will put here because it's learning by doing. If you're supervised in a correct and responsible way you will behave automatically as you should, so it promotes research integrity indirectly." – Junior researcher, social sciences, Germany
Research environment	"Now I would like to come back to the training part because personally I did not have the chance to meet such cases of important fraud. What I have seen in all levels in France and in Greece, is the students not knowing. Misconduct happens, I have seen that, I would say minimum once a year. Very frequently, this is a problem of training we have to talk about it as well, the research environment, the creation

	of pressure, a good competition between labs (who is going to be the first)." - Senior researcher, medical sciences, Greece
	"The way that this is done here is at the lab-level. So it's labs that trains the students, the PhD students, master students, in the way they do things. And this creates a lot of diversion. So so, there are some groups that are very aware of the problem, some others don't care so much, they just look for productivity. And that happens everywhere, I think. So so, the the research integrity in some in some labs is not one thing that they really care about. I'm not saying they don't apply some basic rules, I'm just saying that it's not something they have carefully think about." – Senior researcher, natural sciences, Spain
	"If you don't have the dialogue and the people are feeling like I cannot discuss that I am insecure about this, then you will actually end up with cases that might have gone wrong, simply because the researcher was not sure about the right track and didn't dare to ask. So, it should probably more be about allowing all to learn, rather than the blaming culture." – Industry representative, natural sciences, Netherlands
	"Also, creating an internal atmosphere where that, that doesn't help people go in that direction. So, many PhDs may feel the only way for me to get this pressure off from my PI, or this constant like 'Where is my Nature paper? Where is my science paper?'... yeah, helping inform or train PIs of how to create an atmosphere that doesn't push people into, into that way." – Research support staff, natural sciences, Netherlands
	"What's happened to the university system in the Netherlands, in general since, sort of, neo-liberal turn, right, we're dealing with a culture that is ultimately focused on a structure of competition at the moment and particularly the discourse of excellence, and these kinds of things are breeding a particular kind of research environment and also an emphasis on, you know, a sort of mythology maybe, of what is valued and what isn't, and I think that if you're really wanting to take the big perspective on that, that is actually really killing a lot of, you know, research integrity in a different kind of way... it's the vanity of the excellent discourse" – Senior researcher, humanities, Netherlands
Trust versus oversight	"You should have that. The some committee of university or administrate, comes and says 'Let me see your document'. Although it looks like a bureaucracy the key is in the documents. If you don't have documents, you have nothing." – Senior researcher, medical sciences, Croatia
	"So, you have to train, I think training is the best way to engage; to have responsible researchers, instead of saying: You have to do this, you have to do this, you have to do this as a funder. I think, it also creates more distrust towards, towards researchers instead of the thing you want to achieve [which] is to have responsible researchers." – Privacy officer, natural sciences, Netherlands
	"For me what we need is education and then we have to clear about the later point of implementation. But we must not ignore the educational, cultural point. I think this is very important, otherwise if you not cover the previous step of culture and education, we end up with simple check lists and we should avoid that." – Senior researcher, medical sciences, Greece

	"People are trusted and that's where I see the lack of level of trust here. They're given the training, you might say, and then trusted that they are all doing their job properly without the need to, every time, three, four times a semester, check whether they are applying what they should be." – Senior researcher, humanities, Netherlands
	"But in general requiring courses is part of a general suspicion against people working here, and I want the best for my group, and for the people in my group and I have no interest in harming anybody, neither with my research nor with my colleagues. I may take on decisions but I don't know whether a course would really be the answer to that." - Senior researcher, humanities, Spain
Reward and incentive structures	"I think it also has to do, something with structures. If the structures are not good then you can train as long as you like, but you're never going to change what is going to happen really, in the grassroots." – Senior researcher/confidential counselor, humanities, Netherlands
	"Yes, definitely. I always think that it's not a lack [of awareness] of guidelines, but the way that the pressure upon researchers and the way that researchers are rewarded or not rewarded by funding, by other things, that causes most of the problems, that's in my personal opinion." – Senior researcher, medical sciences, Belgium
	"That we push our researchers to publish, publish, publish and that they get the salary, let's say, that pays on the number of publications that they publish every year… This is, it is enormously important to do something about that in Europe, especially the young PhDs and Postdocs who are really suffering from short-term contract and short-term contract, and trying to get a scientific job. They are really under pressure to perform and to give their best results. So, when you come in the grey-area they are stretching the grey-area maybe. And they are very; it is very easy for them, not easy for them but it is easy to push them across the border because of this pressure, and I think that is mainly something; we should try to do something in Europe. Because, the system is killing the whole community I guess." – Funder, natural sciences, Netherlands
	"<u>Senior researcher</u>: No, it [i.e. competitiveness] is not a barrier for training. <u>Funder</u>: It's not a barrier for training. <u>Senior researcher</u>: It's even, even more […] <u>Funder</u>: You should, because this is a very competitive, so, environment, you need to train and need to get people aware, so that… <u>Secretary of RE committee</u>: And you can also use that competitiveness. So, the reason that it is competitive is because there is a reward for doing it, like that." – Natural sciences, Netherlands

Table A.7 continued

Chapter 5

5. Using co-creation methods for RI guideline development – How, what why and when?

Abstract

Existing research integrity (RI) guideline development methods are limited in including various perspectives. While co-creation methods could help to address this, there is little information available to researchers and practitioners on how, why and when to use co-creation for developing RI guidelines, nor what the outcomes of co-creation methods are. In this paper, we aim to address this gap. First, we discuss how co-creation methods can be used for RI guideline development, based on our experience of developing RI guidelines. We elaborate on steps including preparation of the aims and design; participant sensitization; organizing and facilitating workshops; and analyzing data and translating them into guidelines. Secondly, we present the resulting RI guidelines, to show what the outcome of co-creation methods are.

     Thirdly, we reflect on why and when researchers might want to use co-creation methods for developing RI guidelines. We discuss that stakeholder engagement and inclusion of diverse perspectives are key strengths of co-creation methods. We also reflect that co-creation methods have the potential to make guidelines implementable if followed by additional steps such as revision working groups. We conclude that co-creation methods are a valuable approach to creating new RI guidelines when used together with additional methods.

Keywords: co-creation, guideline development, research policy, research integrity, online research methods

1. Introduction

Research integrity (RI) is about conducting research according to the highest ethical and professional standards (Boehme et al., 2016). RI is important to ensure the trustworthiness and quality of research. RI is thought to be the responsibility of all research stakeholders, including researchers, research institutions, funders, and journals (Bouter, 2018). To help research stakeholders address their RI responsibilities, in the past years various guidance documents on RI (e.g. All European Academies, 2017) have been produced. Guidelines show research stakeholders how to conduct, organize, support or regulate research practices. Guidelines can take various forms, such as providing a checklist, a list of recommendations, or best practices.

     Guidelines are developed using a combination of evidence-based methods and expert discussion (Brouwers et al., 2010; McAlister et al., 2007; Qaseem et al., 2010; Trepanier et al., 2021). There are various ways to organize expert discussion, with the most common being informal (Fretheim et al., 2006; Murphy et al., 1998), and formal consensus approaches (James & Warren-Forward, 2015). Informal consensus approaches entail the use of working groups or panel discussions (Fretheim et al., 2006; Murphy et al., 1998). In contrast, formal consensus approaches, such as Delphi

studies and the nominal group technique, use structured means of reaching consensus (James & Warren-Forward, 2015).

As Fretheim and colleagues explain (2006), informal consensus methods are not ideal because they lack transparency, and are prone to undesirable group dynamics such as certain voices dominating and biasing the discussion. Despite these limitations, many research integrity guidance-providing documents are currently based on such methods (ALLEA 2017; ENERI 2019; NESH 2019). James and Warren-Forward (2015) explain that some formal consensus approaches, such as Delphi studies, are transparent and use various strategies to reduce the influence of group dynamics on decision making. However, they also have limitations. Nie and colleagues (2020) argue that the focus on reaching agreements among the group majority might lead to missing out on diverging views. Yet, such views could be particularly vital for developing guidelines that are sensitive to the specific needs of diverse users across countries and disciplinary fields.

In the past two decades, new methods for engaging participants have been developed in the field of industrial design, often referred to as 'co-creation methods'. Sleeswijk Visser and colleagues (2005) describe a form of co-creation methods used to engage a wide range of expert and non-experts to express and reflect on earlier experiences. This is a generative design research approach to co-creation, which engages stakeholders not as research subjects, but as partners who are 'experts of their experience'; promotes out-of-the-box thinking; and makes stakeholders' tacit values explicit (Sanders and Stappers, 2012). Additionally, this form of co-creation steps beyond 'stakeholder consultation' as criticized by Arnstein (2020), in that stakeholders' views are not merely considered as research 'data' to take into account. Van Woezik and colleagues (2016) explain that co-creation methods are especially valuable for dealing with complex problems. These are problems where multiple intertwining factors and stakeholders are involved and there are no easy solutions, hence requiring flexible approaches which take into account various perspectives. Considering that guidelines often deal with complex problems involving multiple stakeholders and relevant factors, it could be that co-creation methods are helpful for guideline development. As co-creation methods are increasingly being expanded from the fields of industrial design and marketing to fields in the social sciences (Brandsen et al., 2018; Langley et al., 2018), there is a rise in the availability of public co-creation tools and resources (Foster Open Science, n.d.; GoNano, n.d.; SISCODE, n.d.). This is valuable for researchers interested in developing guidelines.

However, there is a gap in the literature regarding how researchers can use these tools and resources to specifically design RI guidelines using co-creation methods. Furthermore, experience-based information on the value of using co-creation methods for RI guideline development is also lacking. In this paper we aim to expand on the knowledge base regarding using co-creation methods for guideline development, by sharing insight with other researchers and practitioners about using co-creation to develop RI guidelines, based on our experiences with developing RI guidelines using online co-creation methods. First, we provide some reflections on how co-creation methods can be used to develop guidelines, using insights gained from our own experience of using co-creation methods to develop guidelines targeted at research institutions and funders on how to foster RI. Secondly, we present the

resulting guidelines to show what the outcome of co-creation methods are. Thirdly, we elaborate on why and when co-creation methods can be used for RI guideline development, based on our own experiences and insights as researchers, as well as by sharing the perspectives and insights from participants included in our guideline development process

2. How to use co-creation methods for RI guideline development?

In our guideline development process, we aimed to develop guidelines for research institutions and funders across Europe on RI topics not currently addressed by high quality publicly available existing documents, together with lead users using co-creation methods (Labib et al., 2020; Lechner et al., 2020). Our intention was to create guidelines which addressed the responsibilities of institutions and funders at the organizational level, namely the level of rectors, deans, directors, RI officers, policy staff, and advisors. The guidelines that we intended to develop for research institutions addressed the topics: 1) RI education and training, 2) building a responsible research environment, and 3) fostering responsible supervision; while the guideline targeted at funders focused on the topics: 4) safeguarding the independence of funded research, 5) selecting and evaluating proposals responsibly, and 6) monitoring funded projects. Prior to developing the guidelines, we had conducted several studies to explore the gaps and lacunas of current practices as institutions and funders (Gaskell et al., 2019; Labib, Evans, et al., 2021; Labib, Roje, et al., 2021; Mejlgaard et al., 2020; Sørensen et al., 2021). We did not pre-specify the format of the guidelines before the workshops, as we intended to address this issue during the co-creation process, and have participants decide on the most appropriate format.

There are various approaches to using co-creation methods – methods which engage users in interactive exercises involving role-playing, story-telling, card games, drawing, and other techniques promoting creativity (De Couvreur & Goossens, 2011; Lee et al., 2018; Sanders & Stappers, 2008). These include approaches focused on the development of user-centered products and services in the commercial sector (e.g. the development of shaving products Sleeswijk Visser et al., 2005); in addressing public sector questions which require novel ideas (for instance on how to become a better elementary school teacher, or how to create better healthcare services, (e.g. Sanders & Stappers, 2012 pp. 88-89, 106-111); and to engage members of the general public in matters of research and innovation (GoNano, n.d.; Robinson et al., 2020; e.g. SISCODE, n.d.). However, none of these approaches were fully adequate for the purpose of RI guideline development, because the tools created in other contexts – for instance, exercise toolkits, sensitization materials, and card games – were not aimed at creating concrete guidance documents. To meet our needs, we needed to develop our own approach to co-creation methods, which allowed for developing and discussing RI guidelines usable by research institutions and funders across Europe, and ensured the appropriateness of all tools for our specific target group of research stakeholders.

Due to the COVID-19 pandemic, it was not physically possible to organize workshops where research stakeholders from different countries in Europe could come together in real life to work on the guidelines. Therefore, we found it most convenient to organize the workshops in an online environment, as this allowed for the inclusion of

participants from countries across Europe. At the time, there was less published literature about using online methods for co-creation than is available now two years into the pandemic (e.g. Dexter et al., 2013). We worked together with two co-creation experts (PJS and KB) – one a professor in Design, and the other a researcher and professional facilitator – to combine their methodology expertise with the rest of the team's topic expertise to design the RI guideline co-creation methods .

Steps to co-creating RI guidelines

Step 1: Preparation

In our experience, using co-creation methods involves extensive preparatory work (Sanders & Stappers, 2012), which should not be underestimated. Preparation involves the following steps which are elaborated further below: a) setting clear aims, b) designing the method, as well as c) selecting a suitable recruitment strategy for finding participants.

a) Aims

Researchers have the option to choose between a more exploratory aim (such as reflecting on how supervision can be improved) and a more concrete outcome oriented aim (e.g. a guideline on supervision for research institutions) (Bhalla, 2016; Ida, 2017; Liu et al., 2018; Nambisan & Nambisan,). The former could be helpful in allowing participants to openly explore the general problem at hand, and jointly agree on an outcome based on this initial exploration (Nambisan & Nambisan, 2013). While this approach is more participatory, it requires sufficient workshop time for exploring the problem at hand and jointly constructing a project aim. As such, it may be most suitable in cases where there is a lack of available literature on the problem at hand. Alternatively, an outcome oriented aim (particularly if not based on prior research with stakeholders), is at risk of not sufficiently taking into account stakeholders' actual needs and preferences, but can be much more efficient and easier to work with (Nambisan & Nambisan, 2013).

We decided to go with the latter option since we had decided on creating our RI guidelines – i.e. the intended outcomes of the co-creation process– based on an earlier extensive deliberation process supported by multiple empirical steps, in which we already consulted with various stakeholders (Labib et al., 2020; Labib, Roje, et al., 2021; Lechner et al., 2020). In addition to setting a concrete product-oriented aim (i.e. to create RI guidelines on the pre-specified topics), we also set two additional exploratory aims: investigate which guideline formats participants prefer, as well as delve into potential implementation issues of the guidelines. This allowed us to not only make steps towards producing the guidelines, but also helped us to look forward to how the guidelines might be implemented in practice.

b) Methods

One of the dominant approaches to using co-creation methods is the 'Double Diamond' (Figure 1A), referring to a 4 step process to co-creation including: 1) discovering new

ideas and opportunities, 2) defining a creation strategy by filtering, selecting and discarding ideas, 3) developing the ideas chosen in step 2 into a product, and 4) delivering a product, including testing and launching (Design Council, 2015, 2021). As such, the 'Double Diamond' approach consists of alternating divergent and convergent stages, where participants first go through a process of opening up and creating many ideas, and then closing in by filtering and selecting ideas (Design Council, 2015, 2021; Stelzle et al., 2017). This ensures that a broad range of options is considered in the (product or service) development process, promoting the inclusion of novel and innovative ideas, as well as safeguarding that of all potential options, only the highest quality ideas (as defined by participants) are included in the final product or, in our case, guidelines (Stelzle et al., 2017).

In our work, we adapted the Double Diamond model to design our guideline development process (Figure 1B). We decided to organize four workshops for each of the six topics we aimed to create guidelines for; therefore we conducted twenty-four workshops in total. The first two of the workshops per topic were focused on creating content for the guidelines, whereas the last two workshops were focused on refining the guideline content. Within each workshop, there was a divergent and convergent step where participants first had to develop a wide range of ideas, and then select and prioritize some ideas over others. The first two workshops focused on the first diamond, while the second two workshops focused on the second diamond. In addition to preparing and facilitating the workshops, the researchers' role was to draft the first version of the guidelines after the content creation workshops, and revise the guidelines after the content refinement workshops.

c) Recruitment strategy

Participant selection and recruitment is similar in co-creation methods compared to other qualitative research methods (Sanders & Stappers, 2012). One notable difference is that it is typical for workshops using co-creation methods to include fewer participants – two to six people – than other methods, such as focus groups (Sanders & Stappers, 2012). This is because close and intensive collaborations between a small group is necessary to allow sufficient room and time to discuss diverse ideas and to come to conclusions (Sanders & Stappers, 2012). Due to challenges in keeping everyone engaged and active in the online environment, our experience is that it is even more important in the online setting to include only a small number of participants per workshop.
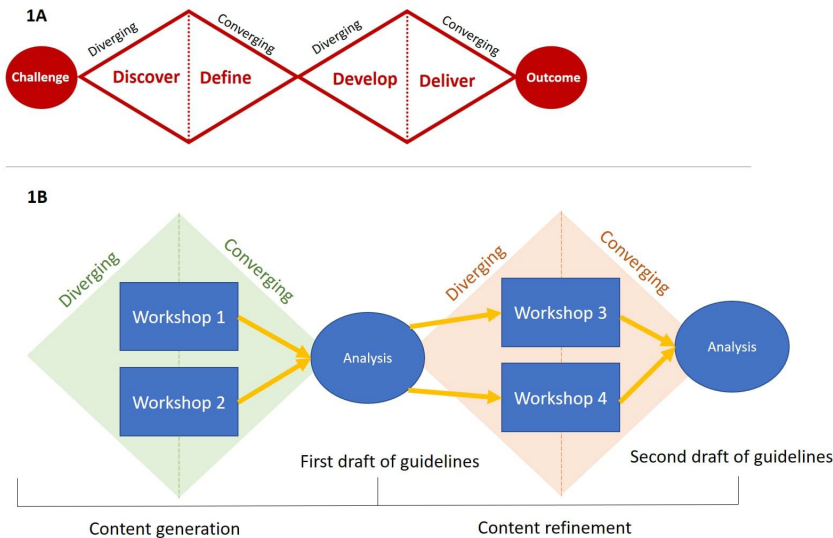
Figure 1 Co-creation process of diverging and converging. Figure 1A visualizes the Double Diamond design approach shown by the Design Council. Figure 1B shows our adaptation of the Double Diamond design approach. Each of our workshops in the content generation set (i.e. workshops 1 and 2)) consisted of a divergent phase where participants developed a wide range of ideas for the guidelines, and a convergent phase where a selection of ideas was made and prioritized. Following the content generation step, the researchers analyzed the data and compiled the first version of the guidelines. Each of the workshops in the subsequent content refinement set (i.e. workshops 3 and 4) consisted of a divergent phase where participants evaluated the guidelines, and a convergent phase where participants came to agreements about refinements needed in the guidelines. Following the content refinement workshops, the researchers analyzed the data and refined and finalized the guidelines.

Taking these considerations into account, we aimed to recruit four participants per workshop, so as to keep the group small but still allow for input from diverse perspectives. We identified and subsequently invited participants who would be future lead users of the guidelines (i.e. RI officers, educators, researchers, funders, policy makers, administrators, etc. from various parts of Europe) using our networks, as well as through snowballing. We aimed to include participants with diversity in country, gender, and position. To allow for some continuity across workshops, we included one to two participants in both the 'content creation' and 'content refinement' workshops for that RI guideline topic. This led to the inclusion of seventy-five participants in total across our twenty four workshops (i.e., six RI guideline topics, with four workshops for each topic), with twenty one participants taking part in both a 'content creation' and 'content refinement' workshop; for more details, please see Pizzolato et al. (2021). We had two to seven participants per workshop.

To familiarize participants with the online tools used for the workshop, we organized a fifteen-minute one-on-one call with each participant prior to the workshop to test the online tools and practice using them. This was to minimize potential problems that might arise during the workshops due to technical issues, and to

therefore safeguard the quality of the workshops. Participants' familiarity with online tools and tech-savviness can influence the quality of workshop collaborations (Fuglerud et al., 2021; Wallgren et al., 2021).

Step 2: Sensitization

As Sanders and Stappers (2012) explain, 'creativity', a key element in the divergent phase of workshop using co-creation methods, does not happen instantaneously. Instead, creativity is a process requiring sufficient preparation in terms of priming and activation to ensure that individuals can generate a wide range of ideas, link initially separate ideas into new combinations, and make associations between interconnected information (Sanders & Stappers, 2012). Because of this, it is standard practice before a workshop using co-creation methods to 'sensitize' participants, i.e. give participants some tasks to complete in preparation of the workshop (Sleeswijk Visser et al. 2005). The general advice regarding sensitization is to provide participants with a task which engages them to think about concepts related to the workshop aim, without necessarily specifying the exact aims of the workshop (Sanders & Stappers, 2012; Sleeswijk Visser et al., 2005). This is considered important to not restrict participants' thinking process, allowing for 'out-of-the box' ideas to be formed before the workshop (Sleeswijk Visser et al. 2005). Researchers can expect that many participants will not complete all sensitization exercises before the workshop. Nonetheless, mere exposure to the sensitization exercise instructions can be helpful in the – conscious or unconscious – activation of ideas and priming among participants.

Especially presuming that the workshop participants, i.e. research stakeholders, might be more accustomed to more 'analytic' types of workshops (which are focused more on critical thinking or convergent processes) than 'creative' workshops (which are focused on opening up to different ideas), we found it important to carefully design simultaneously stimulating and serious sensitization exercises and materials that would foster creativity in our workshops. For the 'content creation workshops', we designed 'inspirations' – small pieces of text or visual depictions of ideas related to the workshop topic –, which we circulated to participants one week prior to the workshop (https://osf.io/8cs42/). We asked participants to browse through the inspirations, select three which they found most striking, and explain why they found them striking (please see https://osf.io/6sqau/ for more details). This was in order to 'sensitize' them before a workshop. The 'inspirations' were intentionally designed to be 'ambiguous', or allow room for different interpretations, as ambiguity is considered a valuable tool in co-creation methods for nurturing richer discussions (Gaver, Beaver, and Benford 2003). For instance, for our workshops focusing on guidelines for RI education, one of our 'inspirations' was a picture of a devil and angel heart (Figure 2). This could be interpreted in various – potentially conflicting – ways such as: 1) that RI education helps researchers become good in research, and prevent them from being bad, 2) that researchers have both good and bad tendencies, and 3) that RI trainings treat research in black and white terms of good and bad, rather than seeing the nuances involved in doing research. We piloted the exercise with colleagues before sending them to our participants, as suggested by Sleeswijk Visser et al. (2005) to check whether they work as expected.

Figure 2 *Example inspiration sent to participants in the RI education workshops.*

<u>Step 3: Workshop exercises</u>

Because workshops using co-creation methods are focused on creating something with a group (Galvagno and Dalli 2014), the organization and facilitation of such workshops requires some specific considerations to optimize development of outcomes and foster creativity. Virtual collaborative software programs such as MIRO (MIRO 2021) or MURAL (MURAL n.d.) allow for real-time interaction between participants and have many useful built in tools, such as sticky notes (Busse and Kleiber 2020; Kaur, Kaur, and Blomkamp 2021). In our workshops, we used Zoom (Zoom Video Communications 2021) to connect with our participants, and MIRO (MIRO 2021) to interact and create our guidelines collaboratively.

To allow sufficient time to meet our workshop objectives, and yet account for the limitations involved in doing online work (e.g. becoming fatigued more easily and finding it more difficult to concentrate), the duration of each of our co-creation workshops was 3-3.5 hours long. To ensure that the workshop addressed the challenge at hand, and led to a concrete outcome, our workshops were broken down into a number of smaller exercises, with each exercise building on the previous one (as discussed in Sanders & Stappers, 2012; Sleeswijk Visser et al., 2005). More specifically, each of our workshops contained four to five exercises, ranging between 15-45 minutes in duration. Although we found various toolboxes providing standard exercises that can be used in a co-creation workshop (e.g. Hyper Island, n.d.; Skalska, 2017), we found it difficult to use existing exercises as most were not suitable for our aims. Instead, we opted to look at available toolboxes as inspiration for designing our own exercises. An example of an exercise used in our workshops can be found in Figure 3.

Each of our workshop exercises' objective contributed to the overall workshop, and thereby also the overall co-creation aims. For instance, the aim of the workshop from Figure 3 was to create content for guidelines on RI education, while the specific goal of the exercise shown was to discuss the content for RI education specifically targeted at the level of bachelor, master and PhD students. To stimulate creativity, but also safeguard the final outcome being produced (Stelzle, Jannack, and Noennig 2017), each exercise was composed of divergent and convergent elements. Individual elements – such as steps 1, 3 and 5 shown in Figure 3 – were helpful for generating initial ideas, while group exercises – such as steps 2, 4, and 6 – were particularly valuable for creating more and better ideas, through building on individuals' ideas through recombination, transformation and merging (Chung 2018). While exercises were outcome oriented, we also asked participants to explain why they selected

certain ideas or made certain choices, which allowed for a deeper understanding of stakeholders' tacit needs and values (e.g. steps 2 and 4 in Figure 3).
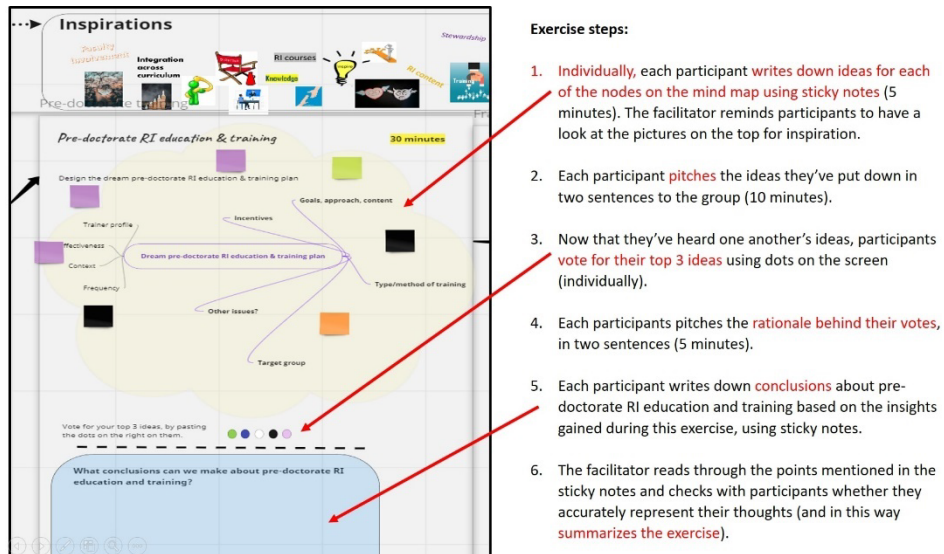
Figure 3 *Example exercise used in one of our co-creation workshops. This exercise focused on RI education for bachelor, master and PhD students, and was part of the content creation workshops for RI education and training.*

Using MIRO allowed us to visualize all our exercises before the start of the workshops and place all our materials on one virtual board. We asked participants to use pictures of the inspirations used in the sensitization exercise to think of ideas, sticky notes to write down ideas, and dots to vote on ideas selected in convergent exercise steps. In the content creation workshops, the exercises focused on generating and selecting various ideas for the guidelines at hand, as well as is in discussing the preferred guideline format (see https://osf.io/8x3b2/ for examples). Alternatively, the content refinement workshops provided participants with the opportunity to comment on any gaps, inconsistencies, discrepancies, disagreements, or other issues in the drafted guidelines available, as well as to reflect on potential implementation concerns (see https://osf.io/kx8dj/ for examples). Prior to each set of workshops, we piloted different workshop exercises with colleagues to check that they would work as expected.

In each workshop, there was at least one facilitator and one co-facilitator present in the video call. The facilitator was responsible for moderating the session, whereas the co-facilitator helped with technical issues and any other problems during the session. In workshops in which more participants joined than initially expected (i.e more than five), there was also a second co-facilitator who assisted the co-facilitator. Having at least one co-facilitator in the session was crucial for the success of the online workshops, since co-facilitators could help participants struggling with the online tools. Because of the structured nature of the workshops, we wrote detailed facilitator instructions to ensure that the facilitators were well prepared (Appendix 1). However,

since it is not possible to predict exactly how workshops will proceed (Pointon 2018), facilitators were also instructed to be flexible and adapt the workshop program when necessary – in consultation with our co-creation expert (KB) – without compromising on the workshop objectives.
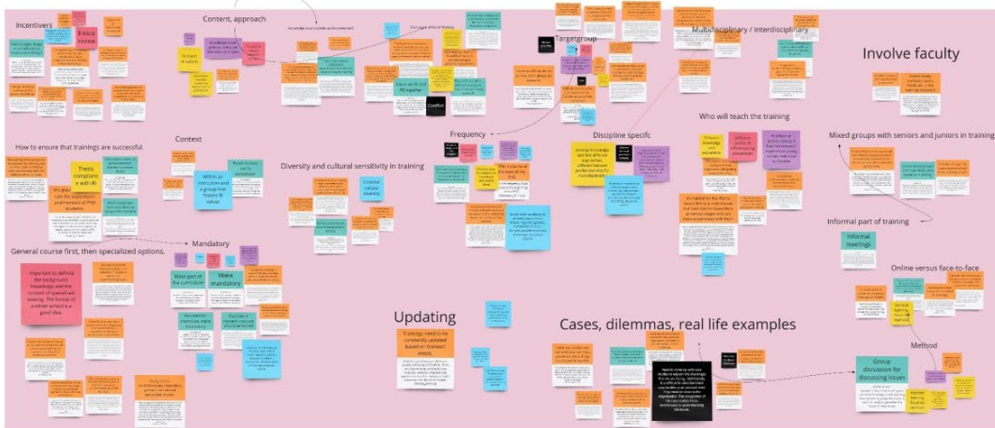
Step 4: Analysis

As is common with qualitative research, co-creation methods generate a substantial amount of data (Sanders & Stappers, 2012; Sleeswijk Visser et al., 2005). This consists of not only the workshop transcripts, but also the actual products of the workshop (e.g. the ideas on the MIRO board). As explained by Sanders and Stappers (2012), researchers have three options regarding how in-depth they conduct their analysis: i) 'inspiration only' (i.e. immersion in data without rigorous analysis), ii) traditional 'database' (i.e. line by line coding using software), and iii) 'analysis on the wall' (i.e. clustering data on a wall). Although it might seem that a traditional database approach is ideal as it is most rigorous, as explained by Sleeswijk Visser (2005) this approach "does not offer an inspiring and flexible workspace for analyzing fragmentary information about context of product use….and [does] not encourage the team to view data with empathy". Therefore, we used an 'analysis on the wall' approach, where data is clustered into groups on a real-life or virtual 'wall', rather than on a database, (Sanders & Stappers, 2012). The rationale for this decision was twofold: this approach prevents (particularly the visual) data from becoming hidden in a database, and is time-efficient.
 We used inductive and deductive thematic analysis to analyze the results from the 'content creation' and 'content refinement' workshops, respectively. The deductive themes used for the later analyses were based on the main guideline items produced earlier in the guideline development process. After analyzing the data in a small group (of 2-5 researchers per workshop topic), we visualized the results in analysis posters (example shown in Figure 4). This was to keep the analysis results close to the situations discussed by participants during the workshops. When developing the guidelines, we looked at the results of the analysis posters to write and revise each guideline items. We formulated the guidelines as a list of recommendations, each based on the results of the analysis (Pizzolato et al. 2021). An example of the second version of the guidelines – which was the direct output of the workshops – can be found in Appendix 3, while the most updated versions are available on the Open Science Framework (e.g. https://osf.io/z7m3v/).
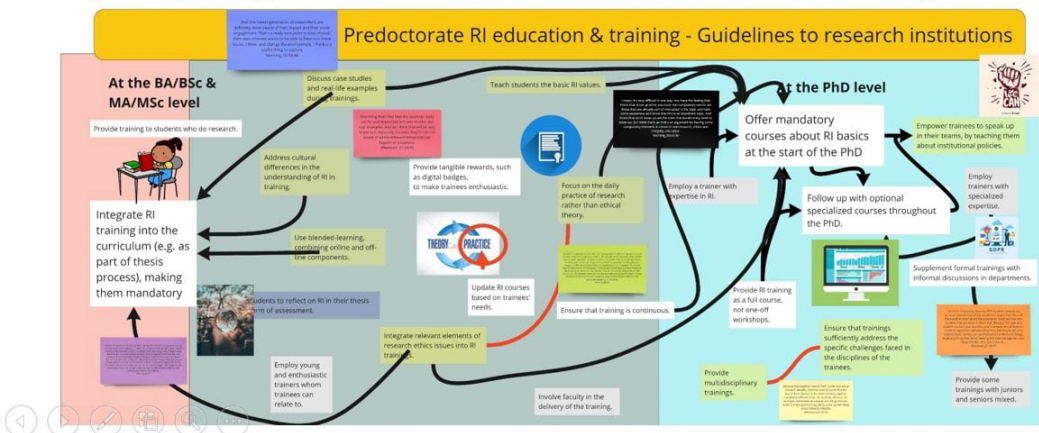
Figure 4 *Example overview of the analysis process used in the SOPs4RI co-creation process. Step 1 shows an overview of how the data were clustered by the group in an 'analysis on the wall' approach. Step 2 shows the resulting poster, based on which a first draft of a guideline on RI education for bachelor, master and PhD students was developed.*

3. What do co-creation methods result in when applied to RI guideline development?

Through the use of co-creation methods, we developed RI guidelines for research institutions on 1) RI education and training, 2) responsible supervision and leadership, 3) research environment; as well as RI guidelines for research funders on 4) selecting and evaluating proposals, 5) preventing unjustifiable interferences, and 6) monitoring funded projects. The guidelines take into account the diversity of stakeholders involved

in the research process (such as the needs of researchers across ranks, as well as differences among institutions and funders). They provide recommendations that are practical, and yet can be tailored to different research contexts. Furthermore, they are built on the consensus of the co-creators involved in developing them. Table 1 provides a breakdown of the guidelines created per topic. In this section, we provide an overview of the content of the guidelines for each topic to help readers understand *what* the co-creation methods we applied in RI guideline development result in. An example of the second version of the guidelines – which was the direct output of the workshops – can be found in Appendix 3, while the most updated versions are available on the Open Science Framework (e.g. https://osf.io/z7m3v/). We delve deeper into the specific guideline content for each topic in separate manuscripts (e.g. Pizzolato et al., 2022).

The guidelines for research institutions

RI education and training

In line with recommendations in the literature (e.g. Fanelli, 2019), the co-creation workshop participants considered it important that RI education and training is provided to all research stakeholders, including students, junior and senior researchers, as well as others involved in the research endeavour (e.g. ombduspersons, research managers, RI officers, policy staff). During the co-creation workshops, the participants discussed specific recommendations that would be appropriate for the RI education of different stakeholders, based on their own experiences with RI education. For instance, they suggested full RI courses for PhD students, small workshops for more senior researchers, and peer-to-peer learning events for other RI stakeholders. Furthermore, they emphasized the importance of approaches to learning about RI that fall outside the scope of 'formal training' in the classroom. This allowed us to create guidelines on RI education and training that capture various approaches to learning about RI, and are specific to the needs of various stakeholders. An infographic of the finalized guidelines can be found on OSF: https://osf.io/6zbqc.

Responsible supervision and leadership

The co-creation workshop participants highlighted research institutions' responsibilities regarding communicating the responsibilities of and requirements for good supervisors and leaders, but also emphasized the need for research institutions to provide adequate support and training to supervisors and leaders to achieve these. In this way, the resulting guidelines went beyond outlining what responsible supervision and leadership mean, but actually focused on how institutions can empower supervisors, PhD students and research leaders (e.g. principle investigators) to ensure responsible supervision practices. This included recommendations on providing structures for peer-to-peer support, paying sufficient attention to researchers' well-being, and providing bodies to consult in cases of conflict. The finalized guidelines have been visualized on this infographic which can be found on OSF: https://osf.io/8n5ud.

Table 1: Breakdown of RI guidelines developing per topic using co-creation methods

Topic	# of guidelines	Guideline content		Target of guidelines
RI education and training	4	-	RI training of bachelor, master and PhD students	Research institutions
		-	RI training of post-doctorate and senior researchers	
		-	RI training of other RI stakeholders	
		-	Continuous RI education	
Responsible supervision and leadership	3	-	Supervision	Research institutions
		-	PhD guidelines	
		-	Leadership	
Research environment	4	-	Culture building	Research institutions
		-	Adequate education and skill straining	
		-	Managing competition and publication pressure	
		-	Diversity and inclusion	
Selecting and evaluating proposals	3*	-	RI plan	Research funders
		-	Methodological requirements	
		-	Diversity and inclusion	
Preventing unjustifiable interferences	4*	-	What counts as an unjustifiable interference?	Research funders
		-	Interference by funders	
		-	Interference by commercial influences	
		-	Interference by political/other external influences	
Monitoring of funded projects	3*	-	Execution of the research grant	Research funders
		-	RI requirements	
		-	Financial monitoring	

This is the number of guidelines created as a result of the co-creation workshops for this topic. However, after the co-creation workshops, at a guideline revision step, it was decided that these guidelines would all be merged into one longer guideline to cover the entire topic.

Research environment

The guidelines on building a responsible research environment provide practical steps that institutions can take to addressing this key, but less tangible, issue for RI (as seen by participants in another study we conducted, Labib, Roje, et al., 2021). One of the guidelines focused on the general question of how to create community building for a responsible research environment, and this guideline addressed various areas that institutions need to address such as conducting responsible research assessments; creating an open, safe, diverse and inclusive research culture; providing researchers

with appropriate support structures and training; and dealing with competition and publication pressure. While the guideline was broad, participants provided concrete in-practice examples that could serve as inspiration for institutions on implementing these points. One such example was to publish institutional staff survey results, including negative comments, so as to create more transparency in the institution. The other three guidelines under the topic of responsible research environment, provided more detailed recommendations on specific aspects that were already highlighted under the general guideline on community building. The guideline focused on diversity and inclusion, for instance, emphasized the importance of taking an intersectional approach to diversity that accounts for different types of diversity (e.g. race, gender, class) and how they intersect, and provided concrete suggestions on how to do this (for instance by including diverse researchers in a bottom up way when developing diversity policies in the institution). A more detailed overview of the guidelines can be found on the OSF: https://osf.io/jcpgq.

Guidelines for research funders

The guidelines for research funders have been visualized on this infographic, available on OSF: https://osf.io/q2wra.

Selecting and evaluating proposals

When creating the guidelines for this topic, participants emphasized the difficulty to standardize detailed recommendations given the large diversity in funders in terms of size, funding streams (i.e. governmental or private), and culture. However, they agreed on some basics that apply across funders, such as requiring proposals to include a plan on how to safeguard RI; paying sufficient attention to the methodology section of submitted proposals; and removing biases from the selection and evaluation process (e.g. by ensuring that the language used to communicate to grant applicants is inclusive).

Preventing unjustifiable interferences

Co-creation workshop participants highlighted that not all interferences in the research process are unjustified and came to agreement about which are and are not justified. For instance, they concluded that the funder can influence the research agenda, but that interference in the publication of results is unjustified. In addition to recommending having a clear definition of 'unjustifiable interferences' for each research funder, the guidelines on this topic also provide recommendations on how to prevent such interference and deal with it. These recommendations address various phases of the research and funding process, from the moment of selecting and evaluating proposals (in which preventing conflicts of interests is considered crucial), to providing guidelines about projects co-funded by commercial parties, and addressing how to keep researchers independent in the publication step of the research process.

<u>*Monitoring of funded projects*</u>

Co-creation workshop participants emphasized that while having funders monitor funded projects is crucial as a way to increase the trustworthiness of research, this monitoring process should be collaborative and cooperative (rather than employ a policing approach). While participants provided various recommendations on how to monitor projects in a manner that is fruitful and minimally bureaucratic (e.g. providing a checklist of points that researchers can report on to the funder), they also emphasized the need to create a quality assurance system for the monitoring process. Therefore the monitoring guidelines provide both 'what to monitor' recommendations for funders, as well as recommendations on 'how to' ensure that this is done in a cooperative and productive manner.

4. Why and when to use co-creation methods for guideline development?

Since the use of co-creation methods for RI guideline development is a novel approach, we frequently interrogated the advantages and disadvantages of co-creation methods and their suitability for RI guideline creation. In this section, we reflect on why and when it is suitable to use co-creation methods for developing guidelines. For this, we are combining our own experiences with those of our participants. To learn about our participants' reflections, we conducted a set of informal interviews with one participant from each of our workshops, with the aim to explore how participants evaluated co-creation methods for developing RI guidelines. More details about the interviews, including the interview guide, characteristics of interviewees (including their interviewee numbers, demarcated as 'IN'), and interview procedures can be found in Appendix 2.

Why use co-creation methods for RI guideline development?

<u>*Stakeholder engagement*</u>

Our interviewees identified close stakeholder engagement as a key benefit to co-creation methods. They indicated that the stimulating exercises of our workshops kept them closely engaged and willing to contribute to the RI guidelines. Interviewees expressed that our workshops were: "interactive and colorful, and not boring…" (IN 3), "quite remarkable" (IN 6), "insightful and interesting" and "innovative" (IN 7), "fun" and "rewarding" (IN 8). One of our interviewees mentioned that when "you have [people] on board [and engaged], they will implement [the guidelines] because they like it and because they contributed to it" (interviewee number, IN 18). Another participant remarked that engaging stakeholders actively is especially necessary to get "buy-in", i.e. to ensure that all stakeholders support the guidelines (IN, 14). Others mirrored these views by stating that stakeholder engagement is especially valuable to increase the likelihood that the guidelines will be actively used by institutions after the development process is complete. There were also some participants who appeared less comfortable with the workshops, more specifically to the online setting of the

workshops, with a few remarking that MIRO was difficult to use. Others expressed that while their initial reaction to the workshop invitation was hesitant due to the online setting, they were then surprised to experience the online workshops as interesting and productive.

These results are promising, especially because we were initially concerned – also due to challenges outlined in Deserti, Rizzo, and Smallman (2020) – that it might be difficult to stimulate research stakeholders such as policy makers to embrace the creative aspects of co-creation methods (e.g. work with images to create a broad range of new ideas and alternative perspectives). Features of co-creation methods that make it particularly suitable for engaging stakeholders involve a) giving stakeholders the opportunity to create outputs based on their own needs, which they can then use themselves, and b) using interactive and playful exercises, by design (Sanders & Stappers, 2012). This level of stakeholder engagement in our co-creation process is not fully surprising given that co-creation methods have also been previously used successfully with other groups of analytical, or 'critical', participants, including CEOs, healthcare workers, and policy makers (e.g. Agrawal, Kaushik, & Rahman, 2015; Fuster Morell & Senabre Hidalgo, 2020; Kimbell & Bailey, 2017; Sanders & Stappers, 2012; Waseem, Biggemann, & Garry, 2018). Even stakeholders who might initially be hesitant about participating in creative workshops can be stimulated to be creative and engage in 'serious play', using and offering the appropriate tools to evoke creativity (e.g. 'inspirations' and stimulating workshop exercises) and the reassurance that the 'play' will lead to productive and valuable outputs for the participants (Hinthorne & Schneider, 2012; Sanders & Stappers, 2012).

We encountered a few difficulties with ensuring that all participants were comfortable online. However, we experienced the use of a whiteboard interactive software like MIRO to facilitate engaging workshop sessions virtually positively, since it allowed participants from different parts of the world – time differences allowing – to collaborate together on one platform in real time, using various creative tools (Busse & Kleiber, 2020; Kaur, Kaur, & Blomkamp, 2021). Since MIRO requires some familiarization before it can be used optimally, we found it helpful to organize one-on-one calls with each participant ahead of time to explain the tool and help them practice with it. It might be, however, that participation in more than one workshop is necessary for all participants to feel comfortable with tools like MIRO, and that simpler tools are needed when engaging one-time participants who are not very tech-savvy (e.g. Google, n.d.) . However, this has to be weighed against the inconveniences that simpler tools might present, such as limitations for facilitators and not providing all technical options needed for a workshop.

Inclusion of diverse perspectives

Co-creation methods were also considered valuable by our interviewees in stimulating the inclusion of a broad range of perspectives in the RI guidelines. As put by one of our interviewees, discussing diverse views is important "to prevent skipping some steps and starting with a one-sided perspective" (IN 1). Another interviewee also appreciated that "co-creation is not only meant to see what is mainstream but what are possibly dissenting views", since when it comes to RI guidelines, "minority views are as

important as majority views", as they are likely "more problematic, more novel, and innovative" (IN 6). Our workshop participants appreciated that co-creation methods actively encourage participants to share their diverse and unique perspectives in various ways, including starting workshops by first encouraging participants to be open to various ideas – and even be stimulated to "dream" about an ideal guideline (IN 8) – and only afterwards select ideas which are more practical and feasible; giving all participants "time to reflect on an issue" individually and then stimulating them to share these reflections in discussion with the group (IN 22); the framing of questions in ways that lead to different interpretations; combining visual and textual elements during the workshops (as people "respond differently when [they] have cartoons or words" (IN 22)); and "focusing on real-life experiences" of the participants to "come up with bigger perspectives" (IN 9). However, our interviewees also highlighted that the diversity of perspectives that can be included in a workshop is limited to the characteristics of the workshop participants. We received some criticism that despite there being sufficient diversity among our participants in terms of gender, stakeholder type, and country in Europe, we did not include enough participants from junior ranks, countries outside Europe, and diverse cultural backgrounds in our co-creation workshops.

Indeed, compared to other guideline development methods we have previously used, such as Delphi studies, we also found co-creation methods to more actively focus on evoking a broad range of ideas. These results confirm that co-creation methods are suitable for addressing complex problems, since they actively promote the inclusion of various perspectives (van Woezik et al., 2016). However, the results also suggest that to include a broad range of perspectives, both diversity in the demographic characteristics of participants, as well as the use of techniques that evoke various perspectives among a specific group (e.g. using phrasings that lead to different interpretations), are needed. However, including diversity in all dimensions of the demographic characteristics of participants will require a larger number of participants and, hence, workshops. The decision as to whether to hold workshops online or in-person will also have an influence on the diversity of the participants included in the workshops.

When it comes to guideline development, including diverse perspectives is helpful to ensure that the guidelines are sensitive to the needs of all relevant stakeholders. For RI guidelines, this includes research stakeholders across disciplines, countries, and institutions. Of course, at the end of the guideline development process, many ideas will need to be abandoned so that only the highest quality ideas are used. Yet, starting out with an open approach and allowing diverse users to define priorities allows guidelines to address the most important needs of all stakeholders.

When use co-creation methods for developing RI guidelines?

Our advice to other researchers and practitioners is to use co-creation methods early in the guideline development process. Early use of co-creation methods allows for – as described by one participant (IN 22) – guideline "details [and nuances] that would be missed in a different setting". Additionally, early use of co-creation methods in the RI guideline development process – as we did – allows for a timely understanding of the level of agreement about the RI guideline content among stakeholders, as well as for

consensus building. Many of our participants, for instance, said that they were satisfied to see that after discussions in the workshops, many points of agreement emerged and remaining differences in opinion – although well represented in the final outputs – were small and mainly related to the specific context in which the participants worked in (e.g. country, institution type).

Co-creation methods can be followed up by additional steps in the guideline development process to ensure that guidelines are well refined. This was highlighted by our interviewees, one of whom mentioned that the guidelines are still "very general", whereas they "should be specific and… offer a way of action" (IN 17), and another who suggested that further steps are needed to write "things [i.e. the guideline items] in a clear and academic language" (IN 6).

Although the interviews were held before the interviewees had the chance to see the guidelines resulting from the workshops, these results already indicate that the workshops led to an abundance of ideas for the guidelines addressing important aspects of the workshop topic, which needed fine tuning later. This is in line with what we observed, since at the end of our guideline development process, we had a comprehensive set of guidelines for each of our 6 RI topics, which included agreements formed by participants across workshops. Furthermore, the workshops helped to elucidate differences in how institutions and funders from various countries approach RI (e.g. regarding laws, definitions, existing infrastructures and policies, etc.). In line with the interviewee's concerns, while the resulting guidelines were comprehensive, they were less 'actionable' (i.e. ready to be used). This is because our workshops produced a lot of ideas, but 1) the organization of these ideas (including merging, regrouping, and simplifying ideas) was not finished after the workshops, 2) the formulation of the ideas into concrete recommendations needed further fine-tuning (e.g. some items needed to be made general enough to be implementable across different institutions and funders in Europe). This is not surprising; given that workshop discussions using co-creation methods are focused on broader ideas even during convergent steps, there is less room for fine-tuning the details of the guideline formulations during workshops.

This suggests that guideline developers should use co-creation methods as the first major phase of the RI guideline development process – akin to the 'fuzzy front end' of design described by Sanders and Stappers (2012) – to bring diverse and out-of-the-box perspectives to the floor, and then follow up with other methods to finalize the RI guidelines (such as expert working groups, surveys, and consensus methods). In line with this view, we used additional steps after the co-creation methods to finalize our guidelines, including a small expert working group to refine the guidelines, input from additional experience and content experts, as well as piloting of the guidelines. We expect that using such a multi-stage guideline development process helped to engage diverse stakeholders closely throughout the guideline development process to incorporate diverse perspectives and safeguard the quality of our guidelines and promote their implementation. Furthermore, such an approach allows the joint development of not only the guideline content, but also the format, as well as an early exploration into potential implementation challenges and opportunities. However, we also acknowledge that this process costs substantial time and human resources, which might not always be available, particularly in smaller RI guideline development projects.

5. Conclusions

In this paper, we reflect on what we learned about co-creation methods when developing RI guidelines to share insights with other researchers and practitioners on how, why and when they can use co-creation methods for developing RI guidelines, as well as sharing what the outputs of co-creation methods can be. Regarding the 'how' question, we discuss that careful and extensive planning is required to prepare co-creation methods. This includes setting a clear and suitable research aim, designing the guideline development process using alternating diverging and converging steps, and recruiting diverse participants into a number of small and intense workshops. Additionally, we discuss the importance of sensitizing participants prior to the workshop to prepare them for creativity, as well as organizing and facilitating engaging and structured workshop exercises to stimulate productivity. Finally, we discuss that researchers have a variety of options regarding how to analyze their data in order to develop the guidelines, depending on the time available and purpose of the analysis.

Regarding the question of 'what' co-creation methods can result in, we present an overview of the resulting guidelines for the six topics we addressed using co-creation methods: 1) RI education and training, 2) responsible supervision and mentoring, 3) research environment, 4) selecting and evaluating funding proposals, 5) preventing unjustifiable interferences in the research process, and 6) monitoring of funded projects. The results show that co-creation methods help to develop RI guidelines that are sensitive to the needs of diverse RI stakeholders.

As to 'why' use co-creation for guideline development, in our view, co-creation methods are unique and valuable to the guideline development process. They are particularly helpful in terms of engaging stakeholders closely throughout the guideline development process, as well as evoking a broad range of ideas and including diverse perspectives in the guidelines. This allows for the development of guidelines that meet diverse stakeholders' actual needs.

To address the 'when' question, our experiences indicate that co-creation methods are most helpful at the early phase of the guideline development process. We would recommend guideline developers to use a multi-stage approach to co-creating guidelines; co-creation methods likely need to be followed up by additional guideline development methods (e.g. expert working groups, consensus methods, etc.) to further organize the ideas generated by co-creation, and make guidelines precise, actionable and implementable.

Supplementary materials

Appendices I-III can be found here: https://osf.io/rkap7/

Declarations

Conflict of interest

The authors have no conflicts of interest to declare.

Funding

Authors' contributions

KL drafted and revised the co-creation study protocol and design, conducted the follow-up evaluation interviews, analyzed the data, and drafted and revised the manuscript. DP, IL, and NE revised the co-creation study protocol and design, analyzed the data and revised the manuscript. PJS, KB, GW, LB, and KD revised the co-creations study protocol and design, and revised the manuscript. JT revised the co-creation study protocol and design, conducted the interviews, analyzed the data, revised the manuscript, and supervised the work. The workshops were designed by KL, in collaboration with all other authors. The workshops were overseen by KB. Facilitators included NE, JT, and KL, and co-facilitators included DP, IL, and NS and BT (please see acknowledgements).

Acknowledgements

References

Agrawal, A. K., Kaushik, A. K., & Rahman, Z. (2015). Co-creation of social value through integration of stakeholders. Procedia-Social and Behavioral Sciences, 189, 442–448.

All European Academies. (2017). The European Code of Conduct for Research Integrity. All European Academies. https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

ALLEA. (2017). European Code of Conduct for Research Integrity.

Arnstein, S. (2020). "A Ladder of Citizen Participation": Journal of the American Institute of Planners (1969). In The City Reader (pp. 290–302). Routledge.

Bhalla, G. (2016). "Collaboration and Co-Creation: The Road to Creating Value." The Marketing Journal.

Boehme, O., Föger, N., Hiney, M., Peatfiled, T., & Petiet, F. (2016). Research Integrity Practices in Science Europe Member Organisations. https://doi.org/10.5281/zenodo.5060196

Bouter, L. (2018). Fostering responsible research practices is a shared responsibility of multiple stakeholders. Journal of Clinical Epidemiology, 96, 143–146.

Brandsen, T., Steen, T., & Verschuere, B. (2018). Co-creation and co-production in public services: Urgent issues in practice and research. In Co-Production and Co-Creation (pp. 3–8). Routledge.

Brouwers, M. C., Kho, M. E., Browman, G. P., Burgers, J. S., Cluzeau, F., Feder, G., Fervers, B., Graham, I. D., Grimshaw, J., & Hanna, S. E. (2010). AGREE II: advancing guideline development, reporting and evaluation in health care. Cmaj, 182(18), E839–E842.

Busse, B., & Kleiber, I. (2020). Realizing an online conference: Organization, management, tools, communication, and co-creation. *International Journal of Corpus Linguistics, 25*(3), 322–346.

Chung, W. (2018). Effective ideation method for collective creativity. *International Conference on Applied Human Factors and Ergonomics,* 153–161.

De Couvreur, L., & Goossens, R. (2011). Design for (every) one: co-creation as a bridge between universal design and rehabilitation engineering. *CoDesign, 7*(2), 107–121.

Deserti, A., Rizzo, F., & Smallman, M. (2020). Experimenting with co-design in STI policy making. *Policy Design and Practice, 3*(2), 1–15.

Design Council. (2015). *The design process: What is the double diamond.* https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond

Design Council. (2021). *What is the framework for innovation? Design Council's evolved Double Diamond.* https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond

Dexter, M., Atkinson, P., & Dearden, A. (2013). *Open design and medical products: irreconcilable differences, or natural bedfellows?*

ENERI. (2019). *Recommendations for the Investigation of Research Misconduct.*

Fanelli, D. (2019). *MLE on Research Integrity – Training & Education – Thematic Report No 4.* https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE RI No4 - Training and education.pdf

Foster Open Science. (n.d.). *Menu of Co-creation Tools Title.* Retrieved October 19, 2021, from https://www.fosteropenscience.eu/content/menu-co-creation-tools

Fretheim, A., Schünemann, H. J., & Oxman, A. D. (2006). Improving the use of research evidence in guideline development: 5. Group processes. *Health Research Policy and Systems, 4*(1), 1–4.

Fuglerud, K. S., Halbach, T., & Snaprud, M. (2021). Involving Diverse Users for Inclusive Technology Development. *IADIS International Conference on Interfaces and Human Computer Interaction 2021 (Part of MCCSIS 2021).*

Fuster Morell, M., & Senabre Hidalgo, E. (2020). Co-creation applied to public policy: a case study on collaborative policies for the platform economy in the city of Barcelona. *CoDesign,* 1–20. https://doi.org/10.1080/15710882.2020.1854313

Galvagno, M., & Dalli, D. (2014). Theory of value co-creation: a systematic literature review. *Managing Service Quality.*

Gaskell, G., Ščepanović, R., Buljan, R., Utrobičić, A., Marušić, A., Reyes Elizondo, A., Kaltenbrunner, W., Labib, K., & Tijdink, J. (2019). *D3.2: Scoping reviews including multi-level model of research cultures and research conduct* (A. Marušić, Ed.).

Gaver, W. W., Beaver, J., & Benford, S. (2003). Ambiguity as a resource for design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 233–240.

GoNano. (n.d.). *Involving societal stakeholders as a source of creativity in research - A co-creation toolkit for researchers and engineers.* Retrieved October 19, 2021, from http://gonano-project.eu/wp-content/uploads/2020/05/GoNano_Co-Creation-toolkit_DEF.pdf

Google. (n.d.). *Breng leren tot leven met Jamboard.* https://edu.google.com/products/jamboard/

Hinthorne, L. L., & Schneider, K. (2012). Playing with purpose: using serious play to enhance participatory development communication. *International Journal of Communication, 6,* 2801–2824.

Hyper Island. (n.d.). *Hyper Island Toolbox.* Retrieved October 27, 2021, from https://toolbox.hyperisland.com/

Ida, E. (2017). The Role of Customers' Involvement in Value Co-creation Behaviour is Value Co-creation the Source of Competitive Advantage? *Journal of Competitiveness, 9*(3), 51–66.

James, D., & Warren-Forward, H. (2015). Research methods for formal consensus development. *Nurse Researcher, 22*(3).

Kaur, S., Kaur, M., & Blomkamp, E. (2021). Identifying best practice in remote co-design. *Tensions, Paradoxes and Plurality Conference, 173*(72), 615–616.

Kimbell, L., & Bailey, J. (2017). Prototyping and the new spirit of policy-making. *CoDesign*, *13*(3), 214–226.

Labib, K., Evans, N., Roje, R., Kavouras, P., Reyes Elizondo, A., Kaltenbrunner, W., Buljan, I., Ravn, T., Widdershoven, G., Bouter, L., Charitidis, C., Sørensen, M. P., & Tijdink, J. (2021). Education and training policies for research integrity: Insights from a focus group study. *Science and Public Policy*, scab077. https://doi.org/https://doi.org/10.1093/scipol/scab077

Labib, K., Pizzolato, D., Taraj, B., Lechner, I., Claesen, N., Evans, N., Stappers, P. J., Bergema, K., Dierickx, K., & Tijdink, J. (2020). SOPs4RI co-creation workshops protocol. https://osf.io/8upmb/

Labib, K., Roje, R., Bouter, L., Widdershoven, G., Evans, N., Marušić, A., Mokkink, L., & Tijdink, J. (2021). Important topics for fostering research integrity by research performing and research funding organizations: a Delphi consensus study. *Science and Engineering Ethics*, *27*(47), 1–22. https://doi.org/10.1007/s11948-021-00322-9

Langley, J., Wolstenholme, D., & Cooke, J. (2018). 'Collective making' as knowledge mobilisation: the contribution of participatory design in the co-creation of knowledge in healthcare. *BMC Health Services Research 2018 18:1*, *18*(1), 1–10. https://doi.org/10.1186/S12913-018-3397-Y

Lechner, I., Tijdink, J., Sørensen, M. P., Ravn, T., Bendtsen, A.-K., Labib, K., Kaltenbrunner, W., Reyes Elizondo, A. E., & Kavouras, P. (2020). D4.3: Second version of SOPs and guidelines.

Lee, J.-J., Jaatinen, M., Salmi, A., Mattelmäki, T., Smeds, R., & Holopainen, M. (2018). Design choices framework for co-creation projects. *International Journal of Design*, *12*(2).

Liu, Z., Ming, X., Song, W., Qiu, S., & Qu, Y. (2018). A perspective on value co-creation-oriented framework for smart product-service system. *Procedia Cirp*, *73*, 155–160.

McAlister, F. A., Van Diepen, S., Padwal, R. S., Johnson, J. A., & Majumdar, S. R. (2007). How evidence-based are the recommendations in evidence-based guidelines? *PLoS Med*, *4*(8), e250.

Mejlgaard, N., Bouter, L., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A.-K., Charitidis, C. A., Claesen, N., Dierickx, K., Domaradzka, A., Reyes Elizondo, A., Foeger, N., Hiney, M., Kaltenbrunner, W., Labib, K., Marušić, A., Sørensen, M. P., Ravn, T., Ščepanović, R., … Veltri, G. A. (2020). Research integrity: nine ways to move from talk to walk. *Nature Publishing Group*.

MIRO. (2021). An online visual collaborative platform. https://miro.com/

MURAL. (n.d.). Free your team's imagination. Retrieved October 22, 2021, from https://start.mural.co/free-forever?utm_medium=paid-search&utm_source=adwords&utm_campaign=201101-Core_Brand&utm_adgroup=MURAL_-_Branded&utm_campaign_id=11265145092&utm_content=mural board&utm_adgroupid=109231331743&gclid=CjwKCAjwwsmLBhACEiwANq-tXLx3TrFcB

Murphy, M. K., Black, N. A., Lamping, D. L., McKee, C. M., Sanderson, C. F., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment*, *2*(3), i–88.

Nambisan, S., & Nambisan, P. (2013). Engaging citizens in co-creation in public services. *IBM Center for Business Development*.

NESH. (2019). Guidelines for Research Ethics in the Social Sciences, Humanities, Law and Theology.

Nie, R., Tian, Z., Wang, J., & Luo, H. (2020). An objective and interactive-information-based feedback mechanism for the consensus-reaching process considering a non-support degree for minority opinions. *Expert Systems*, *37*(5), e12543.

Pizzolato, D., Dierickx, K., Tijdink, J., Labib, K., Lechner, I., Bonn, N. A., Taraj, B., Claesen, N., & Gaskell, G. (2021). D4.4: Report on the Co-Creation Workshops.

Pizzolato, D., Labib, K., Skoulikaris, N., Evans, N., Roje, R., Kavouras, P., Aubert Bonn, N., Dierickx, K., & Tijdink, J. (2022). How can research institutions support responsible supervision and leadership? *Accountability in Research*. https://doi.org/https://doi.org/10.1080/08989621.2022.2112033

Pointon, M. (2018). 5 tips for facilitating your next design workshop. UX Collective. https://uxdesign.cc/5-tips-for-facilitating-design-workshops-290808a674c4

Qaseem, A., Snow, V., Owens, D. K., & Shekelle, P. (2010). The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. Annals of Internal Medicine, 153(3), 194–199.

Robinson, D. K. R., Simone, A., & Mazzonetto, M. (2020). RRI legacies: co-creation for responsible, equitable and fair innovation in Horizon Europe. Journal of Responsible Innovation, 1–8. https://doi.org/10.1080/23299460.2020.1842633

Sanders, E., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. Co-Design, 4(1), 5–18.

Sanders, E., & Stappers, P. J. (2012). Convivial toolbox: generative research for the front end of design. BIS Publishers.

SISCODE. (n.d.). SISCODE toolbox for co-creation journeys. Retrieved October 19, 2021, from https://siscodeproject.eu/wp-content/uploads/2019/09/toolkit-27092019-1.pdf

Skalska, K. (2017). Designer toolbox: Design exercises for not only design teams, part 1. https://medium.com/product-design-foodora/designer-toolbox-design-exercises-for-not-only-design-teams-part-1-f67b94508d86

Sleeswijk Visser, F., Stappers, P. J., Van der Lugt, R., & Sanders, E. (2005). Contextmapping: experiences from practice. CoDesign, 1(2), 119–149.

Sørensen, M. P., Ravn, T., Marušić, A., Elizondo, A. R., Kavouras, P., Tijdink, J. K., & Bendtsen, A.-K. (2021). Strengthening research integrity: which topic areas should organisations focus on? Humanities and Social Sciences Communications, 8(1), 1–15.

Stelzle, B., Jannack, A., & Noennig, J. R. (2017). Co-design and co-decision: Decision making on collaborative design platforms. Procedia Computer Science, 112, 2435–2444.

Trepanier, L., Reyes, A., Stamoulos, C., Beauchamp, S., Dagenais, C., Ciquier, G., & Drapeau, M. (2021). Can We Develop Evidence-Based Guidelines Without Research Expertise? Administration and Policy in Mental Health and Mental Health Services Research. https://doi.org/https://doi.org/10.1007/s10488-021-01110-0

van Woezik, A. F. G., Braakman-Jansen, L. M. A., Kulyk, O., Siemons, L., & van Gemert-Pijnen, J. E. W. C. (2016). Tackling wicked problems in infection prevention and control: a guideline for co-creation with stakeholders. Antimicrobial Resistance & Infection Control, 5(1), 1–12.

Wallgren, P., Babapour, M., & Eriksson, S. (2021). HOW TO MAKE ADVANCED ONLINE USER STUDIES MEANINGFUL. Proceedings of the Design Society, 1, 1787–1796.

Waseem, D., Biggemann, S., & Garry, T. (2018). Value co-creation: The role of actor competence. Industrial Marketing Management, 70, 5–12.

Zoom Video Communications, Inc. (2021). Video conferencing, web conferencing. https://zoom.us/

Chapter 6

Labib K, Evans N, Pizzolato D, Aubert Bonn NA, Widdershoven G, Bouter L, Konach T, Langendam M, Dierickx K, Tijdink JK. Co-creating research integrity education guidelines for research institutions.
This contents of this chapter are under review in the journal *Science and Engineering Ethics*

6. Co-creating research integrity education guidelines for research institutions

Abstract

To foster research integrity (RI), research institutions should develop a continuous RI education approach, addressing various target groups. To support institutions to achieve this, we developed RI education guidelines together with RI experts and research administrators, exploring similarities and differences in recommendations across target groups, as well as recommendations about RI education using approaches other than formal RI training. We used an iterative co-creative process. We conducted four half-day online co-creation workshops with 16 participants in total, which were informed by the RI education evidence base. In the first two workshops, participants generated ideas for guidelines' content, focusing on different target groups and various approaches to RI education. Based on this content we developed first drafts of the guidelines. Participants in the third and fourth workshop refined those drafts. We next organized a working group which further prioritized, reorganized, and optimized the content of the guidelines.

We developed four guidelines on RI education focusing on a) bachelor, master and PhD students; b) post-doctorate and senior researchers; c) other RI stakeholders; as well as d) continuous RI education. Across guidelines, we recommend mandatory RI training; follow-up refresher training; informal discussions about RI; appropriate rewards and incentives for active participation in RI education; and evaluation of RI educational events. Our work provides experience-based co-created guidance to research institutions on what to consider when developing a successful RI education strategy. Each guideline will be offered as a distinct, publicly available tool in our toolbox (https://sops4ri.eu/toolbox/) which institutions can access, adapt and implement to meet their institution-specific RI education needs.

Trial registration: https://osf.io/zej5b

Keywords: education, training, research integrity, responsible conduct of research, guidelines, institutions

Background

Research integrity (RI) can be defined as doing research according to high professional, methodological and ethical standards (Boehme et al., 2016), and is crucial for producing trustworthy research findings. Fostering RI is the joint responsibility of multiple stakeholders (Bouter, 2018) because RI is influenced by various individual, institutional, and systemic factors. These include researchers' personal character traits and ethical decision making skills (Tijdink et al., 2016), the departmental research culture (Haven et al., 2019; Joynson & Leyser, 2015), availability of responsible leadership (Pizzolato et al., 2022), and assessment criteria for funding, hiring and promotion (Aubert Bonn & Bouter, 2021; Titus et al., 2008). Since researchers and their behaviors are highly

dependent on the infrastructures, procedures, support systems, and research environments present at their institution, research institutions, in particular, play an important role in fostering RI (Mejlgaard et al., 2020).

One of the core responsibilities of research institutions is to provide RI education and training (All European Academies, 2017; Mejlgaard et al., 2020), with some countries even having legal mandates for researchers or research institutions receiving public funding, such as the US (Kalichman, 2014). RI education is thought to shape knowledge, skills, and attitudes towards RI and thereby increase awareness about responsible research practices (RRPs) and questionable research practices (QRPs) (Kalichman & Plemmons, 2007; Labib, Roje, et al., 2021), and contribute to a better research culture (Kalichman, 2014). QRPs consist of practices that do not count as outright misconduct but can hamper the quality of research, e.g. selective reporting, hypothesizing after results are known, p-hacking, or poor supervision. The terms 'education' and 'training' are often used interchangeably and there are numerous ways to define them (Masadeh, 2012). In this paper, we use the term 'RI education' to refer to all approaches used to develop researchers' cognitive and moral understanding of, and skills related to, RI. On the other hand, we say 'training' when addressing specific formal instructional events used for RI education (e.g. courses, workshops). Thus, we see RI training as an important aspect of RI education.

There is an increasing provision of RI trainings globally (Abdi, Pizzolato, et al., 2021; Evans et al., 2021; Kalichman, 2014; Mejlgaard et al., 2020), but these are typically developed without being part of a general overarching institutional RI education strategy (Kalichman, 2014; Kalichman & Plemmons, 2007), and as such there is a risk that trainings are experienced as one-off events which have little impact on participants' long-term behavior (Barnes et al., 2006). Furthermore, most existing RI educational events target PhD students, even though research shows that diverse stakeholders also see the need for targeting other students (i.e. at the bachelor and master level), researchers across ranks, as well as other institutional stakeholders involved in research such as institutional leaders and RI policy makers (Labib, Evans, et al., 2021). Because of the diverse needs of various research stakeholders (Labib, Evans, et al., 2021), it might be that different educational strategies are required for different stakeholders (e.g. students as opposed to senior researchers). Moreover, it might be that RI training is not sufficient in providing adequate RI education, considering that RI education also takes place in informal ways, such as through supervision and socialization in the research process (Labib, Evans, et al., 2021).

It could be valuable and efficient for research institutions to develop an RI education strategy which includes educational approaches tailored to different target groups (including students, researchers and other institutional stakeholders) and allows for continuous RI education (Barnes et al., 2006; Labib, Evans, et al., 2021). RI education guidelines, entailing recommendations and best practices, can provide considerations for institutions on what to include in their institutional RI education strategy. By 'guidelines', we refer to documents containing guidance, and by 'recommendations' we refer to the specific items in the guidelines. To ensure that guidelines are sensitive to stakeholders' actual RI education needs, they should be focused on practice and incorporate the perspectives and experiences of various research stakeholders. A co-creative approach to developing the guidelines, where

stakeholders are not only consulted but also directly involved in the guideline development process, is helpful to achieve this (den Breejen et al., 2012; Labib, Pizzolato, et al., 2021).

Together with various research stakeholders, we used an iterative co-creative methodology, which resulted in co-created guidelines on RI education for research institutions. In this paper, we describe the development of these guidelines and reflect on them by focusing on three questions: 1) Which recommendations are applicable across various RI education target groups?; 2) Are there any specific recommendations that are applicable to some target groups but not others?; and 3) What additional recommendations to research institutions, i.e. institutional officials and decision makers, are needed to increase awareness about RI in the institution, other than providing RI training?

Methods

The guidelines presented in this paper are the result of a combination of iterative steps used to co-create guidelines on a number of distinct RI topics (Figure 1). Here, we focus on the methods and results relating specifically to the guidelines on the topic of RI education. To obtain a first overview of potential considerations to include in these guidelines, we used insights from several previous empirical studies, which can be seen as preliminary steps in our guideline development process (Gaskell et al., 2019; Labib, Evans, et al., 2021; Labib, Roje, et al., 2021; Ščepanović et al., 2021; Sørensen et al., 2021). We next conducted four co-creation workshops together with various stakeholders to develop the guidelines, and then formed a working group to further revise and operationalize the developed guidelines.
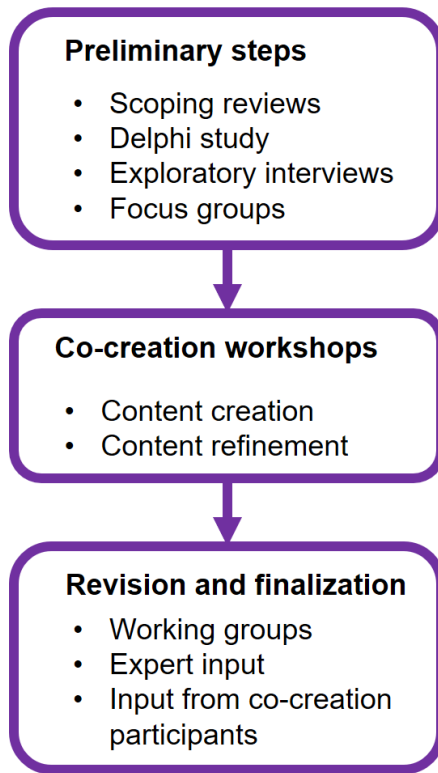
Figure 1 *Guideline co-creation process. Since the preliminary steps have been discussed in detail elsewhere (Labib, Evans, et al., 2021; Labib, Roje, et al., 2021; Roje et al., 2021, 2022; Ščepanović et al., 2021; Sørensen et al., 2021), this manuscript elaborates more on the co-creation workshops and revision and finalization steps of the guideline development process.*

Preliminary steps

We identified available recommendations on the topic of RI education, as well as gaps, by: 1) performing scoping reviews on best practices for RI promotion (including RI education) (Ščepanović et al., 2021) and the factors for successful implementation of these (Roje et al., 2022); 2) conducting 23 interviews with RI experts (Roje et al., 2021); and 3) a Delphi consensus-study with 68 research policy makers and research leaders across Europe (Labib, Roje, et al., 2021). Informed by these studies, we then conducted 30 focus groups with researchers and other research stakeholders from different disciplines and countries in Europe (Labib, Evans, et al., 2021) to explore their perspectives and preferences regarding RI education. Based on the insights gained, we compiled a comprehensive list of possible recommendations for research institutions on RI education (Lechner et al., 2020), which were represented as 'inspirations' (elaborated on further below) and served as input for a set of co-creation workshops.

Co-creation workshops

We conducted four co-creation workshops to jointly develop the RI guidelines together with various research stakeholders. The workshop methods have been described in detail elsewhere (Labib, Pizzolato, et al., 2021). The workshops included active involvement – rather than mere consultation – of stakeholders from the onset of the guideline development process (Labib, Pizzolato, et al., 2021). We followed co-creation workshop approach as elaborated on by Sanders and Stappers (2012), where stakeholders are engaged in creative workshops to jointly develop user-centered outputs. The workshops stimulated stakeholders to reflect on their experiences with RI education through the use of various interactive exercises making use of visual and textual stimuli to create ideas for guidelines, and then discuss these with others to build on each other's ideas, prioritize ideas, and make joint conclusions (Labib, Pizzolato, et al., 2021). The methods were aimed at incorporating the actual needs and perspectives of stakeholders (Sanders & Stappers, 2012), and can be considered particularly valuable for eliciting a broad range of ideas (Labib, Pizzolato, et al., 2021).

Participants

We used a purposive recruitment strategy to identify and invite participants who were potential lead users of the guidelines (i.e. have a responsibility in their implementation). Participants included RI experts and research administrators representing different countries, professional roles, and genders. We aimed to recruit 4-6 participants per workshop to allow for in-depth discussions (Labib, Pizzolato, et al., 2021). Other relevant stakeholders' specific needs in relation to RI education – including those of junior researchers and PhD students – identified in the preliminary steps (e.g. focus groups), were fed into the co-creation workshops as preparatory material that workshop participants received before joining the workshops. To invite participants, we simultaneously 1) approached contacts from our networks via email, followed by snowballing, and 2) approached people listed in internal databases of RI experts (e.g. ENERI, https://eneri.eu/; EARMA, https://www.earma.org/). We recruited 16 participants in total trough this strategy (Table 1).

Table 1: Characteristics of participants

Characteristics	Number of participants
Participating in each workshop	
Workshop 1	4
Workshop 2	5
Workshop 3	4*
Workshop 4	5*
Total	16
Gender	
Female	10
Male	6
Stakeholder type**	
Research manager	6
Senior researcher	2
Research head	1
RI coordinator	6
Publisher	1
Country	
Belgium	2
Finland	1
Germany	1
Ireland	2
Italy	1
Lithuania	1
Netherlands	2
Spain	3
Sweden	1
Switzerland	1
UK	1
Total number of countries	11

*One of these participants had also contributed to the first or second workshop
** We categorized participants' roles based on their job titles and positions. Research manager includes stakeholders with job titles such as research manager, research support manager, graduate education officer, research integrity officer/manager, and assistant to ombudsperson. Senior researcher includes researchers who are assistant, associate or full professors. Research head refers to researchers with positions as department, faculty, or institution wide leads (e.g. department heads, rectors). RI coordinator includes those with job titles such as RI or ethics coordinator, research coordinator, and scientific coordinator. Publisher refers to participants primarily representing a research publisher.

Workshop set-up and organization

The workshops were approved by the institutional review board of KU Leuven under dossier number G-2020011945. Prior to taking part in the workshops, participants received an information leaflet and signed an informed consent form. A detailed

workshop protocol can be found on the Open Science Framework (https://osf.io/8upmb/). We conducted four co-creation workshops to develop the RI education guidelines: workshops one and two were dedicated to content creation and workshops three and four were dedicated to content refinement (Figure 2, with more details available in Appendix I). We created 'inspirations' – images or short pieces of text representing different recommendations – based on a compilation of existing ideas and recommendations about RI education from the preliminary steps (Lechner et al., 2020). The 'inspirations' served as 'sensitization' material; 'sensitization' primes participants to various ideas and promotes creativity in co-creation workshops (Sanders & Stappers, 2012; Sleeswijk Visser et al., 2005). The 'inspirations' sensitized participants to ideas about RI education elicited in the preliminary steps, and promoted different interpretations, as well as encouraging discussion of out-of-the-box, new ideas (Labib, Pizzolato, et al., 2021). We sent the resulting 'inspirations' to all participants.

A week later, we conducted the first two workshops. During each workshop, we asked experts to create the content of RI education guidelines separately for 1) students, 2) researchers, and 3) other stakeholders, to ensure that relevant differences between target groups could be addressed. Furthermore, each workshop had a section dedicated to the role of providing directed advice and counseling for RI as a form of teaching about RI, to address forms of RI education that do not fall under the category of formal RI training. During these content creation workshops, we focused on generating a broad range of ideas for the content of the guidelines. Based on the discussions in the content creation workshops, we drafted a first version of the guidelines which we sent to participants of the content refinement workshops (i.e., workshops three and four). During the content refinement workshops, we asked participants to provide general comments, additions and concerns about the guidelines, for instance regarding redundancies, gaps, lack of clarity, conflicting statements, etc. Following the content refinement workshops, we revised the guidelines further and sent them to participants to provide any additional comments or suggestions. Further details about the workshop proceedings, technical details, and facilitation can be found in Appendix I.

Each workshop lasted approximately 3 hours. Due to the COVID-19 pandemic, we conducted all workshops online using the virtual collaborative whiteboard software program MIRO (https://miro.com), and the video conference program Zoom (https://zoom.us/). The workshops were led by a facilitator (NE, JT or KL), as well as one or two co-facilitators (DP, BT, IL, NS), and they were audio and video recorded, and transcribed. The role of the facilitator throughout the workshops was to guide the process of co-creation and ensure the inclusion of all participants' ideas, without providing input to the content. The detailed program for each workshop can be found on the Open Science Framework (https://osf.io/9bztf/).
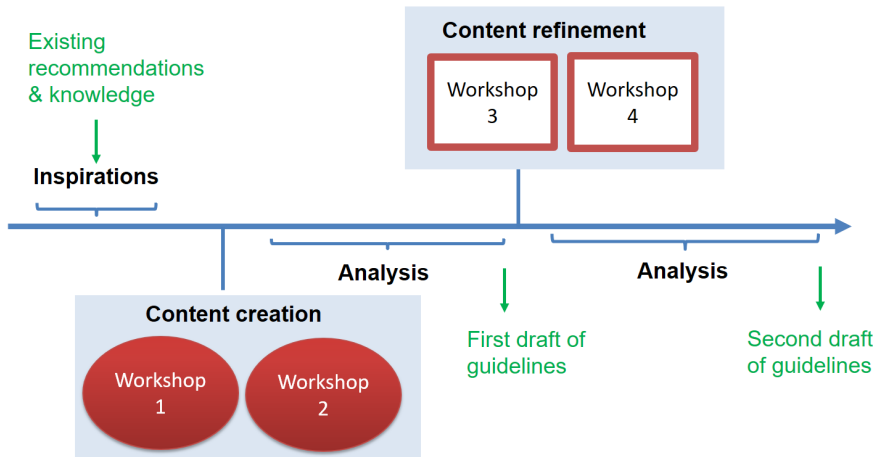
Figure 2 *Guideline co-creation project process*

<u>Analysis</u>

KL, IL, and JT used inductive (Boyatzis, 1998) and deductive thematic analysis (Crabtree & Miller, 1992) to analyze the results of the 'content generation' and 'content refinement workshops', respectively, through an analysis-on-the-wall approach using MIRO as described by Sanders and Stappers (2012). The analysis was done collaboratively by the coders as described in Appendix I as is common for co-creation methods (Sanders & Stappers, 2012), so as to include various perspectives in the coding and enrich the interpretation and construction of codes and themes. Differences between coders therefore contributed to a more nuanced understanding of themes, and contributed towards an iterative analysis process. A detailed code book including the theme and subtheme labels, and detailed descriptions and illustrative quotes for each was collaboratively developed by KL and NE (https://osf.io/y3c5n/). The code book was created per topic discussed in each workshop, namely the RI education of students, researchers, other RI stakeholders, and RI counseling and advise (which came to encompass all informal RI education approaches). Using the insights from the analysis, we developed the first and second draft of the RI education guidelines. To assess thematic saturation, we compared the insights gained during the content creation workshops (i.e. the analysis results and resulting guidelines) with the recommendations compiled from the preliminary steps, based on the views of other stakeholders (i.e. Lechner et al., 2020). Any points from the previous steps which had not been discussed in the content creation workshops were added to the guidelines, marked in a different color (as can be seen in Appendix II), and fed into the content refinement workshops, so that the content refinement workshop participants could comment on them. A detailed description of these analysis steps can be found in Appendix I.

Revision and finalization of the guidelines

After the co-creation workshops, the precise formulations of the recommendations in the guidelines needed further revision in order to be clear and usable (Labib, Pizzolato, et al., 2021). We organized a revision working group, composed of KL, GW and TK, which prioritized, reorganized, and optimized the draft recommendations in the guidelines (please see https://osf.io/f9ghj/for more details). The working group aimed to increase the understandability, implementability, and comprehensiveness of the guidelines. While revising the guidelines in the working group, we scrutinized the similarities and differences in the recommendations across the RI educational target groups, to ensure that the recommendations for each target group were relevant and appropriate.

We had intended to create four guidelines on topics that were determined together with stakeholders in one of the preliminary steps of the research (Labib, Roje, et al., 2021), namely the RI education of 1) students and researchers without a doctorate; 2) researchers with a doctorate; 3) research support staff and RI teachers; and 4) RI counseling and advice (as a form of education that falls outside of formal training). However, based on feedback from other research in the preliminary steps (Labib, Evans, et al., 2021) and the co-creation workshop participants about these categories, during the guideline revision process, we revised the categories to the RI education of 1) bachelor, master and PhD students, 2) post-doctorate and senior researchers, 3) other RI stakeholders, and 4) Continuous RI education. This required some substantial changes to the 4th category; we reworded the title of the guideline on RI counseling and advice to explicitly include all forms of informal RI education and included insights from the workshops that were not only about RI counseling and advice, but also about other ways of raising awareness about RI within an institution. Additionally, we removed RI counseling and advice recommendations that were not related to RI education (e.g. those dealing with misconduct), based on participants' suggestions.

Following this, we sent the guidelines to three external stakeholders with expertise in RI (MvdH, JPB, and MM), for feedback on how to further refine the guidelines to improve their implementability. We instructed experts to provide us with concrete feedback that we used to revise the guidelines directly. To ensure that the feedback from participants of the co-creation project was well considered and understood in the revision process, we also sent the revised guidelines to the co-creation workshop participants as a final member check (Thomas, 2017). All but one of the participants responded with approval of the guidelines or some additional suggestions – which we then integrated into the guidelines. Suggestions from the experts and participants were mostly related to refining the formulations in the recommendations (e.g. to recommend the use of 'diverse learning environments', rather than 'blended learning' for RI training). Other suggestions were related to implementation, and were not aimed at revising the guidelines themselves. A full overview of the suggestions received can be found on OSF: https://osf.io/we6pq.

<u>Results</u>

To adequately address the needs of students, researchers and research support staff, as well as incorporate various formal and informal approaches to RI education, we co-created four guidelines on RI education. Each guideline focused on a specific topic that we decided on together with various stakeholders (Labib, Roje, et al., 2021; Ščepanović et al., 2021; Sørensen et al., 2021), and then finalized in the further guideline development process. The guidelines focus on: a) RI education of bachelor, master and PhD students (https://osf.io/z7m3v/); b) RI education of post-doctorate and senior researchers, including all researchers with a doctorate ranging from early career researchers to full professors (https://osf.io/6d9ta/); c) RI education of other institutional RI stakeholders (e.g. ombudspersons, research integrity officers, etc., https://osf.io/ya3qj/); and d) continuous RI education (https://osf.io/ambg3/). Guideline c focuses on stakeholders, other than researchers, who play an important role regarding RI. This includes people involved in developing and implementing RI policy at the institution, handling complaints, raising awareness about RI, or providing support or information to researchers for good research practice (e.g. RIO, ombudsperson, data management officer). The exact tasks and responsibilities of RI stakeholders targeted in guideline c differ per institution and country in Europe, so the guidelines do not provide descriptions of each role in detail.

Guidelines a-c focus on the steps institutions can take to provide successful education to various target groups, while guideline d focuses on approaches for RI education other than formal RI training. To ensure that the guidelines are flexible enough to be incorporated in different institutional and country settings, we refrained from prescribing specific training aims or content in them, nor a specific theory about RI education. Instead, the guidelines provide higher level recommendations which need to be further operationalized and tailored to the local context of the institution. Institutions with many resources and already existing RI education policies in place may be able to include many of the recommendations directly, while those with fewer resources or existing RI infrastructure will need to phase the recommendations in slowly over time. We do not provide instructions for institutions on how to do this, because the order and manner of implementing recommendations will depend on the specific institutional context, but we provide 'in practice examples' for different recommendations which can serve as inspiration for institutions on where to get started. The main recommendations from each guideline are shown in Table 2, whereas the full versions are available in Appendices IV-VII (where the 'in practice examples' can also be found). As can be seen, there were commonalities in the guidelines across target groups, but there were also some important points of distinction that merit discussion.

Guideline title	**a. Guidelines for research institutions on the RI education of bachelor, master, and PhD students**
Recommendations	1. Integrate mandatory RI education into the bachelor and master curriculum 2. Deliver a mandatory RI course at the start of the PhD trajectory 3. Provide PhD students with follow-up elective courses on RI 4. Organize opportunities to discuss RI informally 5. Provide train-the-trainer education and basic qualifications for RI trainers 6. Use diverse learning environments, combining online and in-person elements in RI education 7. Focus on students' actual experiences with research rather than merely addressing theory in RI education 8. Motivate and reward students to actively take part in RI education 9. Evaluate educational programs
Guideline title	**b. Guidelines for research institutions on the RI education of post-doctorate and senior researchers**
Recommendations	1. Deliver mandatory training about RI for researchers starting new positions 2. Provide researchers with follow-up specialized training on RI 3. Involve senior researchers in the RI training of students and junior researchers 4. Organize opportunities to discuss RI informally 5. Provide train-the-trainer education and basic qualifications for RI trainers 6. Use diverse learning environments, combining online and in-person elements in RI education 7. Consult with researchers about their RI education needs and tailor education accordingly 8. Motivate and reward researchers to actively take part in RI education 9. Evaluate educational programs
Guideline title	**c. Guidelines for research institutions on the RI education of institutional RI stakeholders**
Recommendations	1. Provide institutional RI stakeholders who are not performing research with basic RI training 2. Organize events where RI stakeholders come together to ask questions, exchange experiences and discuss how to work together on RI

3. Provide train-the-trainer education and basic qualifications for RI trainers
4. Organize follow-up educational events when RI policies and regulations change
5. Provide opportunities for peer-to-peer learning about RI
6. Motivate and reward various RI stakeholders to actively take part in RI education
7. Evaluate educational programs

Guideline title	*d. Guidelines for research institutions on continuous RI education*
Recommendations	1. Provide researchers with educational RI resources to consult when needed
	2. Show institutional commitment to provide continuous RI education
	3. Provide researchers with contact persons who can support continuous RI education, by providing low-threshold, disciplinary-specific advice on day-to-day RI questions
	4. Develop policies to foster responsible supervision and leadership
	5. Develop policies for building a responsible research environment

Table 2 continued

I. Commonalities across target groups

1. Initial mandatory RI training

Across target groups, the co-creation participants recommended mandatory RI training to ensure that everyone in the institution is well-informed about RI. More specifically, co-creation participants thought that RI training should be mandatory when starting a new academic degree program (e.g. a bachelor, master, or PhD trajectory) or a new research position (e.g. new professorship, new postdoctoral research contract, etc.). Furthermore, participants stressed the importance of having RI trainers undergo train-the-trainer courses to ensure that they are not only aware of RI theory, but are also equipped with the necessary didactic skills and tools to train students and researchers. Participants also highlighted that other RI stakeholders such as ombudspersons and RI officers would benefit from educational activities about RI, although they did not explicitly mention formal training for this. In our revision working group, we proposed to extend the recommendation for formal basic training also to these RI stakeholders when starting new positions, to ensure that they have sufficient knowledge of RI to be able to support researchers with RI.

2. Follow-up RI training

The co-creation workshop participants recommended that all target groups should be provided with periodic follow-up RI training. At the bachelor and master level, they

suggested that discussing RI in depth during the thesis research phase would be most appropriate. For PhD students, it was thought that follow-up courses on discipline-specific RI topics such as data management would be most helpful in supporting students' research practice. Similarly, the co-creation workshop participants suggested that repeating follow-up disciplinary-specific training every 2-3 years to address specific RI issues such as new developments in research, would also be valuable for post-doctorate and senior researchers to keep up with the newest regulations and policies, as well as to refresh their knowledge and skills on RI. The same reasoning applied to other RI stakeholders such as RI officers, ombudspersons, and policy staff, to suggest that institutions should provide new training for these target groups every time new policies and regulations are introduced.

3. Informal discussions about RI

The importance of informal RI discussions were highlighted during our co-creation workshops, where many participants thought that discussing RI experiences and problems in informal meetings together with colleagues, supervisors and supervisees, would be valuable for continuous RI peer-to-peer learning. However, some participants were concerned that it might be difficult for institutions to coordinate informal meetings. Therefore, in our revision working group, we recommended that institutions should stimulate and support departments and teams to organize informal events and integrate RI questions in them (e.g. by providing institutional awards for the best RI events), rather than coordinate this process themselves. The co-creation workshop participants agreed that RI education can contribute to a more responsible research culture, while an open research culture is a prerequisite for fruitful interactions during RI education. To take this consideration into account, all the RI education guidelines recommend that institutions develop policies that foster a responsible research environment (addressing community building, https://osf.io/7fn2x; diversity and inclusion, https://osf.io/fwa5c; managing competition and publication pressure, https://osf.io/ya3qj; adequate education and skills training to researchers, https://osf.io/2p3vf), as discussed by the co-creation workshop participants and our revision working group.

4. Motivation, incentives and rewards

To motivate students and researchers to actively take part in RI education, our co-creation workshop participants suggested that RI educational events should emphasize the importance of RI (e.g. for research quality) and use a positive approach to RI (focused on promoting responsible research rather than discussing misconduct). This can involve highlighting the importance of RI education for researcher productivity, and professional and scientific success. A positive approach entails focusing on the real-life challenges faced in research practice rather than only teaching RI theory, telling trainees what to do, or focusing on the prevention of research misconduct. The co-creation workshop participants also stressed that institutions should provide suitable incentives and rewards to ensure students and researchers are actively engaged in RI education (e.g. free lunches, certificates, promotions). Our co-creation workshop participants additionally highlighted that it is not only researchers and students who

should be rewarded for taking part in RI education and contributing towards improved RI, but also other RI stakeholders such as RI trainers and officers. Although motivation, incentives and rewards were recommended for all target groups, participants stressed that these should be tailored specifically for the target group as the same incentives and rewards might not work for everyone.

<u>5. Evaluation of educational events</u>

Evaluation was seen as crucial for the continuous improvement and update of RI education. Due to potential feasibility challenges in conducting objective outcome evaluations measuring changes in researcher behavior in every research institution, the workshop participants thought that process evaluations by research institutions would also be informative for evaluating educational events. They suggested using qualitative and quantitative measures for process evaluations. For instance, participants suggested to evaluate educational events using experiential data, such as how useful students perceive the event to be, as well as quantitative data not related to 'effectiveness', such as the number of individuals registering and attending optional RI trainings.

II. Points of distinction between target groups

<u>1. Bachelor, master, and PhD students</u>

Our co-creation participants thought that incentivizing RI education for bachelor, master and PhD students is relatively easy compared to other target groups. They suggested that providing students with tangible incentives – for example digital badges or other incentives tailored to students in different stages of their educational trajectory, disciplinary backgrounds and institutions – for completing trainings would suffice in motivating students to actively engage with RI education. Furthermore, they recommended providing all students with a substantial number of contact hours focused on RI (e.g. in the form of a complete course for PhD students), as this would ensure sufficient familiarization with RI at the start of their education about research, and would not be difficult to mandate.

<u>2. Post-doctorate and senior researchers</u>

Our co-creation workshop participants suggested that motivating post-doctorate and senior researchers to participate in RI education is difficult since researchers in these career stages are increasingly busy and have competing priorities. They stressed that institutional RI education policies should sufficiently address this concern to ensure engagement with RI education among researchers across seniority levels. Multiple recommendations on how to do this are offered in the guidelines. These include suggestions to consider RI and RI education in promotions and career assessments. The guidelines also offer other simpler suggestions such as labelling training as a 'Masterclass' rather than a 'training' to make them sound more appealing to researchers. To reduce the burden that mandatory trainings would impose, many of the co-creation workshop participants suggested to provide this target group with

small training events (e.g. 1-2 hour workshops), rather than full courses. It was also thought that a 'bottom-up' approach to RI training would be valuable to make RI training attractive and relevant for post-doctorate and senior researchers. Such an approach would involve consulting researchers beforehand to capture the RI topics they need support and help with (i.e. conducting a training needs analysis), and tailoring RI trainings accordingly. Tailoring trainings to trainees' needs was thought to be especially relevant considering that very experienced senior researchers might have different training needs compared to researchers in earlier stages of their career.

3. Other institutional RI stakeholders

Co-creation workshop participants mentioned that peer-to-peer learning is likely most suitable for the RI education of other institutional RI stakeholders (e.g. ombudspersons, RI officers, trainers, policy makers, etc.). Participants suggested that research institutions foster peer-to-peer learning by supporting the organization of peer consultation meetings and other informal events where various RI stakeholders can come together to share experiences about RI and learn from each other. However, participants did recommend formal RI training for RI trainers, focusing not only on RI theory but also on didactical skills. They highlighted that national and European level support groups and networks for institutional RI stakeholders would be valuable to address the lack of availability of RI resources at some institutions, as well as fostering the sharing of experiences across institutions and countries. Many participants suggested that it would be helpful to share RI cases and educational materials in such a network in order to learn from each other and avoid 'reinventing the wheel'. In our revision working group, we operationalized this recommendation to directly address research institutions, by suggesting that institutions provide RI stakeholders with opportunities to engage in peer-to-peer learning (e.g. by hosting networking events, providing funds or time to employees, etc.).

III. Additional measures to increase awareness about RI

Our co-creation workshop participants stressed that raising awareness about RI requires more than one-off RI trainings; therefore, in our revision working group, we decided to dedicate one of the RI education guidelines to continuous RI education to highlight this concern. Stimulated by advice from one of the experts we consulted in the guideline revision process, we decided to explicitly state that our definition of 'RI education' is broad and entails all means of creating awareness about RI – rather than only constituting formal education – in the preamble of the guidelines. This guideline on continuous RI education incorporates the co-creation workshop participants' recommendations regarding institutional commitment to RI education, provision of necessary educational resources, creation of policies in the institution on building a responsible research environment, inclusion of responsible supervision and mentoring, and provision of low-threshold advice to researchers about RI through informal RI 'champions' or 'stewards' as a means of increasing awareness about RI. Regarding our recommendations in this guideline for building a responsible research environment and fostering responsible supervision, our continuous RI education guideline links to more

detailed guidelines that we are developing in separate topics on the themes of research environment and supervision (Pizzolato et al., 2021, 2022).

Discussion

We co-created institutional guidelines together with various research stakeholders regarding the RI education of a) bachelor, master and PhD students, b) post-doctorate and senior researchers, and c) other institutional RI stakeholders, as well as d) continuous RI education. These guidelines can help research institutions develop an overarching strategy for RI education that includes various educational approaches and addresses all relevant target groups. Across target groups, our guidelines indicate that institutions should organize continuous RI education using multiple formal and informal educational events (e.g. workshops, courses, informal discussions, etc.) and using target-group-appropriate incentives and rewards to actively motivate trainees to stay engaged in RI practices, for instance by including participation in RI education in promotion procedures for senior researchers. Furthermore, the guidelines suggest that education should focus on the concrete needs and practical challenges that participants deal with – a finding supported by a recent systematic review (Katsarov et al., 2022) – and use regular process evaluations to ensure constant updating and improvement. Moreover, our guidelines highlight that a holistic RI education approach will not only require provision of formal RI trainings, but also additional educational approaches (e.g. responsible supervision) so as to support continuous education.

Given that current RI education often consists of stand-alone courses on RI (Abdi, Pizzolato, et al., 2021), the implementation of continuous RI education will require substantial effort and commitment by research institutions to organize, design and deliver additional RI training events to various target groups. Although this could be perceived as a high burden by research institutions (Sørensen et al., 2021), we believe this commitment is necessary given that it is highly unlikely that a single course or workshop will be sufficient in influencing trainees' perceptions and behaviors relating to RI (Kalichman, 2014). However, further empirical research on the effects that RI education has on researchers' behavior is urgently needed to confirm this. To increase the feasibility of providing continuous RI education, institutions could make use of learning approaches utilizing different learning mediums where possible; they could consider using already existing openly accessible online RI trainings and resources, they could integrate relevant RI discussions in existing research courses, workshops, and department meetings, or they could cooperate with external trainers and institutions to share the provision of RI education. This recommendation is supported by evidence suggesting that blended learning approaches are highly effective for ethics instruction (Todd et al., 2017).

Almost every stakeholder we talked to, not only in the co-creation workshops, but also in the preliminary steps of the research process, agreed that some form of mandatory education on RI was needed across institutions. What some disagreed about was the form that this training should take (e.g. a full course versus a one hour workshop to update researchers about a new research development). While it can be assumed that mandating RI training for senior researchers is likely to meet resistance (Fanelli, 2019; Labib, Evans, et al., 2021), our co-creation workshop participants

recommend mandatory RI training for senior as well as junior researchers. This is based on the view that mandatory training is the only way to ensure that all researchers – rather than only those who already consider RI as important – take part in RI education. Furthermore, training more senior researchers, especially regarding how to lead, manage and mentor their research teams responsibly and foster responsible research practices that warrant rigor, reproducibility and research quality, plays a crucial role in shaping institutional cultures including rigor, reproducibility, and integrity (Antes et al., 2016, 2019; McIntosh et al., 2020; Pizzolato et al., 2022). Our participants provided suggestions about how to provide training that is appropriate to different contexts (e.g. a session for seniors researchers to discuss the implications of new laws and policies, such as the General Data Protection Regulation, GDPR).

To reduce potential resistance and to ensure that trainees are actively engaged in RI education, the RI education guidelines stress the importance of providing suitable incentives and rewards for participating in RI training (e.g. tying RI education to promotions, using tangible rewards, etc.). Our guidelines further suggest to tailor incentives and rewards to their target group: a finding that is in line with existing literature suggesting that effective incentives and rewards are different for junior than for senior researchers (Fanelli, 2019; Labib, Evans, et al., 2021). Rewarding and incentivizing participation in RI education is also in line with other existing initiatives in the research community which state that researcher evaluations should consider a broader range of contributions and should value responsible research practices (Aubert Bonn & Bouter, 2021; Moher et al., 2020).

Motivation to actively participate in training will also depend on the extent to which the RI training appeals to the needs of each target group. Therefore, our guidelines stress the importance of providing RI education that focuses on the specific needs and challenges of the education target group. Focusing on real life cases of RI dilemmas that come up in research practice when teaching students about RI can help increase the relevance of RI training and has been suggested by others as well (Fanelli, 2019; Kalichman, 2014; Katsarov et al., 2022). Our recommendation to use a training-needs-analysis to ensure that post-doctoral and senior researchers can determine what should be included in their RI trainings and how, rather than following trainings focusing on methods and context predetermined as relevant for them by trainers, has to our knowledge, not been discussed in previous literature. However, we believe that such an approach is important, particularly to prevent researchers from perceiving RI training as a box-ticking exercise (Labib et al., 2022). Especially considering that researchers of various disciplines and ranks (e.g. full professors as compared to less experienced post-doctorate researchers) may have different needs (ENERI, 2017), using such a bottom-up, tailored approach to RI education is likely to be valuable, albeit the associated financial and time costs present challenges for implementation.

To ensure that RI education is continuously updated and improved over time, the RI education guidelines emphasize the importance of evaluating RI educational events. However, our results also suggest that evaluating educational events on their effects on researcher behavior will likely be difficult, if not impossible, indicating the need for institutions to engage in subjective process evaluations (e.g. on perceptions of training usefulness) over outcome-oriented evaluations (e.g. relating to changes in actual behaviors). This approach to evaluation might seem unsatisfactory for trainers

who would like to develop RI trainings based on solid empirical outcome research, as well as for institutions who would like to know that their RI education policies are actually impacting research practice. However, we would argue that it is not the responsibility of single research institutions to provide full insight into what makes RI education effective; a focus on subjective process evaluations is more feasible and can still provide valuable information to trainers. For instance, evaluations on subjective data like stakeholder experiences can provide valuable information about the contextual mechanisms and processes that influence the success of educational initiatives (Hamza et al., 2020). Of course, if institutions have the possibility and means to also conduct behavioral outcome-oriented evaluations, that can be beneficial in ensuring that RI education improves trainee learning, skills development, and behavior change.

Relying on process outcomes in evaluations for those institutions unable to conduct behavioral outcome-oriented evaluations is likely to be more acceptable when the behavioral effects of the educational approach have already been documented in the literature. There are substantial current efforts to find strategies to measure RI training effectiveness on outcomes such as improvements in moral reasoning and changes in behavior (e.g. Abdi, Fieuws, et al. 2021; Katsarov et al. 2020; Watts et al. 2017); these can provide institutions with information on the effectiveness of RI education and supplement institutional efforts in process evaluations of various educational programs.

Strengths and limitations

The guidelines we discuss in this paper are – to our knowledge – the first to provide an overview of what to include in research institutions' overarching RI education strategy. The guidelines are a result of an iterative co-creative research process, involving various potential lead-users from different parts of Europe. The co-creation workshop methods we have used have allowed us to incorporate a wide range of perspectives in the guideline, including heterogeneity in participants from high and low resource institutions and countries in Europe. Further research is needed to explore the implementability of the guidelines in other settings, and in low and middle income countries.

As such, the guidelines have been developed with a focus on incorporating various research stakeholders' actual RI education needs and perspectives. The qualitative approach to the guideline development process allowed us to understand stakeholders' perspectives about RI education in depth and in a nuanced way. The active co-creators of the project represent the guideline lead-users (e.g. RI officers and trainers) rather than end-users (e.g. junior and senior researchers). It was necessary to focus on lead-users since the intensity of co-creation workshops limits the number of participants that could be included in them (Sanders & Stappers, 2012). Despite the fact that we only have a limited number of participants from different stakeholder groups in the co-creation workshops, we consider that the guidelines still provide a comprehensive and diverse user input given the engagement of both lead-users and end-users in our preliminary work (Labib, Evans, et al., 2021; Labib, Roje, et al., 2021;

Roje et al., 2021). Furthermore, our approach might be limited by the fact that we did not provide an open call for feedback on the guidelines.

It would be valuable to obtain insights on a larger sample of experts' thoughts on the importance, relevance and feasibility of the guidelines, using quantitative means. Actual testing of the guidelines in a number of research institutions will be necessary to further refine the guidelines and make them implementable on a large scale. Such testing can provide insights about how the recommendations can be implemented with few resources. A pilot study can also help to create a more comprehensive and robust set of 'best practice' examples for the recommendations in each guideline. Institutions interested in using the guidelines will need to take into account costs, local capacity, cultural issues, and context-specific factors during implementation of the guidelines (Horbach & Sørensen, n.d.; Konach et al., 2022). Implementation of the recommendations will likely vary between institutions which already provide some RI educational programs and those that do not.

Conclusions

Our work provides experience-based co-created guidance to research institutions on important considerations for developing a successful RI education strategy. Our guidelines on RI education address the needs of students, researchers and other RI stakeholders, and take into account various approaches to RI education. In the guidelines, we recommend mandatory RI training; follow-up refresher training; informal discussions about RI; appropriate rewards and incentives for active participation in RI education; and evaluation of RI educational events across target groups. Each of our four guidelines can be considered a distinct tool that institutions can access, adapt and implement to meet their institution-specific RI education needs. Research institutions across and outside of Europe can use our guidelines as tools to strengthen their RI education efforts and consequently contribute towards better quality and more trustworthy research.

List of abbreviations

RI – Research integrity
RRPs – responsible research practices
QRPs – questionable research practices

Declarations

Ethics approval

The workshops were approved by the institutional review board of KU Leuven under dossier number G-2020011945. Prior to taking part in the workshops, participants received an information leaflet and signed an informed consent form.

Consent for publication

Not applicable

Availability of data and materials

Due to privacy reasons, the co-creation workshop transcripts used for this article are not publicly available. However, we have made the anonymized code book, including quotes, publicly available on the Open Science Framework: https://osf.io/y3c5n/. We excluded all descriptions about the characteristics of the quote owners (e.g. country, role, gender), to prevent the identifiability of the data, considering that the majority of the co-creators are acknowledged in this paper.

Competing interests

The authors have no conflicts of interest to declare.

Authors' contributions

KL contributed to the preliminary steps of the guideline development process, designed the co-creation workshops and analyzed the resulting data, led the working group for the refinement of the guidelines, and drafted and revised the manuscript. NE contributed towards the preliminary steps of the guideline development process, the co-creation study design and data analysis, and revision of the manuscript. DP contributed towards the co-creation study design and data collection, and revision of the manuscript. NAB contributed towards the co-creation workshop data analysis, designed the protocol used in the revision working group of the guidelines, and revised of the manuscript. GW, LB and KD contributed to the co-creation workshop study design, and revision of the manuscript. GW and LB were additionally involved in the guideline development process. GW, together with TK, was a member of the revision working group of the guidelines; TK also revised the manuscript. ML contributed towards the design of the protocol used in the guideline revision working group and revised the manuscript. JT contributed to the preliminary steps of the guideline development process, revised the co-creation workshop study protocol and collected the resulting data, revised the manuscript, and supervised the work.

We would like to thank Pieter Jan Stappers and Katinka Bergema for their valuable advice on co-creation methodology and their contribution towards the study design, as well as the three experts who provided us with input on the revision of the guidelines: Mariëtte van den Hoven, Julia Priess-Buchheit, and Malcolm MacLeod. We would also like to thank all members of the SOPs4RI project who contributed to the various stages of the guideline development process, and the co-creation project including Iris Lechner, Nik Claesen, Borana Taraj, and Nikolaos Skoulikaris, and Roshni Jegan.

Finally, we would like to send our gratitude to the three anonymous reviewers who helped us to strengthen this manuscript.

References

Abdi, S., Fieuws, S., Nemery, B., & Dierickx, K. (2021). Do we achieve anything by teaching research integrity to starting PhD students? Humanities and Social Sciences Communications, 8(1), 1–7.

Abdi, S., Pizzolato, D., Nemery, B., & Dierickx, K. (2021). Educating PhD Students in Research Integrity in Europe. Science and Engineering Ethics, 27(1), 1–19.

All European Academies. (2017). The European Code of Conduct for Research Integrity. All European Academies. https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

Antes, A. L., Kuykendall, A., & DuBois, J. M. (2019). Leading for research excellence and integrity: A qualitative investigation of the relationship-building practices of exemplary principal investigators. Accountability in Research, 26(3), 198–226.

Antes, A. L., Mart, A., & DuBois, J. M. (2016). Are leadership and management essential for good research? An interview study of genetic researchers. Journal of Empirical Research on Human Research Ethics, 11(5), 408–423.

Aubert Bonn, N., & Bouter, L. (2021). Research assessments should recognize responsible research practices—Narrative review of a lively debate and promising developments. In OSF Preprints. https://doi.org/10.31222/osf.io/82rmj

Barnes, B. E., Friedman, C. P., Rosenberg, J. L., Russell, J., Beedle, A., & Levine, A. S. (2006). Creating an infrastructure for training in the responsible conduct of research: the University of Pittsburgh's experience. Academic Medicine, 81(2), 119–127.

Boehme, O., Föger, N., Hiney, M., Peatfiled, T., & Petiet, F. (2016). Research Integrity Practices in Science Europe Member Organisations. https://doi.org/10.5281/zenodo.5060196

Bouter, L. (2018). Fostering responsible research practices is a shared responsibility of multiple stakeholders. Journal of Clinical Epidemiology, 96, 143–146.

Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. SAGE Publications Inc.

Crabtree, B. F., & Miller, W. F. (1992). A template approach to text analysis: developing and using codebooks. In B. F. Crabtree & W. F. Miller (Eds.), Doing qualitative research (pp. 163–177). SAGE Publications Inc.

den Breejen, E. M. E., Nelen, W. L. D. M., Knijnenburg, J. M. L., Burgers, J. S., Hermens, R. P. M. G., & Kremer, J. A. M. (2012). Feasibility of a wiki as a participatory tool for patients in clinical guideline development. Journal of Medical Internet Research, 14(5), e2080.

ENERI. (2017). Report of Stakeholder/Focus Group Workshop in Athens. https://eneri.eu/wp-content/uploads/2018/06/Report-Stakeholder-Workshop-Athens.pdf

Evans, N., Marusic, A., Foeger, N., Lofstrom, E., van Hoof, M., Vrijhoef-Welten, S., Inguaggiato, G., Dierickx, K., Bouter, L., & Widdershoven, G. (2021). Virtue-based ethics and integrity of research: train-the-trainer programme for upholding the principles and practices of the European Code of Conduct for Research Integrity (VIRT2UE). Research Ideas and Outcomes, 7, e68258.

Fanelli, D. (2019). *MLE on Research Integrity – Training & Education – Thematic Report No 4.* https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE RI No4 - Training and education.pdf

Gaskell, G., Ščepanović, R., Buljan, R., Utrobičić, A., Marušić, A., Reyes Elizondo, A., Kaltenbrunner, W., Labib, K., & Tijdink, J. (2019). *D3.2: Scoping reviews including multi-level model of research cultures and research conduct (A. Marušić (ed.)).* https://www.sops4ri.eu/wp-content/uploads/D3.2_Scoping-reviews-including-multi-level-model-of-research-cultures-and-research-conduct.pdf

Hamza, D. M., Ross, S., & Oandasan, I. (2020). Process and outcome evaluation of a CBME intervention guided by program theory. *Journal of Evaluation in Clinical Practice, 26(4),* 1096–1104.

Haven, T., Tijdink, J., Martinson, B. C., & Bouter, L. (2019). Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. *PloS One, 14(1),* e0210599.

Horbach, S., & Sørensen, M. P. (n.d.). *How to create and implement a Research Integrity Promotion Plan (RIPP).* https://sops4ri.eu/wp-content/uploads/Implementation-Guideline_FINAL.pdf

Joynson, C., & Leyser, O. (2015). The culture of scientific research. *F1000Research, 4(46).*

Kalichman, M. (2014). Rescuing responsible conduct of research (RCR) education. *Accountability in Research, 21(1),* 68–83.

Kalichman, M., & Plemmons, D. K. (2007). Reported goals for responsible conduct of research courses. *Academic Medicine, 82(9),* 846–852.

Katsarov, J., Andorno, R., Krom, A., & van den Hoven, M. (2022). Effective Strategies for Research Integrity Training—a Meta-analysis. *Educational Psychology Review, 34,* 935–955. https://doi.org/https://doi.org/10.1007/s10648-021-09630-9

Konach, T., Horbach, Serge Rochambeau, M., & Taraj, B. (2022). *D7.2: Report on Pilot Studies.* https://sops4ri.eu/wp-content/uploads/D7.2_Report-on-Pilot-Studies.pdf

Labib, K., Evans, N., Roje, R., Kavouras, P., Reyes Elizondo, A., Kaltenbrunner, W., Buljan, I., Ravn, T., Widdershoven, G., Bouter, L., Charitidis, C., Sørensen, M. P., Tijdink, J., Scepanovic, R., Kavouras, P., Elizondo, A. R., Kaltenbrunner, W., Buljan, I., Ravn, T., … Bouter, L. (2021). Education and training policies for research integrity: Insights from a focus group study. *Science and Public Policy, 49(2),* scab077. https://doi.org/https://doi.org/10.1093/scipol/scab077

Labib, K., Pizzolato, D., Stappers, P. J., Evans, N., Lechner, I., Widdershoven, G., Bouter, L., Dierickx, K., Bergema, K., & Tijdink, J. (2021). Using co-creation for guideline development - How why and when? *OSF Preprints.* https://doi.org/10.31219/osf.io/cg3rw

Labib, K., Roje, R., Bouter, L., Widdershoven, G., Evans, N., Marušić, A., Mokkink, L., & Tijdink, J. (2021). Important topics for fostering research integrity by research performing and research funding organizations: a Delphi consensus study. *Science and Engineering Ethics, 27(47),* 1–22. https://doi.org/10.1007/s11948-021-00322-9

Labib, K., Tijdink, J. K., Sijtsma, K., Bouter, L., Evans, N., & Widdershoven, G. (2022). How to combine rules and commitment in fostering research integrity? *MetaArXiV Preprints.* https://doi.org/10.31222/osf.io/sx58q

Lechner, I., Tijdink, J., Sørensen, M. P., Ravn, T., Bendtsen, A.-K., Labib, K., Kaltenbrunner, W., Reyes Elizondo, A. E., & Kavouras, P. (2020). *D4.3: Second version of SOPs and guidelines.* https://sops4ri.eu/wp-content/uploads/D4.3_Second-version-of-SOPs-and-guidelines-1.pdf

Masadeh, M. (2012). Training, education, development and learning: what is the difference? *European Scientific Journal, 8(10).*

McIntosh, T., Sanders, C., & Antes, A. L. (2020). Leading the people and leading the work: Practical considerations for ethical research. *Translational Issues in Psychological Science, 6(3),* 257.

Mejlgaard, N., Bouter, L., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A.-K., Charitidis, C. A., Claesen, N., Dierickx, K., Domaradzka, A., Reyes Elizondo, A., Foeger, N., Hiney, M., Kaltenbrunner, W., Labib, K., Marušić, A., Sørensen, M. P., Ravn, T., Ščepanović, R., … Veltri, G. A. (2020). Research integrity: nine ways to move from talk to walk. *Nature, 586,* 358–360.

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A.-M., Foeger, N., & Dirnagl, U. (2020). *The Hong Kong Principles for assessing researchers: Fostering research integrity. PLoS Biology, 18(7), e3000737.*

Pizzolato, D., Dierickx, K., Tijdink, J., Labib, K., Lechner, I., Bonn, N. A., Taraj, B., Claesen, N., & Gaskell, G. (2021). *D4.4: Report on the Co-Creation Workshops.* https://sops4ri.eu/wp-content/uploads/D4.4_Report-on-the-Co-Creation-Workshops.pdf

Pizzolato, D., Labib, K., Skoulikaris, N., Evans, N., Roje, R., Kavouras, P., Aubert Bonn, N., Dierickx, K., & Tijdink, J. (2022). *How can research institutions support responsible supervision and leadership? Accountability in Research.* https://doi.org/https://doi.org/10.1080/08989621.2022.2112033

Roje, R., Reyes Elizondo, A., Kaltenbrunner, W., Buljan, I., & Marušić, A. (2022). *Factors influencing the promotion and implementation of research integrity in research performing and research funding organizations: a scoping review. Accountability in Research.* https://doi.org/https://doi.org/10.1080/08989621.2022.2073819

Roje, R., Tomić, V., Buljan, I., & Marušić, A. (2021). *Development and implementation of research integrity guidance documents: Explorative interviews with research integrity experts. Accountability in Research, 1–38.*

Sanders, E., & Stappers, P. J. (2012). *Convivial toolbox: generative research for the front end of design. BIS Publishers.*

Ščepanović, R., Labib, K., Buljan, I., Tijdink, J., & Marušić, A. (2021). *Practices for Research Integrity Promotion in Research Performing Organisations and Research Funding Organisations: A Scoping Review. Science and Engineering Ethics, 27(4), 1–20.* https://doi.org/10.1007/s11948-021-00281-1

Sleeswijk Visser, F., Stappers, P. J., Van der Lugt, R., & Sanders, E. (2005). *Contextmapping: experiences from practice. CoDesign, 1(2), 119–149.*

Sørensen, M. P., Ravn, T., Marušić, A., Elizondo, A. R., Kavouras, P., Tijdink, J. K., & Bendtsen, A.-K. (2021). *Strengthening research integrity: which topic areas should organisations focus on? Humanities and Social Sciences Communications, 8(1), 1–15.*

Thomas, D. R. (2017). *Feedback from research participants: are member checks useful in qualitative research? Qualitative Research in Psychology, 14(1), 23–41.*

Tijdink, J., Bouter, L., 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.*

Titus, S., Wells, J. A., & Rhoades, L. J. (2008). *Repairing research integrity. Nature, 453, 980–982.*

Todd, E. M., Watts, L. L., Mulhearn, T. J., Torrence, B. S., Turner, M. R., Connelly, S., & Mumford, M. D. (2017). *A meta-analytic comparison of face-to-face and online delivery in ethics instruction: the case for a hybrid approach. Science and Engineering Ethics, 23(6), 1719–1754.*

Watts, L. L., Medeiros, K. E., Mulhearn, T. J., Steele, L. M., Connelly, S., & Mumford, M. D. (2017). *Are ethics training programs improving? A meta-analytic review of past and present ethics instruction in the sciences. Ethics & Behavior, 27(5), 351–384.*

Supplementary materials

Appendix I: Detailed methodology used for the SOPs4RI co-creation workshops

An overview of the demographics of participants included in each co-creation workshop can be found in Table 1.

Table 1: Characteristics of participants included in each co-creation workshop

Number of participants	Countries	Stakeholder types
Workshop 1		
4	Belgium, Sweden, Netherlands, Ireland	RI coordinator, Research manager
Workshop 2		
5	Netherlands, Lithuania, Spain, Italy	Senior researcher, Research head, RI coordinator
Workshop 3		
4	Sweden, Spain, Finland, Switzerland	RI coordinator, Research manager, Publisher
Workshop 4		
5	Netherlands, Belgium, Ireland, UK, Germany	Senior researcher, Research manager

The co-creation process used to develop the RI education recommendations consisted of a number of steps, including: creating inspirations, content creation, a first round of analysis, content refinement, and a second round of analysis (Figure 1). The role of the facilitator throughout the workshops was to guide the process of the co-creation and ensure the inclusion of all participants' ideas, without providing input to the content. Some details on each step can be found below, while (1,2) provides the full overview of the methods used for the workshops.
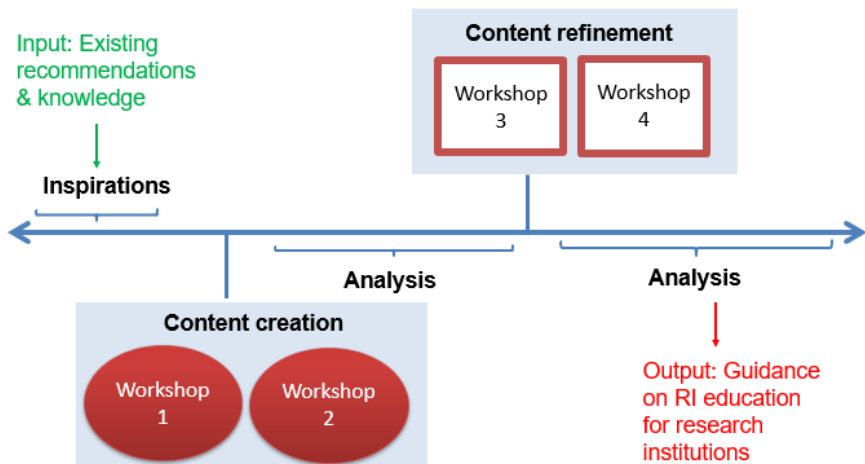
Figure 1 Guideline co-creation process in SOPs4RI project

<u>*Inspirations*</u>

To make use of existing recommendations about RI education compiled by the SOPs4RI consortium (3), we created 'inspirations' – images and/or short pieces of text representing different recommendations – which we used to evoke ideas in the workshop participants without steering them into specific directions, as they could be interpreted in multiple ways (https://osf.io/8dzxg/). For instance, we used the text 'knowledge' and the image of a crystal ball to represent the possibility of focusing on the acquisition of knowledge, or reflective skills in education, respectively. To ensure participants' familiarity with the inspirations during the workshops, the inspirations were sent to participants a week before the workshop. During this time, participants were asked to reflect on the inspirations, select three which they found most striking, and provide a rationale for their choices on the MIRO board (Figure 2).

Prior to attending the first workshops, participants were sent inspirations to stimulate ideas about the workshop topic.

As part of the preparation (i.e. 'sensitization') for the workshops, participants were asked to select 3 striking inspirations, paste them into the boxes shown here, and explain why they found them striking. These were then used during the introduction round of the actual workshops; each participant had to explain their selections to the other participants in the workshop, as an ice-breaker.

Please note that the names of the participants (on the sticky notes) have been blurred.
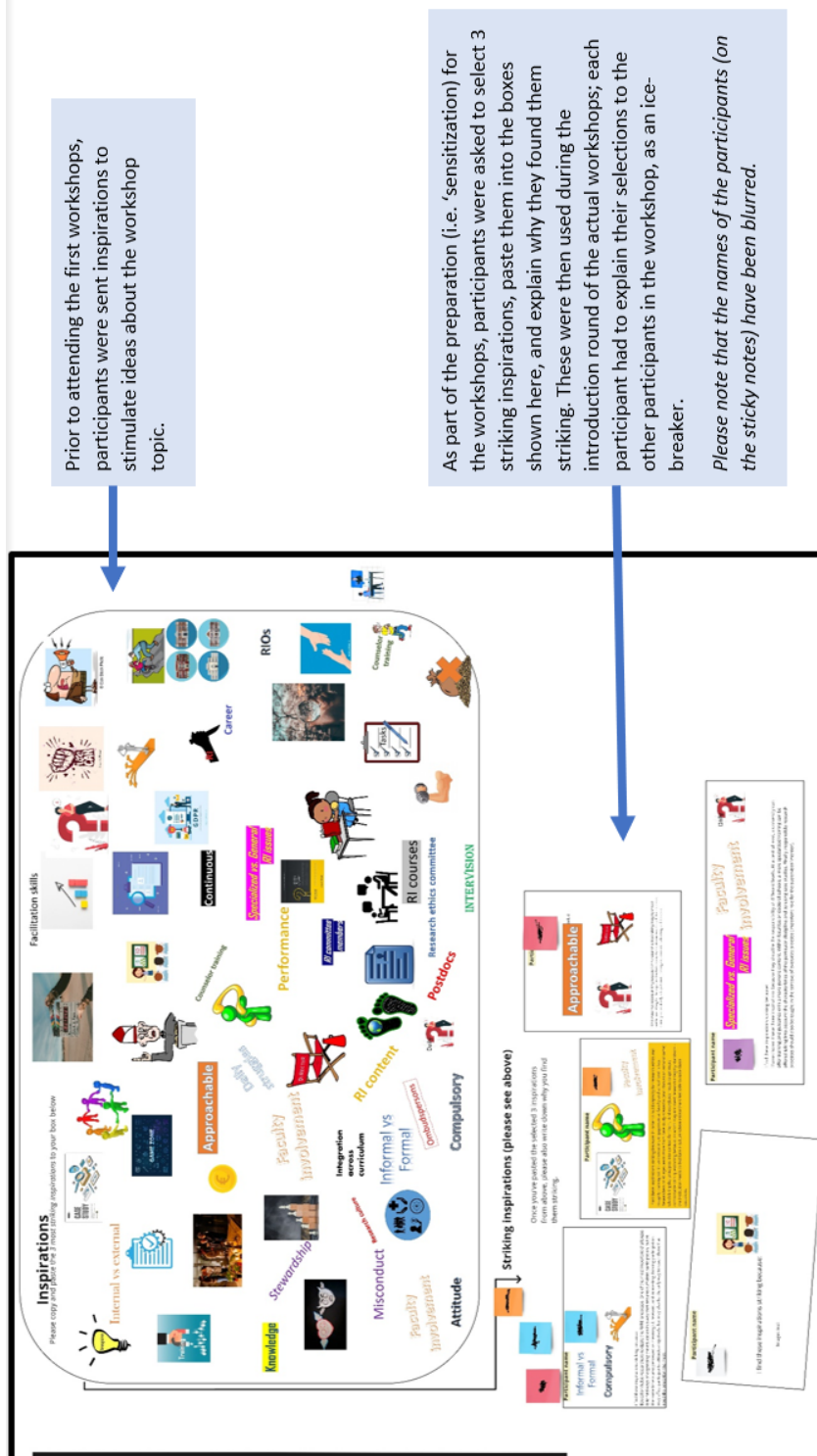

Figure 1 *Use of inspirations in preparation of the co-creation workshops (i.e. the sensitization phase)*

Content creation

During the 'content creation' workshops, we asked participants create the guideline content of four different topics related to RI education. We did this by asking participants to dream about what an ideal institutional education plan would be for: 1) students (including at the bachelor, master and PhD level), 2) researchers (post-doctorate to senior level), and 3) other research stakeholders (e.g. ombudspersons, RI officers and trainers); as well as to discuss how the institution should ideally provide 4) RI counseling and advice. Participants were encouraged to look at the inspirations for ideas (Figure 3). To optimize individual ideation (i.e. collect a large breadth of ideas from all participants), we first asked participants to individually think of as many ideas as possible and write them down on the exercise board. To foster an interchange of ideas and experiences, we then facilitated a collective discussion of their ideas. After these discussions, they could vote for the most important idea discussed, allowing for the revision of original ideas based on the interactions within the group. The group was asked to summarize the joint insights at the end of each exercise. Any differences of opinion were highlighted to be discussed in subsequent workshops. The full program for these workshops can be found here: https://osf.io/9bztf/. Following the workshop, we sent a summary of the workshop conclusions to the participants as a member check.
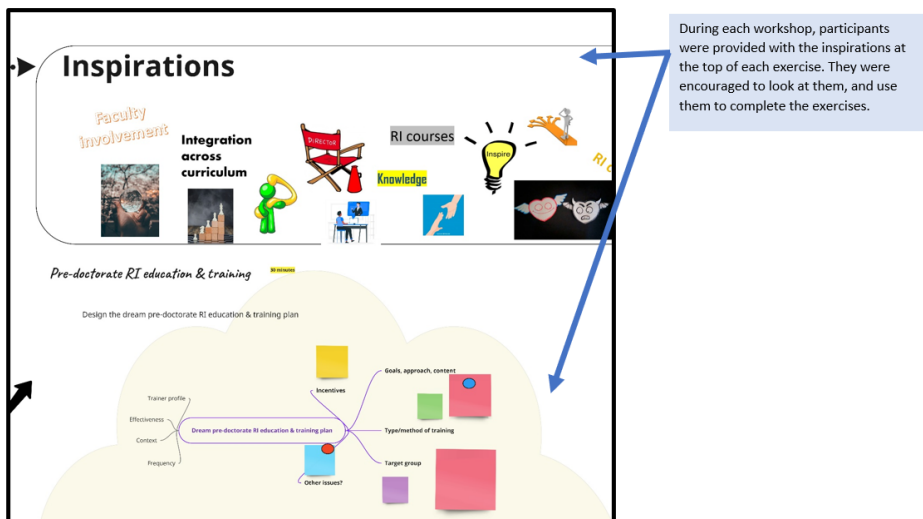
Figure 3 Example exercise used during the content creation workshops.

First round of analysis

We analyzed the data using inductive thematic analysis (4) through an analysis-on-the-wall approach as described by Sanders and Stappers (2012). We used transcripts of the workshops that we generated automatically via Amberscript (5). Two researchers – KL and IL – independently read through the workshop transcripts to identify quotes about the RI education recommendations per topic (i.e. 1) students, 2) researchers, 3) other

RI stakeholders, and 4) RI counseling and advice), and pasted them on a MIRO board for analysis; KL also listened to the recordings of the workshop. The visual outputs of both workshops were also copied onto this MIRO board. The researchers met to conduct an 'analysis workshop', during which they clustered all the data (i.e. quotes and visual outputs) per topic into themes. Each theme was subsequently assigned a label. The analysis results were visualized in an analysis poster per topic. Using the insights from the analysis, i.e. phrasing the themes in prescriptive terms and organizing them in a coherent order, KL developed a first draft of the RI education recommendations per target group, in collaboration with JT. When formulating the recommendations, we made them general enough that any context-related specificities were excluded. Discrepancies or differences in opinion among participants not related to context were highlighted in the guidelines for discussion in the follow up workshops. To ensure that this first draft adequately incorporated insights from the preliminary steps of the guideline development process (3), we added any additional recommendations from the preliminary work not yet discussed in the co-creation workshops to the guidelines, and marked these clearly as not originating from the workshops. The guidance documents were sent to participants of the 'content refinement' workshops a week in advance. In the preparation phase of the 'content refinement' workshops, participants were asked to reflect on how these guidelines would impact their institution.

Content refinement

The 'content refinement' workshops represented the 'convergent' phase of the co-creation process as described by Stelzle, Jannack, and Noenning (6). During these workshops, we focused on refining the RI education recommendations by asking participants to comment on the draft recommendations per topic (i.e. 1) students, 2) researchers, 3) other RI stakeholders, and 4) RI counseling and advice) (example shown in Figure 3). Participants were invited to provide general comments, additions and concerns about the recommendations (e.g. redundancies, gaps, unclarities, conflicting statements, etc.), rather than focus on specific terms and formulations. Additionally, they were invited to provide some best practice examples that could accompany the recommendations, as well as flag potential implementation challenges and opportunities. As in the content generation workshops, we first facilitated individual ideation, followed by a group interchange and building up of ideas, and the group summarized joint insights at the end of each exercise. We asked participants how to deal with any differences of opinion in the guidelines. A summary of the conclusions of the workshops were sent to participants following the workshop, as a member check.

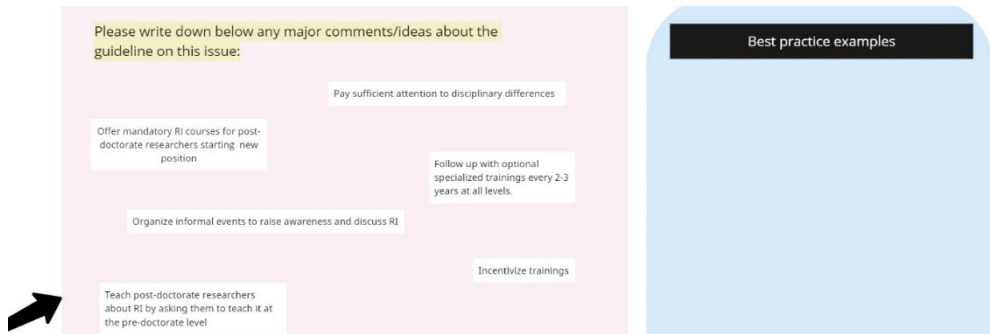**Figure 4** *Example exercise used during the content refinement workshops*

Second round of analysis

Deductive thematic analysis (7) was used for the analysis of the content refinement workshops, using an analysis-on-the-wall approach (8). After reading through the workshop transcripts (genderated automatically by 5) to identify quotes (with KL also listening to the audio recording of the workshops), and pasting these alongside the workshop visual outputs on a MIRO analysis board, KL, IL and NAB conducted another 'analysis workshop'. During this workshop, they deductively clustered the data into themes corresponding to first draft of the RI education recommendations per topic, as well as two additional themes on 'implementation concerns' and 'best practice examples'. In some cases, the themes were slightly altered or expanded upon. Differences between the results from the content creation and content refinement workshops were visualized in analysis posters per topic. Based on the new insights from the content refinement workshops, KL revised the RI education recommendations. The recommendations were kept general enough that any context-related specificities were excluded. In case of any unresolved discrepancies or differences in opinion among participants, we tried to formulate the guidelines in a way that would allow the guideline users to interpret the guideline and tailor it in the way that they preferred. The revised recommendations were sent to all participants for another member check.

Additional analysis

To ensure the rigor of the analysis, RJ checked and made corrections in the automatically generated transcripts as we had deemed them to be ~90-95% correct in earlier steps. The earlier stages of the data analysis were revisited to scrutinize alignment between the overarching themes, subthemes and corresponding quotes (by KL and NE) as in Fereday and Muir-Cochrane (9). Based on this, a detailed code book including the theme and subtheme labels, and detailed descriptions and illustrative quotes for each, was collaboratively developed per target group by KL and NE (https://osf.io/y3c5n/).

References

1.   Pizzolato D, Dierickx K, Tijdink J, Labib K, Lechner I, Bonn NA, et al. D4.4: Report on the Co-Creation Workshops [Internet]. 2021. Available from: https://sops4ri.eu/wp-content/uploads/D4.4_Report-on-the-Co-Creation-Workshops.pdf

2.   Labib K, Pizzolato D, Stappers PJ, Evans N, Lechner I, Widdershoven G, et al. Using co-creation for guideline development - How why and when? OSF Prepr [Internet]. 2021; Available from: https://osf.io/cg3rw/

3.   Lechner I, Tijdink J, Sørensen MP, Ravn T, Bendtsen A-K, Labib K, et al. D4.3: Second version of SOPs and guidelines [Internet]. 2020. Available from: https://sops4ri.eu/wp-content/uploads/D4.3_Second-version-of-SOPs-and-guidelines-1.pdf

4.   Boyatzis RE. Transforming qualitative information: Thematic analysis and code development. Thousand Oaks: SAGE Publications Inc; 1998.

5.   Amberscript. Transform your audio and video to text and subtitles [Internet]. 2021. Available from: https://www.amberscript.com/en/#

6.   Stelzle B, Jannack A, Noennig JR. Co-design and co-decision: Decision making on collaborative design platforms. Procedia Comput Sci. 2017;112:2435–44.

7.   Crabtree BF, Miller WF. A template approach to text analysis: developing and using codebooks. In: Crabtree BF, Miller WF, editors. Doing qualitative research. Newbury Park, CA: SAGE Publications Inc; 1992. p. 163–77.

8.   Sanders E, Stappers PJ. Convivial toolbox: generative research for the front end of design. Amsterdam: BIS Publishers; 2012.

9.   Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. Int J Qual methods. 2006;5(1):80–92.

Appendices II-VII

Appendices II-VII can be found on the Open Science Framework: https://osf.io/gzhd4/

Section 3: Reflecting on implementation

Chapter 7

7. How to combine rules and commitment in fostering research integrity?

Abstract

Research integrity (RI) is crucial for trustworthy research. Rules are important in setting RI standards and improving research practice, but they can lead to increased bureaucracy; without commensurate commitment amongst researchers towards RI, they are unlikely to improve research practices. In this paper, we explore how to combine rules and commitment in fostering RI. Research institutions can govern RI using markets (using incentives), bureaucracies (using rules), and network processes (through commitment and agreements). Based on Habermas' Theory of Communicative Action, we argue that network processes, as part of the lifeworld, can legitimize systems – that is, market or bureaucratic governance modes. This can regulate and support RI practices in an efficient way. Systems can also become dominant and repress consensus processes. Fostering RI requires a balance between network, market and bureaucratic governance modes. We analyze the institutional response to a serious RI case to illustrate how network processes can be combined with bureaucratic rules. Specifically, we analyze how the Science Committee established at Tilburg University in 2012 has navigated different governance modes, resulting in a normatively grounded and efficient approach to fostering RI. Based on this case, we formulate recommendations to research institutions on how to combine rules and commitment.

Key words: research integrity, responsible conduct of research, research misconduct, research governance, bureaucracy, markets, networks, lifeworld, systems

1. Introduction

The replication crisis and prominent cases of research misconduct in the past decade have suggested that research is in crisis (Piper, 2020; Titus & Ballou, 2014). The high prevalence of research misconduct (2-8% according to self-reports) and questionable research practices (QRPs) such as inadequate mentorship and p-hacking (30-50%) jeopardizes not only the quality, relevance, and validity of research findings, but also their trustworthiness (Fanelli, 2009; Gopalakrishna et al., 2022; Xie et al., 2021). In response to these concerns, many institutions across the globe have resorted to policies aimed at fostering research integrity (RI) (Mejlgaard et al., 2020), including the adoption of codes of conduct (e.g., All European Academies, 2017; Resnik & Shamoo, 2011), which are translated at an institutional level into rules and procedures. These steps are crucial for setting clear standards about research practice, providing guidance to researchers, as well as encouraging institutions to more actively support researchers to engage in responsible research practices (Bouter, 2020; Mejlgaard et al., 2020).

Rules, however, typically add additional burdens on researchers who are already navigating high workloads, competition, and stress, and these pressures in research may lead researchers to 'cut corners' and engage in QRPs (Haven et al., 2019; Labib,

Evans, et al., 2021; Sørensen et al., 2021). Furthermore, rule-based approaches to fostering RI might lead to a 'check-box' mentality where researchers are more concerned about complying with externally imposed rules and requirements, rather than being intrinsically motivated to reflect on how to engage in research responsibly (Hemminki, 2016; Labib, Roje, et al., 2021). If researchers are not internally motivated to engage in responsible research practices, it is unlikely that they will commit to these practices or that institutional policies will have the desired effect (Davies, 2019). In other words, efforts to foster RI are much needed, but when they focus only on imposing rules, they can be potentially counterproductive. While this challenge has been widely discussed by various RI researchers (e.g. Haven et al., 2019; Labib, Evans, et al., 2021; Sørensen et al., 2021), there is a gap in the literature regarding ideas on how to address it, reflecting on the foundations of rule-based and commitment-based ways of approaching RI, and their mutual relationship. There is a need for a better understanding of how institutions can foster RI through rules, while simultaneously garnering researcher commitment to RI, considering the distinction between various institutional ways of approaching RI issues and concerns.

    To address this gap, it is important to consider that research is inherently a social phenomenon, requiring interaction between individuals who are part of the research community (Anderson et al., 2010); efforts to foster RI should adequately consider the social context of research to cultivate the commitment of researchers to RI. To understand how efforts to foster RI can adequately consider the social element of research, it is helpful to look at three different modes of governance, as described in governance theory, for instance, by Thompson and colleagues (1991): 'markets' (governance through incentives); 'bureaucracies' (governance through rules, also called 'hierarchies'); and 'networks' (cooperative governance). In practice, these modes of governance are often implemented in combination, and do not occur in isolation (Thompson et al., 1991), also in the context of research. We propose that fostering RI will require an appropriate combination of these different governance modes. Habermas' Theory of Communicative Action, with its core concepts of lifeworld and systems, is a theoretical framework which can provide an explanation of how these modes of governance can be combined adequately in institutional efforts to foster RI, resulting in a normatively grounded as well as efficient approach.

    In this paper, we investigate how research institutions should combine rules with researchers' awareness of and commitment to RI. First, we describe the three analytical modes of governance distinguished by theorists such as Thompson et al. (1991). Next, we present Habermas' Theory of Communicative Action and the central concepts lifeworld and system (Habermas, 1981, 1987), in order to reflect on the relationship between the three modes of governance and investigate how they can be combined in a normatively adequate way. We then present a case analysis to see how our theoretical considerations relate to the practice of RI governance. We explore how the Science Committee of the Tilburg School of Social and Behavioral Sciences (TSB) has navigated between different modes of governance, combining lifeworld and system elements. Based on the insights gained from this case analysis, we end with recommendations to research institutions on how to combine rules and commitment in fostering RI.

2. Governance in research

As discussed by governance theorists (e.g., Pahl-Wostl, 2019; Powell, 1990; Thompson et al., 1991), social life consists of the interaction of various individuals and groups. Organizations are responsible for governing these interactions to ensure they contribute towards the goals of the organization. Three modes of governing social life – markets, bureaucracies (or hierarchies), and networks – have been identified by governance theorists (e.g., Pahl-Wostl, 2019; Powell, 1990; Thompson et al., 1991). In governance theory, market refers to an 'automatic' and 'unconscious' way of coordinating social life, which is driven by the self-interest of individuals and groups who are acting and interacting under the guiding hand of market mechanisms (e.g., incentives, demands and prices present in the system). Alternatively, in bureaucracies or hierarchies, social life is coordinated by formal procedures and rules aimed at consciously regulating the interaction between agents. In contrast to this, networks are characterized by governance theorists as the cooperation of equal, rather than hierarchically organized, agents who steer themselves cooperatively. Applying markets, bureaucracies, and networks to the field of RI can be fruitful in understanding the approaches institutions can take to foster RI.

Market, bureaucratic, and network mechanisms can function both outside and within research institutions. Research institutions do not function in isolation but have a complex relationship with other research stakeholders, and the internal governance strategies of the institution are partially determined by external factors. The social practices of research institutions can be externally influenced by market, hierarchical, and network mechanisms of publishers, funders, and governments. It is important to note that especially publishers' and funders' incentive structures are thought to be important in influencing RI practices at the research institution (Bouter, 2018; Titus & Bosch, 2010). This is because the number and impact of publications are important for obtaining funding and furthering careers, and there is a limited availability of funding for research, which triggers high competition.

Markets, bureaucracies, and networks also govern research practice and influence RI within research institutions. Table 1 is adapted from Pahl-Wostl (2019), Powell (1990) and Thompson et al. (1991), and provides the application of the three modes of governance to research institutions. In this context, market governance can be seen as regulating research through incentives, for example, for hiring, evaluating, and promoting researchers. The incentives used by the research institution will likely mirror incentives imposed by journals and funders. For instance, within the current system of research, many institutions consider the number and impact of publications, as well as prior success in obtaining funding as important criteria for hiring, evaluating, and promoting researchers (Aubert Bonn & Pinxten, 2021). Thus, output indicators (e.g., impact factors, Hirsch (H) indices, success in obtaining grants) are used to evaluate research performance. In a market governance mode, researchers are free to operate as they wish to achieve their goals, but those who do not perform well in relation to the measured incentive criteria will likely be forced out of the system, meaning they will not be hired or promoted. This mechanism pushes all researchers to behave in specific ways that increase their chances of 'surviving in research'. This might induce haste resulting in inaccurate work, for instance, focusing on positive outcomes

to publish in top tier journals, rather than focusing on responsible research practices simply because there is no incentive to focus on this (Sijtsma, 2016); hence the 'publish or perish' analogy in the current research system (Grimes et al., 2018). While individuals in market governance structures may be collegial and want to cooperate with each other, these motives do not primarily function as drivers for coordination of actual practice. Coordination is driven, instead, primarily through market incentives and strategic actions, and the climate is competitive.

Bureaucratic governance of RI by research institutions, on the other hand, is rooted in rules and procedures that are mandated by a governing body. There are clear lines of authority and organizational structures to create a formalized research environment. As such, researchers are required to comply with specific rules and regulations, for instance, regarding the ethical or methodological requirements of study protocols, or data management (e.g., how and where data should be collected, stored and archived). Bureaucracies provide a consciously created, procedural, and organizational system, and a chain of command. Monitoring for compliance with rules and regulations is essential in evaluating the functioning of this type of research governance.

In the third mode of research governance – networks – researchers cooperatively govern themselves at the group level by forming mutual agreements. Researchers' actions are based on and oriented towards their relationships of trust and solidarity with others in the network. They engage with each other in reciprocal and mutually beneficial relationships. Contributing to and being supported by the community is a key motivation for individuals in the network. Network governance require processes of negotiation and may be less efficient than markets and hierarchies; yet it fosters commitment and engagement. It is generally accepted that markets, bureaucracies, and networks as modes of governance do not mutually exclude each other and usually occur in hybrid forms (Pahl-Wostl, 2019; Thompson et al., 1991). However, it is not clear yet how the three modes of governance influence each other and can be combined to optimally foster RI; exploring this is crucial to understanding how RI rules can be combined with researcher commitment to RI.

Table 1: Key features of different research integrity governance modes

Mode of governance	Market	Hierarchy	Network
Motives of researchers	Obtaining funding Scoring highly on output indicators (e.g. number of publications in high impact-factor journals, number of grants obtained)	Following the rules of the research institution	Belonging to the network and contributing to the goals of the network
Factors limiting actions	Self-interest	Rules	Relationship with others in the network
Drivers of behavior	Incentives	Formal authority	Mutual trust and solidarity
Who governs?	Journals and funders *(since incentives are externally formulated)*	Governing body	All involved in the network
Conflict resolution	Those who obtain incentivized achievements win	Formal procedures	Aiming for consensus
Monitoring & evaluation	Focused on output indicators	Compliance with rules	Participatory, joint reflection on agreed goals
Tone/climate	Competitive	Formal	Cooperative

3. Combining modes of governance: relations between systems and lifeworld

Markets, bureaucracies, and networks provide us with an analytical distinction between different approaches that institutions can use to foster RI. To obtain an in-depth understanding of how the modes of governance can be optimally combined to foster RI, focusing on their normative foundations, we will use Habermas' Theory of Communicative Action (1981, 1987).

3.1 Communicative and strategic action

Habermas distinguishes between two types of interactions in social practices: communicative action occurs when two or more actors are involved in an interaction that is oriented towards understanding and achieving mutual goals (Habermas, 1998b), while strategic actions are characterized by interactions in which each actor aims to reach their own individual goals (Habermas, 1998a). Communicative action is oriented towards consensus. This process entails the exchange of validity claims. Each speaker

in the interaction presents validity claims that each listener can accept or refute (Habermas, 1981 pp. 70-72, 1998a). More specifically, each speaker makes three claims, as their statement can be assessed regarding (1) truth (i.e., is what the speaker is saying empirically correct?); (2) normative rightness (i.e., is the speaker's claim in line with what is right according to shared values and standards?), and (3) truthfulness (i.e., is the speaker genuine and sincere?) (Habermas, 1981 pp. 70-72, 1998a). For instance, if speaker A says: "Good supervision is very important in research", then speaker B can question the claims of (1) truth (e.g., Does the evidence about supervision and its relationship to research practices show this?), (2) normative rightness (e.g., Is good supervision an important value in our research community? Are researchers entitled to good supervision?), and (3) truthfulness (e.g., Is the speaker genuine with this statement or do they just want people to pay them for the supervision training they have on offer?). If all three claims are accepted, communicative action can continue; if one or more claims are refuted, the speaker will have to support them with reasons or give them up. As such, mutual understanding is built through a process of accepting and rejecting validity claims between two or more actors. Every interaction requiring consensus building and the development of mutual understanding requires the exchange of these three validity claims, although actors in communication do not need to consciously or explicitly reflect on each validity claim separately; instead, much of the process will take place implicitly.

Because this consensus building process is time-consuming, not every interaction that takes place between actors, particularly when it needs to be fast and efficient, can take place through communicative action (Habermas, 1987 pp. 153-189). Sometimes the three validity claims need to be relaxed to allow actors to engage in interactions, which allow them to accomplish goals that do not require reaching mutual understanding. For instance, a faculty head can communicate a policy to the faculty requiring incoming faculty supervisors to complete a supervision training. In this case, reaching mutual understanding between the faculty head and new supervisors is not the goal and there is no need to go through a consensus building process; the goal is rather to ensure that new supervisors are being trained. Actions in which validity claims are relaxed and in which reaching mutual understanding is not a goal but are rather focused on achieving the outcomes of actors in an efficient and fast process can be referred to as 'strategic actions' (Habermas, 1981 pp. 285-295).

3.2 Applying Habermas' Theory of Communicative Action to modes of governance and their mutual relationships

Social practices consist of both communicative and strategic actions (Habermas, 1987 pp. 153-189). Communicative actions are characteristic of the types of interactions that take place in what Habermas refers to as the lifeworld (Habermas, 1987 pp. 113-153). The lifeworld consists of the background resources and contexts in which communicative action takes place, as can be seen in Table 2 (Habermas, 1987 pp. 113-153). As such, network governance would fall within the lifeworld since network governance operates through communicative action. The lifeworld can be contrasted with markets and bureaucracies, which Habermas calls systems. These are modes of interaction among actors in which the three validity claims of truth, normative

rightness, and truthfulness need not be tested, as certain assumptions are taken for granted and a consensus building process is not carried out (Bohman & Rehg, 2017; Habermas, 1987 pp. 310-331). Social relations in markets and bureaucracies are regulated primarily by money and power, respectively (Habermas, 1987 pp. 154). Since incentives and rules are mechanisms by which money and power operate (Habermas, 1987 p. 159, 307), Habermas' concepts of markets and bureaucracies are comparable to the market and bureaucracy (or hierarchy) modes of RI governance outlined in Table 1. Strategic, rather than communicative, actions are dominant in markets and bureaucracies (Habermas, 1987 pp. 153-189). According to Habermas, both lifeworld and systems are necessary for today's social practices (Habermas, 1987 pp. 153-189, 310-331).

According to Habermas, the agreements that are reached in the lifeworld form the basis for and legitimize the systems (Habermas, 1987 pp. 153-189; 1996a pp. 408-409). For instance, bureaucratic decisions such as whether to mandate training for supervisors, should be rooted in prior lifeworld communicative action processes. One may think of consensus building among the research community about the importance of supervision training for research. Similarly, market mechanisms such as incentives should preferably be based on agreements made in the lifeworld about what constitutes good research (e.g., quantity or quality of publications). Networks, as part of the lifeworld are thus needed to legitimize markets and bureaucracies, while markets and bureaucracies allow for an organized and efficient approach to coordinating actions in a complex environment such as that of research (Habermas, 1987 pp. 153-189).

However, an imbalance can occur between systems and the lifeworld, as systems may colonize the lifeworld; in such a case, strategic actions in markets and bureaucracies impinge on the lifeworld and replace communicative action (Habermas, 1987 pp. 310-331). In the context of research, for instance, market mechanisms would colonize the lifeworld when research incentives focused on a high quantity of publications impede on communicative processes in the research community about what entails good research as well as hinder RI. Similarly, strict regulation of supervision agreements could add pressure and paperwork on researchers' already busy schedules, and prevent supervisors and supervisees from engaging in communicative processes about supervision, and reduce the quality of supervision offered.

When the lifeworld is colonized by a system, countervailing power is needed to act against the interference of the system in the lifeworld (Habermas, 1996b). Countervailing power comes from members of the community coming together and becoming actively involved in a deliberative and participatory process aimed at limiting the interference of the system at stake (Habermas, 1996b). For example, if mandatory supervision training becomes a bureaucratic burden that adds little value to the research community's conception of good research, counterforces could stimulate the research community at the department, the faculty, or the university working together to change the training offered so that it addresses supervisors' needs. This could mean discussing with supervisors the issues and problems for which they require support, as well as what format this support could have, and developing training programs together with supervisors, which directly address their concerns. Likewise, when

market mechanisms such as incentives to publish in high impact journals interfere in the effort of producing high quality studies, members of the research community could work together to change the incentives of research to privilege higher quality research. This could involve developing networks within or between institutions - perhaps something akin to the UK Reproducibility Networks Local Network Leads (Local Network Leads, n.d.) – who together with researchers create new standards for how funders, journals, or institutions can assess who to hire, promote, fund, or publish, which then provide better research incentives (Aubert Bonn & Bouter, 2021).

Applying these insights to the question of how to adequately combine rules and commitment in fostering RI, we argue that institutional RI policies should be anchored in agreements formed among networks of researchers engaged in communicative action, and that deliberative participatory countervailing forces are needed to counter any existing or potential colonization of lifeworld processes in the research community by markets and bureaucracies. As such, rules used to foster RI can be considered as justified and necessary, and garner the commitment of researchers and other stakeholders, so long as they are rooted in network processes, and tendencies towards colonization are recognized and acted against.

Table 2: Key features of 'lifeworld' and 'systems'

Domain of the social world →	Lifeworld	Systems	
		Market	**Bureaucracy**
Characteristics	The domain in which actors cooperate with each other based on mutual understanding	Pre-defined modes of coordination among actors, in which consensus building is not necessary, to allow for efficient interactions	
Types of interactions	Dominated by communicative actions	Dominated by strategic actions	
Regulated primarily by	Consensus building and mutual understanding	Money	Power
Relation to modes of governance	Networks fall within the lifeworld	Comparable to the 'market' mode of governance	Comparable to the 'bureaucracy' (or 'hierarchy') mode of governance
Relationship with other domains	The agreements formed in the lifeworld legitimize systems	Can impede the lifeworld, by replacing and taking over lifeworld processes. This is also referred to as the 'colonization of the lifeworld'.	

4. Case: Tilburg Science Committee

What does it mean to anchor market or bureaucratic mechanisms in network processes and how can this play out in the practice of fostering RI? To answer this question, we

present a case focusing on how the Science Committee at the Tilburg School of Social and Behavioral Sciences (TSB) navigated between different governance modes after a highly publicized case of research misconduct was discovered at the institution. This caused a crisis in the thinking about whether research should be monitored to at least some degree. The various internal and external pressures the misconduct case elicited called for immediate action and learning on the fly to accommodate the bureaucratic governance model to the community's preparedness to allow RI rules. The case description is based on a document analysis of the yearly internal evaluation reports of the Science Committee from 2013-2020, as well as interviews with the former and the current chair of the Science Committee. Obtaining information about the case via these sources allowed for triangulation.

4.1 Case description

In 2011, it was discovered that the dean of the TSB – Diederik Stapel – had engaged in research misconduct, especially falsification and fabrication of data as well as other forms of misconduct. The report which contained the results of the investigation into the misconduct case, concluded that Stapel was solely responsible (Levelt Committee et al., 2012). However, the report also highlighted that the prevailing research culture at the time might not have promoted openness and transparency about data sharing, which could explain why Stapel's misconduct was not discovered earlier. The new dean of the TSB, eager to address the RI concerns established a committee titled the 'Science Committee', consisting of researchers from different departments at the TSB with sufficient seniority and knowledge about good research practices , with the goal to improve the culture of RI at the TSB. The exact tasks that the committee would be responsible for were left open to make a quick start and learn on the fly while designing rules of RI that would be deemed acceptable by the research community.

Upon formation, the Science Committee decided not to impose RI policies on the school in a top-down approach, but instead contacted all department heads at the TSB, to cooperatively explore appropriate approaches to improve RI at the school. In the first two years since the establishment of the committee, the main activity of the committee was to engage in communicative processes with the researchers at the various departments of the TSB to explore how the Science Committee could help foster RI at the school. During these communicative processes, the Science Committee learned that some departments at the school, such as the one where the case of misconduct originated, were eager to introduce measures to improve RI practices. Other departments were less eager to implement measures since they considered their type of research to be vastly different from Stapel's and doubted whether any RI policies would be helpful for their research practice.

Based on these conversations, together with the departments, the members of the Science Committee realized that setting up a future-oriented policy for RI at the school involving a variety of different measures in place to make sure that research would be conducted responsibly from the outset (including a data management steward, a privacy officer, a research ethics committee, an open-science framework), would take time and might be met with resistance from researchers perceiving such measures as the school distrusting their research activities. Therefore, it was decided

that the Science Committee should start its activities by using a hindsight approach, retrospectively evaluating data practices as one important dimension of RI at the school through audits of published studies, rather than checking these prospectively and providing advice on all RI matters from the start. The idea here was to first learn from the past about one important RI dimension (i.e., about data management practices from existing publications) and then implement new RI policy based on past learning and other sources in small incremental steps forward.

The Science Committee worked with representatives and volunteers in the various departments to develop guidelines on Data Handling and Methods Reporting (DHMR) at the school, as well as an auditing procedure to evaluate the data of studies that had recently been published at the TSB. Thus, the various departments at the school were actively involved in both the decision process that led to the agreement to create the DHMR and conduct the audits, but also in the actual development of both the DHMR and the audit procedure. The DHMR provided guidance to researchers about the data management standards expected at the audit. The audit procedure entailed a member of the Science Committee meeting with an author of a recently published article from the school to jointly reflect on how the data had been managed and stored for the study at hand and discuss any perceived problems together, with the help of a checklist which served as a tool to guide the discussion (Appendix I). The audits aimed to provide a learning opportunity during which researchers could obtain advice about their practices and ask questions. Although the specific focus of the audits was on discussing data management of studies, the Committee used the audits more broadly to discuss other aspects of RI with researchers as well, such as issues related to supervision and collaboration. Additionally, the Science Committee members would reflect on the insights gained during the audits in yearly evaluation meetings and form recommendations for adaptations for future years, with the goal to improve the DHMR and audit procedure, and as input for additional future RI policies. The studies for auditing were randomly selected, so only a random sample of the publications at the school were audited. For example, in 2000, 14 out of a total of 716 publications at the school were sampled and audited.

The audit procedure was first piloted among a group of volunteer researchers who then provided feedback about the procedures, including the checklist for questions asked during the audit. The Committee revised both the DHMR and the audit procedure based on the comments and feedback from the various departments and the researchers who participated in the pilot. The first official audits were launched in 2015. This reflected both the careful but also somewhat long process of getting a new governance policy in place even if it was fueled by a drastic event that shook many people's confidence in science. The Committee has since held annual meetings to evaluate the results of the audits and reflect on its own functioning to revise and update the auditing procedures. For instance, in the early years of the auditing procedure, it became apparent that papers were often only published after the PhD student who was often the first author had left the institution, making it difficult to conduct the audit together with the former PhD student. Therefore, after consultation with the departments, the TSB now requires PhD students to share data with supervisors prior to submitting their thesis, to ensure that data remain available at the TSB. This is not only relevant for the audits, but also for the RI responsibility of the

institution and supervisors more broadly in terms of having complete information about the data for each study.

The Science Committee has continued its auditing activities at the TSB and has become an integral part of the school's research policy. The annual reports of the audits show that researchers are becoming increasingly familiarized with good data practices, although their implementation of these practices is not always perfect. For instance, the report from 2020 showed that all 14 researchers audited had stored their data package in a secure location, but one of them had only made the data accessible to one person instead of two as required by the DHMR. The insights gained during the audits conducted by the Science Committee throughout the years have contributed not only to better data handling practices at the TSB, but also to combining rules and communication in the development of additional valuable forward-looking RI structures focused on other aspects of RI, which include a research ethics committee, RI training for researchers, and assignment of a privacy officer that guides and advises researchers on how to deal with privacy issues in the research process. These additional structures are now also present in other schools at Tilburg University (2022). These additional structures, like the Science Committee, have a dual function of both educating researchers (e.g., about various ethics considerations that can arise in research) and improving practice, as well as creating bureaucratic control of research practices (e.g., the research ethics committee, not to be confused with the Science Committee, checks whether ethics requirements have been met).

4.2 Case analysis

This case provides an example of how institutions can implement RI policies which are legitimized and supported by communicative, network processes. Prior to the installment of the Science Committee at the TSB, there was little discussion about RI and only few rules existed (Levelt Committee et al., 2012). The dominant governance approach could be said to have been a market one, where researchers' behavior was mostly guided by the incentives present in the system of research, such as publishing many articles in high-impact journals and obtaining funding. However, the Science Committee started processes of communication as a basis for introducing rules at the TSB on data handling. These network processes included (1) cooperatively developing RI rules and procedures at the school with researchers from all departments rather than implementing these in a hierarchical and top-down approach, as well as (2) organizing audits to continuously communicate with researchers, to educate them about data handling, and reflect with them in order to improve the rules in place based on the insights gained during the audits.

The goal of the Science Committee's audits is not focused on compliance to the DHMR and imposition of rules, but rather on learning and improving the research culture more generally. Some Science Committee members therefore find the term 'audit' somewhat misleading because they see the audits as a 'conversation' or 'interview' with researchers about what entails high quality research and how to achieve it, rather than a compliance-focused bureaucratic procedure. Regardless of the terminology chosen, the aim of the audits is to focus on learning and improving research practice. The auditors have a list of questions (Appendix I) they send to

researchers before interviews that asks them to prepare to explain how they store their data. The questions serve as a starting point for communication with researchers and discussion of data practices. Instead of criticizing or penalizing researchers for not storing data perfectly, the Science Committee members treat the audits as a moment for reflection and learning about better data handling and RI practices. The committee evaluates the results of the audits to continuously adapt and improve the support provided by the Science Committee to the TSB researchers, as well as to provide input to future RI policies at the school. For instance, after the audits in 2020, the Science Committee discussed that the audits revealed that researchers in some departments receive more support regarding structuring their data management than in other departments. Based on this reflection, they recommended the TBS make an inventory of the support structures provided by various departments and help to create more coherence between the support offered to researchers across departments. As such, the insights gained during the audits are used by the Science Committee and the TSB more broadly to understand the RI issues at the institution with a view to be better equipped to support researchers with responsible practices, rather than to police practices at the level of individual researchers, unless very serious violations of the DHRM are committed, such as losing data. Thus, although the audit procedure is not 'power free', in the sense that the auditors are able to detect serious violations of the DHMR by audited researchers, and flag those, it is more in line with what Power (2000) would describe a "process which harnesses productive learning and self-help" rather than "adherence to performance measures which serve the audit process and little else".

The resolution to start a Science Committee was initiated by the school's management team and approved by the departments and the TSB council representing both faculty and students. The Science Committee decided to involve all the departments at the school in cooperatively developing policies for RI, rather than imposing policies in a top-down manner. The decision about what the main tasks of the Committee would entail, – that is, conducting audits focused on data management – was made jointly by the Science Committee and the departments at the TSB after a year of consensus building and deliberation. Furthermore, the rules set by the Science Committee on data handling were developed cooperatively with the research community, which contributed to their legitimation. As such, while the Science Committee itself is a formal (and hence bureaucratic) body within the TSB, it continuously uses network processes of deliberation and consensus building together with researchers to develop and legitimize its bureaucratic elements.

The Science Committee's approach of using audits to understand and improve research practices within the school rather than police individual researchers has been especially helpful to the TSB since the introduction of the European General Data Protection Regulation (GDPR) (General Data Protection Regulation, 2016). Because the Science Committee members were already familiar with data practices and challenges at the TSB before the GDPR was implemented, they were in a good position to advise the institution how to help navigate the GDPR and relate it to standards of good research practices. For example, the Science Committee members knew that most of the data sets collected and stored at the TSB were already anonymized and therefore not sensitive. Thus, they advised the legal and policy staff at the TSB to only focus on

those few studies at the school which included sensitive data when doing privacy assessments. In this way, based on knowledge obtained during conversations with researchers in audits, the Science Committee helped to prevent introducing rules when they were unnecessary at Tilburg University because of the GDPR, thus countering the potential of colonization of the research lifeworld by bureaucratic procedures.

5. Discussion and recommendations

5.1 Recommendations

In this article, we explored how institutions can combine rules and commitment in fostering RI. We discussed that approaches to RI governance can be roughly categorized into three modes, including markets (governing through incentives), bureaucracies (governing through rules), and networks (cooperative governance). Some of the causes of RI problems likely have to do with perverse incentives in market governance, and bureaucratic governance such as the implementation of RI rules is often used to remedy these. However, fostering RI requires appropriate combinations of the three governance modes because the modes each have their own strengths and weaknesses. For instance, while the network mode is more collegial and collaborative, it tends to be slower and influenced by the dynamics of the group compared to market and bureaucratic modes of governance.

 We used Habermas' Theory of Communicative Action to discuss how to optimally combine these modes of governance. Namely, we argued that market or bureaucratic mechanisms to foster RI should be rooted in lifeworld processes, that is, in network mechanisms. In other words, our central claim in the paper is that any rules used to foster RI need justification and support by network processes – communicative, consensus-building action – among the research community within and outside of the institution. Furthermore, action is needed to counter tendencies of markets and bureaucracies to interfere in network processes in the research community.

 We discussed how the Science Committee at the Tilburg University's School of Social and Behavioral Science (TSB) provided an example of how this can be done. The Science Committee was established in response to a case of misconduct at the TSB, which likely occurred due to the existence of perverse incentives in the research system as well as limited rules about RI; therefore the market mode of governance was dominant. To counter this, the Science Committee created rules on RI together with the research community at the TSB. The introduction of bureaucratic RI governance (i.e. rules), was legitimized through network processes of consensus building and agreement among the research community. This allowed to foster RI at the school, by combining RI rules with the concomitant researchers' commitment to RI. Based on the insights discussed in this paper, we formulate several recommendations for research institutions on how to combine rules and commitment in fostering RI.

R1. Involve researchers in the development of RI rules

To anchor market and bureaucratic mechanisms in network processes, institutions could involve the research community to cooperatively develop RI rules (Labib, Roje, et al., 2021; Mejlgaard et al., 2020). As an example, we have argued that the TSB's Science Committee has improved RI practices at the TSB by developing rules and procedures about data handling, using deliberation and consensus building with researchers at the various departments of the school. Degn's (2020) work suggesting that the top-down implementation of institutional policies does not influence researchers' approach to RI, supports the idea that network-oriented communicative processes among the research community are necessary for the implementation of policies that the research community owns and commits to. It could be helpful for institutions to involve researchers in the development of rules in a systematic manner, using co-creation or other participatory methods (e.g., see as we have described in Labib, Pizzolato, et al., 2021). Such a systematic approach is important to include all relevant voices in a fair way when developing RI policy, especially to prevent misuse of networks, as group processes have been used in research governance in the past for purposes such as eliminating competition, gatekeeping others' ideas, and promoting self (Sahlin & Eriksson-Zetterquist, 2016).

R2. Approach RI rules and regulations to further the research community's goals, rather than as ends in themselves which need stringent enforcement

Using bureaucracy to further the community's shared goals can help the research community to become intrinsically motivated to engage in responsible research practices, rather than doing so because of potential negative sanctioning, and might thereby prevent a 'check-box' mentality (Labib, Roje, et al., 2021). For instance, audits can serve as a valuable communicative tool for learning, exchanging knowledge, and exploring mutual goals, rather than policing (Gerritsen et al., 2021). At the TSB, the audits are focused on education and reflecting on data handling and RI more generally, and this has helped to increase awareness of RI. Also, unnecessary rules are prevented, countering the tendency of the bureaucratic system to colonize the lifeworld. Bureaucracies that become ends in themselves could lead to a policing culture at the institution where researchers are under pressure to comply with externally imposed rules that they might not believe in (DuBois, 2004). Of course, that is not to say that serious misbehavior should not be addressed appropriately if it arises, as dealing with misconduct and its negative consequences appropriately is also in the interests of the community (Fanelli et al., 2015; Labib, Roje, et al., 2021).

R3. Implement RI policies gradually to allow sufficient time to anchor bureaucracies in network processes, but also use momentum when there is increased attention for RI

Implementation of new policies and initiatives can take time, not only due to procedural issues, but also due to the need for reflection on the values behind the policies and their relationship to community goals and culture (Gerritsen et al., 2021; Vasconcelos et al., 2015). Changing attitudes towards RI and forming agreements with stakeholders across the network about what measures to take can be a slow process

(Coates, 2014). Institutions will likely benefit from gradual implementation of new policies, to allow for sufficient time to foster network processes to create commitment of researchers to these policies (Adams et al., 2014; Horbach & Halffman, 2018; Vasconcelos et al., 2015). However, the need to carefully define and reflect on RI criteria does not exclude the possibility in situations, such as the occurrence of misconduct cases, for policy makers to make top-down decisions about introducing RI policy. After such decisions, the process of implementation requires time and inclusion of stakeholders as can be seen in the case of the TSB, in which the initiative to install a Science Committee to foster RI was taken by the management team in a top-down manner after which the exact tasks and responsibilities of the Committee were developed together with researchers from the departments. Indeed, we see that cases of misconduct elsewhere have also led to important new policies and changes aimed at fostering RI (Jones, 2003; Opel et al., 2011; Resnik, 2014). This suggests that the shock caused by the discovery of cases of misconduct in the research community may be an important catalyst in legitimizing the introduction of new RI policies aimed at preventing such cases in the future, but participation of researchers through network processes is also needed for legitimation.

R4. Legitimize rules through network processes continuously rather than only in the initial implementation stage, for updating and improvement of RI policies

Legitimization of rules requires the involvement of the affected network of researchers not only in the initial creation process of bureaucracies, but also in regular evaluations and updating (Pires, 2011; Weller, 2020). This is important to explore whether the rules serve community goals in practice or if any adjustments are needed, and to reevaluate community goals (Pires, 2011; Weller, 2020). The latter might especially be important when new developments in research, such as the introduction of new regulations (e.g., the GDPR) or standards (e.g., related to open science) emerge. Continuous adaptation and legitimization are especially important in serving as a countervailing force against any potential colonization of communicative processes by market and bureaucratic processes. The TSB case shows the need for continuity of audits to reflect on the basic principles of RI and their application in data management; in addition, policies concerning data management should be related to a wider discussion of the importance of RI for research practice.

R5. Use RI policies to foster awareness and good practice of individual researchers, address challenges and opportunities for RI at the institutional level, and to relate challenges and opportunities to developments in the larger research community

Using RI policies to promote responsible research practice at an individual level, as well as address RI challenges and opportunities at an institutional level as done by the TSB Science Committee, can prevent researchers from experiencing rules as externally imposed inconveniences that they have to obey (Labib, Roje, et al., 2021; Landi et al., 2015; Mejlgaard et al., 2020). Such an approach to organizational culture where the emphasis lies on learning from challenges and mistakes rather than punishing individual persons has been termed as 'just culture' and described as important in increasing the safety and quality of the work at the organization (Dekker, 2009; Khatri et al., 2009). Of

course, this should be balanced with making individuals take accountability for their actions when they have engaged in unacceptable behavior (Dekker, 2009). Additionally, relating the policies to developments in the larger researcher community – for instance by sharing experiences with other institutions (e.g., Sijtsma, 2017) – can be valuable in exchanging insights and joint learning, and therefore in legitimizing RI policies on a broader scale.

5.2 Strengths and limitations

In this study, we combined a theoretical perspective with a case analysis to reflect on how to combine rules with commitment in fostering RI. Our theoretical perspective brought together the conceptual approach of governance theory with Habermas' analysis of societal development, including a foundational analysis and normative view concerning the relationship between lifeworld and systems. Our case analysis, informed by a document analysis and interviews with the former and current Science Committee chairs at the TSB, provided a real-life example of this. The chairs were well informed about the functioning of the Science Committee and likely had a more comprehensive overview than others, including regular Committee members.

Our case analysis is limited to the perspectives of the past and present TSB Science Committee chairs and the yearly evaluation reports of the Science Committee. Although the evaluation reports were intended to be used for internal purposes and only shared with us after the fact, it could be that some of the information in them is limited due to strategic considerations. In addition, the chairs might not be able to recall all prior experiences at the Science Committee. We do not have direct information about researchers' experiences in dealing with the committee, which could provide additional insights on the commitment of TSB's researchers toward RI and if they did indeed experience audits as a form of communicative action. Furthermore, while this case provides important insights into how various governance approaches can be combined to foster RI, it mostly focused on bureaucracies and networks, whereas reshaping market governance can also be valuable in fostering RI, particularly considering that perverse incentives in research are seen as a factor contributing to research misconduct and QRPs (Edwards & Roy, 2017). In fact, there are many initiatives currently to address incentives in research to reward responsible research practices rather than quantity of publications (see Aubert Bonn & Bouter, 2021 for an overview of initiatives). So, next to achieving an adequate combination of rules and commitment, also the relationship between market mechanisms and deliberative processes around good research needs to be taken into account.

5.3 Conclusion

In this study, we argued that institutions should use network processes to develop and legitimize rules that foster RI in research institutions. As an example, we showed how the Science Committee at Tilburg University's TSB has involved the researchers at the school's departments in the process of creating and applying rules and procedures, specifically, regulations on data management. From the perspective of Habermas' Theory of Communicative Action, the Science Committee appropriately combined

different modes of governance, resulting in a normatively grounded as well as efficient approach to fostering RI. Based on our analysis of the Science Committee, we recommend that institutions involve researchers in developing RI rules, use the implementation of rules to further the research community's goals, implement RI policies gradually, legitimize rules continuously through network processes, and use RI policies both at the individual and the institutional level.

Declarations

Ethics approval

The study protocol was sent to the Institutional Review Board of the Amsterdam University Medical Centers (location VUmc), which decided that the study does not fall under the Dutch Medical Research Involving Human Subjects Act (Wet medisch-wetenschappelijke onderzoek met mensen).

Competing interests

KS was dean of the TSB, appointed after the misconduct case discussed in this paper had occurred. All information provided in this paper relating to the misconduct case is factual and can be checked for correctness upon request from the first author. The authors have no other conflicts of interest to declare.

Funding

Authors' contributions

Drafting of article – KL; Article revision – KL, JT, KS, LB, NE, GW; Supervision – GW

Acknowledgements

Supplementary materials

Appendix I: Checklist used during the audits at the TSB by the Science Committee

	Yes	No	N/A	Unclear
1. Is a data package available? If so, where is it stored?				
Explanation:				
2. Does the data package contain the raw data file?				

	Yes	No	N/A	Unclear
(as first made available, in digital or digitizable form, to the Tilburg School of Social and Behavioral Sciences (TSB) staff member)				
Explanation:				
3. If the data consists of non-digital source material: *Is there an indication of where the data is stored and of how it can be traced?*				
Explanation:				
4. If the TSB researcher is not allowed to have the data in their possession, due to ethical guidelines, insurmountable logistical concerns, privacy-related reasons, or formal property rights: a. Has a clear explanation been given to account for the fact that the data is not present in the package?				
b. Have attempts been made to store relevant details (e.g. a random sample from the original data file or a screenshot of part of the raw data) in the data package?				
Explanation:				
5. With respect to the metadata: a. Does the data package or the article include a description of who collected the data, where, and at what location? *(For non-TSB data, this information should be sufficiently specific – in the form of contact details for the other researchers, and clear descriptions of the source of the data acquired by the researcher and the date on which it was downloaded.)*				
b. Is there a broad description or automatically generated document indicating who created which file, and when?				
c. Does the data package or the article include a description of the role of each author (and co-author)? (e.g., "devising and setting up the study", "data collection", "data analysis", and "writing the article")				
d. When external finances / grants have been obtained: has been specified who provided these external finances / grants?				
Explanation:				
6. If data originates from existing databases: Has its origin, version, and date been specified?				
Explanation:				
7. Has the data in the raw data file been sufficiently anonymized and has the confidentiality of the data been properly safeguarded? *(The data package should contain no information that might be used to identify individuals)*				
Explanation:				
8. Does the data package contain all of the digital (or scannable) research materials needed to allow the data collection to be replicated by a colleague with the requisite				

	Yes	No	N/A	Unclear
skills? *(e.g. questionnaires, stimuli, instructional texts, chief experimenter protocols, video materials, simulation study software, computer scripts, logs)*				
Explanation:				
9. Does the data package contain syntaxes, computer scripts or statistical logs, pertaining to the processing of the raw data, that would allow the analyses to be replicated by a colleague with the requisite skills?				
Explanation:				
10. Is the data package accessible to at least two individuals? *(The second individual could be a co-author, supervisor, co-supervisor, or executive official, for example)*				
11. Has the data package been stored in such a way that it cannot be lost or become corrupted? *(this could involve backups to multiple locations, automatic backups, or safeguards to prevent the data from being accidentally overwritten or deleted)*				
Explanation:				
12. Will the data package be retained and remain accessible for at least ten years after the article's definitive publication date?				
Explanation:				

References

Adams, V., Burke, N. J., & Whitmarsh, I. (2014). Slow research: thoughts for a movement in global health. Medical Anthropology, 33(3), 179–197.

All European Academies. (2017). The European Code of Conduct for Research Integrity. All European Academies. https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

Anderson, M. S., Ronning, E. A., Vries, R. De, & Martinson, B. C. (2010). Extending the Mertonian norms: scientists' subscription to norms of research. The Journal of Higher Education, 81(3), 366–393.

Aubert Bonn, N., & Bouter, L. (2021). Research assessments should recognize responsible research practices—Narrative review of a lively debate and promising developments. In OSF Preprints. https://doi.org/10.31222/osf.io/82rmj

Aubert Bonn, N., & Pinxten, W. (2021). Advancing science or advancing careers? Researchers' opinions on success indicators. PloS One, 16(2), e0243664.

Bohman, J., & Rehg, W. (2017). Jürgen Habermas. The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/habermas/#TheComAct

Bouter, L. (2018). Fostering responsible research practices is a shared responsibility of multiple stakeholders. Journal of Clinical Epidemiology, 96, 143–146.

Bouter, L. (2020). What Research Institutions Can Do to Foster Research Integrity. Science and Engineering Ethics, 1–7.

Coates, H. L. (2014). Ensuring research integrity: The role of data management in current crises. College & Research Libraries News, 75(11), 598–601. https://doi.org/https://doi.org/10.5860/crln.75.11.9224

Davies, S. R. (2019). An ethics of the system: Talking to scientists about research integrity. *Science and Engineering Ethics, 25*(4), 1235–1253.

Degn, L. (2020). Integrating integrity: The organizational translation of policies on research integrity. *Science and Engineering Ethics, 26*(6), 3167–3182.

Dekker, S. W. A. (2009). Just culture: who gets to draw the line? *Cognition, Technology & Work, 11*(3), 177–185.

DuBois, J. M. (2004). Is compliance a professional virtue of researchers? Reflections on promoting the responsible conduct of research. *Ethics & Behavior, 14*(4), 383–395.

Edwards, M. A., & Roy, S. (2017). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science, 34*(1), 51–61.

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS One, 4*(5), e5738.

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.

Gerritsen, S., van Melle, A. L., Zomer, L. J. C., Widdershoven, G. A. M., & Voskes, Y. (2021). Communities of Practice in Acute and Forensic Psychiatry: Lessons Learned and Perceived Effects. *Psychiatric Quarterly, 92*(4), 1581–1594.

Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PloS One, 17*(2), e0263023.

Grimes, D. R., Bauch, C. T., & Ioannidis, J. P. A. (2018). Modelling science trustworthiness under publish or perish pressure. *Royal Society Open Science, 5*(1), 171511.

Habermas, J. (1981). *The Theory of Communicative Action 1 - Reason and the Rationalization of Society* (T. McCarthy, Ed.). Beacon Press.

Habermas, J. (1987). *The Theory of Communicative Action 2 - Lifeworld and System: A Critique of Functionalist Reason* (T. McCarthy, Ed.). Beacon Press.

Habermas, J. (1996a). *Between Facts and Norms - Contributions to a Discourse Theory of Law and Democracy* (W. Rehg, Ed.). MIT Press.

Habermas, J. (1996b). Civil Society and the Political Public Sphere. In W. Rehg (Ed.), *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (pp. 329–387). MIT Press.

Habermas, J. (1998a). Social action, purposive action, and communication. In M. Cooke (Ed.), *On the pragmatics of communication* (pp. 105–183). MIT Press.

Habermas, J. (1998b). Some further clarifications on the concept of communicative rationality. In M. Cooke (Ed.), *On the pragmatics of communication* (pp. 307–342). MIT Press.

Haven, T., Bouter, L., Smulders, Y. M., & Tijdink, J. (2019). Perceived publication pressure in Amsterdam: Survey of all disciplinary fields and academic ranks. *PloS One, 14*(6), e0217931.

Hemminki, E. (2016). Research ethics committees in the regulation of clinical research: comparison of Finland to England, Canada, and the United States. *Health Research Policy and Systems, 14*(1), 1–12.

Horbach, S. P. J. M., & Halffman, W. (2018). The changing forms and expectations of peer review. *Research Integrity and Peer Review, 3*(1), 1–15.

Jones, A. H. (2003). Can authorship policies help prevent scientific misconduct? What role for scientific societies? *Science and Engineering Ethics, 9*(2), 243–256.

Khatri, N., Brown, G. D., & Hicks, L. L. (2009). From a blame culture to a just culture in health care. *Health Care Management Review, 34*(4), 312–322.

Labib, K., Evans, N., Roje, R., Kavouras, P., Reyes Elizondo, A., Kaltenbrunner, W., Buljan, I., Ravn, T., Widdershoven, G., Bouter, L., Charitidis, C., Sørensen, M. P., & Tijdink, J. (2021). Education and training policies for research integrity: Insights from a focus group study. *Science and Public Policy, scab077.* https://doi.org/https://doi.org/10.1093/scipol/scab077

Labib, K., Pizzolato, D., Stappers, P. J., Evans, N., Lechner, I., Widdershoven, G., Bouter, L., Dierickx, K., Bergema, K., & Tijdink, J. (2021). Using co-creation for guideline

development - How why and when? OSF Preprint. https://doi.org/10.31219/osf.io/cg3rw

Labib, K., Roje, R., Bouter, L., Widdershoven, G., Evans, N., Marušić, A., Mokkink, L., & Tijdink, J. (2021). Important topics for fostering research integrity by research performing and research funding organizations: a Delphi consensus study. Science and Engineering Ethics, 27(47), 1–22. https://doi.org/10.1007/s11948-021-00322-9

Landi, M. S., Shriver, A. J., & Mueller, A. (2015). Consideration and checkboxes: incorporating ethics and science into the 3Rs. Journal of the American Association for Laboratory Animal Science, 54(2), 224–230.

Levelt Committee, Noort Committee, & Drenth Committee. (2012). Flawed science: The fraudulent research practices of social psychologist Diederik Stapel.

Local Network Leads. (n.d.). Retrieved July 6, 2022, from https://www.ukrn.org/local-network-leads/

Mejlgaard, N., Bouter, L., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A.-K., Charitidis, C. A., Claesen, N., Dierickx, K., Domaradzka, A., Reyes Elizondo, A., Foeger, N., Hiney, M., Kaltenbrunner, W., Labib, K., Marušić, A., Sørensen, M. P., Ravn, T., Ščepanović, R., … Veltri, G. A. (2020). Research integrity: nine ways to move from talk to walk. Nature Publishing Group.

Opel, D. J., Diekema, D. S., & Marcuse, E. K. (2011). Assuring research integrity in the wake of Wakefield. British Medical Journal, 342(342:d2).

Pahl-Wostl, C. (2019). The role of governance modes and meta-governance in the transformation towards sustainable water governance. Environmental Science & Policy, 91, 6–16.

Piper, K. (2020). Science has been in a "replication crisis" for a decade. Have we learned anything? https://www.vox.com/future-perfect/21504366/science-replication-crisis-peer-review-statistics

Pires, R. R. C. (2011). Beyond the fear of discretion: Flexibility, performance, and accountability in the management of regulatory bureaucracies. Regulation & Governance, 5(1), 43–69.

Powell, W. W. (1990). Neither market nor hierarchy: Network forms of organization. Research in Organizational Behavior, 12, 295–336.

Power, M. (2000). The audit society—Second thoughts. International Journal of Auditing, 4(1), 111–119.

Resnik, D. B. (2014). Scientific misconduct and research integrity. Handbook of Global Bioethics, 799–810.

Resnik, D. B., & Shamoo, A. E. (2011). The Singapore Statement on Research Integrity. Accountability in Research, 18(2), 71–75.

Sahlin, K., & Eriksson-Zetterquist, U. (2016). Collegiality in modern universities–the composition of governance ideals and practices. Nordic Journal of Studies in Educational Policy, 2016(2–3), 33640.

Sijtsma, K. (2016). Playing with data—or how to discourage questionable research practices and stimulate researchers to do things right. Psychometrika, 81(1), 1–15.

Sijtsma, K. (2017). Never waste a good crisis: Towards responsible data management. Keynote Lecture on the 5th World Conference on Research Integrity.

Sørensen, M. P., Ravn, T., Marušić, A., Elizondo, A. R., Kavouras, P., Tijdink, J. K., & Bendtsen, A.-K. (2021). Strengthening research integrity: which topic areas should organisations focus on? Humanities and Social Sciences Communications, 8(1), 1–15.

Thompson, G., Frances, J., Levacic, R., & Mitchell, J. (1991). Markets, hierarchies and networks: the coordination of social life (1st Editio). Sage Publications.

Tilburg University. (2022). Conduct and integrity. https://www.tilburguniversity.edu/about/conduct-and-integrity

Titus, S., & Ballou, J. M. (2014). Ensuring PhD development of responsible conduct of research behaviors: Who's responsible? Science and Engineering Ethics, 20(1), 221–235.

Titus, S., & Bosch, X. (2010). Tie funding to research integrity. Nature, 466(7305), 436–437.

Vasconcelos, S. M. R., Sorenson, M. M., Watanabe, E. H., Foguel, D., & Palácios, M. (2015). Brazilian science and research integrity: Where are we? What next? Anais Da Academia Brasileira de Ciências, 87, 1259–1269.

Weller, A. (2020). Exploring practitioners' meaning of "ethics,""compliance," and "corporate social responsibility" practices: A communities of practice perspective. *Business & Society, 59(3)*, 518–544.

Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of research misconduct and questionable research practices: A systematic review and meta-analysis. *Science and Engineering Ethics, 27(4)*, 1–28.

Levelt Committee, Noort Committee, & Drenth Committee. (2012). Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. https://www.tilburguniversity.edu/sites/default/files/download/Final report Flawed Science_2.pdf

Local Network Leads. (n.d.). Retrieved July 6, 2022, from https://www.ukrn.org/local-network-leads/

Mejlgaard, N., Bouter, L., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A.-K., Charitidis, C. A., Claesen, N., Dierickx, K., Domaradzka, A., Reyes Elizondo, A., Foeger, N., Hiney, M., Kaltenbrunner, W., Labib, K., Marušić, A., Sørensen, M. P., Ravn, T., Ščepanović, R., … Veltri, G. A. (2020). Research integrity: nine ways to move from talk to walk. *Nature, 586*, 358–360.

Opel, D. J., Diekema, D. S., & Marcuse, E. K. (2011). Assuring research integrity in the wake of Wakefield. *British Medical Journal, 342(342:d2)*.

Pahl-Wostl, C. (2019). The role of governance modes and meta-governance in the transformation towards sustainable water governance. *Environmental Science & Policy, 91*, 6–16.

Piper, K. (2020). Science has been in a "replication crisis" for a decade. Have we learned anything? https://www.vox.com/future-perfect/21504366/science-replication-crisis-peer-review-statistics

Pires, R. R. C. (2011). Beyond the fear of discretion: Flexibility, performance, and accountability in the management of regulatory bureaucracies. *Regulation & Governance, 5(1)*, 43–69.

Powell, W. W. (1990). Neither market nor hierarchy: Network forms of organization. *Research in Organizational Behavior, 12*, 295–336.

Resnik, D. B. (2014). Scientific misconduct and research integrity. *Handbook of Global Bioethics*, 799–810.

Resnik, D. B., & Shamoo, A. E. (2011). The Singapore Statement on Research Integrity. *Accountability in Research, 18(2)*, 71–75.

Sijtsma, K. (2017). Never waste a good crisis: Towards responsible data management. Keynote Lecture on the 5th World Conference on Research Integrity. https://wcrif.org/images/2017/documents/2. Tuesday May 30, 2017/1. Aula/K. Sijtsma - Never waste a good crisis - Towards responsible data management.pdf

Sørensen, M. P., Ravn, T., Marušić, A., Elizondo, A. R., Kavouras, P., Tijdink, J. K., & Bendtsen, A.-K. (2021). Strengthening research integrity: which topic areas should organisations focus on? *Humanities and Social Sciences Communications, 8(1)*, 1–15.

Thompson, G., Frances, J., Levacic, R., & Mitchell, J. (1991). Markets, hierarchies and networks: the coordination of social life (1st Editio). Sage Publications.

Titus, S., & Ballou, J. M. (2014). Ensuring PhD development of responsible conduct of research behaviors: Who's responsible? *Science and Engineering Ethics, 20(1)*, 221–235.

Titus, S., & Bosch, X. (2010). Tie funding to research integrity. *Nature, 466(7305)*, 436–437.

Vasconcelos, S. M. R., Sorenson, M. M., Watanabe, E. H., Foguel, D., & Palácios, M. (2015). Brazilian science and research integrity: Where are we? What next? *Anais Da Academia Brasileira de Ciências, 87*, 1259–1269.

Weller, A. (2020). Exploring practitioners' meaning of "ethics,""compliance," and "corporate social responsibility" practices: A communities of practice perspective. *Business & Society, 59(3)*, 518–544.

Xie, Y., Wang, K., & Kong, Y. (2021). Prevalence of research misconduct and questionable research practices: A systematic review and meta-analysis. *Science and Engineering Ethics, 27(4)*, 1–28.

Discussion

Chapter 8

8. General discussion

In this PhD thesis, I[4] aimed to provide guidance to research institutions on fostering research integrity (RI) by exploring the question: ***How can research institutions develop policies to foster, and raise awareness about, RI?*** To do this, I posed six research questions, divided into three sections, namely 1) setting the agenda, 2) developing guidelines, and 3) reflecting on implementation. The research questions, per section, are shown in table 1.

Table 1: Research questions in this PhD thesis

Setting the agenda
1. *What are current practices of RI promotion in research institutions?*
2. *Which topics should be addressed in institutional RI policies?*

Developing guidelines
3. *What are researchers' and other stakeholders' views and preferences regarding how research institutions can develop and implement better RI education and training policies?*
4. *How can RI guidelines be co-created together with lead users?*
5. *What should be included in the RI education and training policies of research institutions?*

Reflecting on implementation
6. *How can research institutions combine the implementation of RI rules with researcher commitment to foster RI?*

This chapter provides a summary of the findings of the thesis, reflects on the main take-away messages, discusses the strengths and limitations of the work, and provides some directions for future research.

Summary of findings

Setting the agenda

The investigation into what are current practices of RI promotion at research institutions showed that while there are already many institutional practices for RI promotion globally, most of them focus on researcher, rather than institutional, responsibilities for fostering RI (*Answer to Research Question, ARQ, 1*). The results also indicated that many current practices are focused on the biomedical sciences, with less attention on research in other disciplinary fields. Regarding which topics should be included in institutional policies, consensus was reached among research policy experts

---

[4] *In order to not confuse the reader by continuously switching between 'I' to denote personal reflections and views and 'we' to denote collaborative work, I will use 'I' throughout the general discussion chapter. However, it should be noted that the research conducted in this PhD trajectory was a collaborative effort, as is reflected in the co-authors that are involved in the content chapters.*

and research leaders on a comprehensive list of 12 topics that are essential for fostering responsible research practices (RRPs) at research institutions *(ARQ2)*. The highest ranked topic among these in terms of importance was 'RI education and training', followed by 'Supervision and mentoring', 'Dealing with breaches of RI', and 'Supporting a responsible research process' (Table 1).

Table 1: Ranked list of topics which institutional policies on research integrity <u>should</u> address

Rank (by importance)	Topic
1	RI education and training
2	Responsible supervision
3	Dealing with breaches of RI
4	Supporting a responsible research process (e.g. through quality assurance)
5	Research ethics procedures
6	Data management
7	Conflicts of interest
8	Research culture
9	Publication and communication
10	Updating and implementing the RI policy
11	Intellectual property issues
12	Collaborative research among institutions

Developing guidelines

To provide insights to research institutions on how to foster RI, I specifically zoomed in on developing guidelines on the topic of RI education. I chose this focus because, despite being one of the most discussed and highly valued approaches to fostering RI, there is little guidance for institutions on how to provide RI education in a way that is sensitive to stakeholders' needs. The results show that researchers and other research stakeholders across disciplinary fields in Europe support the provision of continuous RI education which targets all researchers (across ranks) and other institutional stakeholders *(ARQ 3)*. This was important in informing the guideline development process. I used co-creation methods, a systematic approach to jointly develop guidelines together with stakeholders on RI education. I reflected on the value of using such a systematic approach to joint development of guidelines and concluded that co-creation methods allow for close stakeholder engagement and inclusion of diverse perspectives *(ARQ 4)*. Here, I also elaborated on the various steps of co-creation methods, including preparation, sensitization, workshop organization and facilitation, and analysis and guideline drafting and revision. I then shared the contents of the resulting guidelines on the RI education of a) bachelor, master and PhD students, b) post-doctorate and senior researchers, c) other RI stakeholders; as well as d) continuous RI education *(ARQ 5)*. The guidelines recommend mandatory RI training;

follow-up refresher training; informal discussions about RI; appropriate rewards and incentives for active participation in RI education; and evaluation of RI educational events across target groups.

Reflecting on implementation

The question of how research institutions can combine the implementation of RI rules with researcher commitment to foster RI led to the proposition that institutions can use and combine market (governance through incentives), bureaucracy (governance through rules) and network (cooperative governance in a group) mechanisms to foster RI. Using Habermas' Theory of Communicative Action, I reasoned that institutions can use bureaucratic and market mechanisms (such as rules and incentives, respectively) to foster RI, as long as these are rooted in network processes (i.e. the involvement of stakeholders in the development and improvement of rules or incentives), and that countervailing forces are needed to counter the interference of rules and incentives in network processes *(ARQ 6)*.

Main take-away messages

There are four main lessons that I take-away from the research done in this thesis: 1) Framing matters, 2) Tailoring is essential, 3) Box-checking is a danger, and 4) RI is a journey.

1.   Framing matters

Regardless of which RI stakeholder I talked to as part of this research, all seemed to agree that researchers want to engage in RRPs, but need support from their institutions to do so. At the same time, it was reassuring to see that many institutions across Europe are motivated to put policies in place to support their researchers with RI (even if for some, this motivation may be partially external and stem from the fact that research funders such as the European Commission are increasingly incentivizing institutions to implement RI policies, 1). Yet, the framing of RI can often be off-putting for researchers and institutions. When RI is framed as what 'researchers should not do', or even more positively 'what they have to do' – be it in a RI course setting or institutional guideline for researchers – researchers can feel controlled, intimidated, accused, and interfered with (2 (Chapter 4),3 (Chapter 6)). Approaching and communicating RI as a positive goal that everyone can strive towards (rather than treating it in terms of absolutes, of 'good' versus 'bad' behavior) will likely help institutions to get goodwill and commitment from researchers towards RRPs, as well as reduce questionable research practices (QRPs), and increase insight about how to deal with moral 'gray areas' – situations where there is no clear right or wrong answer – in research (2 (Chapter 4),3 (Chapter 6),4 (Chapter 7)).

The importance of framing is illustrated by the following example: my collaborators and I received negative reactions from not only researchers, but also institutions, to the name of the consortium that this PhD project was part of (Standard Operating Procedures for Research Integrity, SOPs4RI); stakeholders repeatedly told

consortium members that most RI policies cannot be reduced to a set of standard operating procedures (SOPs). Promoting RRPs requires institutions to support researchers in actively reflecting and deliberating on RI rather than merely providing them with set procedures to follow and ensuring compliance (2 (Chapter 4),5 (Chapter 3)). This opened my eyes to the importance of the framing of RI and – despite the fact that the European Commission call for the grant which funded my PhD project explicitly asked for the development of RI SOPs – my collaborators and I have made an effort to both omit the words 'SOPs' from the guidelines we developed, and also ensure that the guidelines we provide are in the spirit of supporting institutions in supporting researchers rather than telling institutions what to tell researchers (3 (Chapter 6)). Interestingly, some participants in the research even shared discomfort with the term 'RI' itself, saying it has negative connotations, and that it should be avoided in some cases (for instance, RI workshops offered to senior researchers could refer to 'research practice' rather than 'research integrity') to prevent resistance from researchers (2 (Chapter 4),3 (Chapter 6)). While I have the impression that the issue of framing is often implicitly assumed in the RI literature (e.g. 6–8), the need to make RI more attractive and frame it more positively through marketing and communication has not been previously explicitly addressed.

Of course, framing RI positively is not just a matter of word choice and marketing, but also has to do with the actual approach to RI policy that institutions take. While the majority of efforts to foster RI are still targeted at individual researchers' responsibilities, there has been a shift in recent years towards also addressing what institutions and other research stakeholders need to do in order to foster RI (8–11). While that is a welcome change, solely focusing on RI bodies (such as ombudspersons) and RI training – the majority of current practices focused on RI promotion at the institutional level (12 (Chapter 2)) – when creating RI policies is not adequate in framing RI as supporting researchers. RI bodies are important to address any conflicts or problems that might arise in the research process, but they are not sufficient in preventing problems from arising (5 (Chapter 3)). Education, alternatively, is considered valuable in preventing RI problems by creating awareness about RRPs and QRPs (5 (Chapter 3),13). However, especially considering that much of existing RI training efforts are currently focused on PhD students (14), it would be undesirable if institutions only offer training as a means to foster RI, since this would again leave the responsibility of doing responsible research on individual researchers (15). Instead, education – consisting of both formal training and informal RI awareness raising approaches, such as good supervision – can be integrated into the research endeavor, and combined with other RI policies to address both systematic and individual factors influencing RI (2 (Chapter 4)). For instance, it is not sufficient for institutions to provide adequate RI education if they do not take any means to reduce the high pressure that many researchers experience to obtain a large number of publications (15,16); RI education should be complemented with institutional policies aimed at reducing publication pressure and hyper-competition. Emphasizing systemic factors in RI policy (for instance, by addressing all the topics highlighted in Chapter 4) – rather than solely focusing on individual researchers – is necessary to frame RI positively, prevent researchers from feeling like they are being 'policed', and allow for the internalization of RRPs (4 (Chapter 7),8).

2.   Tailoring is essential

When I started this PhD trajectory, it was not clear how best to address contextual differences when developing guidance to all research institutions from across Europe on how to foster RI. I partially expected that my research results would provide some indication about differences between disciplinary fields or institutional types, which could then be directly taken up in the guidelines being developed (to indicate, e.g. that in the humanities, institutions can do X, while in the natural sciences, they can do Y). However, while conducting the research, it became apparent that there are more ways in which institutions can differ from each other than can be possibly categorized (just a few of which include disciplinary field, private or public status, legal requirements, region, type of researchers, size, historical background, cultural specificity, and type of funding available) (5 (Chapter 3)). The research steps also highlighted that despite these differences, diverse research stakeholders across Europe agree on the basic steps necessary to foster RI and raise awareness about it, and most differences of opinion relate to how to operationalize those basic steps into different contexts. This suggests that while higher level RI guidance (i.e. guidance discussing basic steps, which need further operationalization to be put into practice) might be relevant for institutions across Europe, there is no one size-fits-all solution to RI regarding how guidance is to be implemented (17 (Chapter 5)) at the institutional level; every institution will need to tailor policies to the needs of the researchers, stakeholders, and other important considerations (such as national laws) in the local context (5 (Chapter 3)). Stakeholders highlighted the importance of not being prescriptive and directive, and warned us against developing guidelines that are very detailed and try to account for each context of research. Therefore, the recommendations available in our guidelines – while much more concrete and workable than what is available in aspirational codes of conduct on RI – are still general enough and leave room for interpretation and tailoring depending on the local context. This was done intentionally to account for stakeholders' concerns that it is important that the implementers of the guidelines themselves have the freedom to tailor the guidelines to the specific context and needs of the institution that they work in.

The need to tailor RI policies to the specific context of the institution leaves open the question of how this can optimally be done by institutions. In this research, I argued that co-creation, jointly developing policies and guidelines with all relevant stakeholders (including researchers, policy makers, administrators, and staff), is key in providing legitimacy to institutional RI efforts, addressing stakeholders' actual needs, and ensuring that stakeholders actively embrace RI policies (4 (Chapter 7)). Co-creation is also important for addressing differences across institutions. Institutions aiming to implement our guidelines can co-create their implementation strategy together with relevant institutional stakeholders (including researchers), to ensure that the implementation is sensitive to the local context (3 (Chapter 6),4 (Chapter 7),5 (Chapter 3),18). For instance, stakeholders can be involved in the decision-making process about the roles and tasks of institutional RI bodies (4 (Chapter 7)).

The term 'co-creation' is often used to refer to the joint development of policies and guidelines with stakeholders. I distinguish 'co-creation' from 'co-creation methods', which refers to a specific and systematic approach to co-creation, based on generative

design research methodology (17 (Chapter 5)). While the importance of co-creation is already discussed extensively – albeit with different words – in the RI literature (e.g. 19), most current approaches to RI are not systematic. Where expertise, time, and resources allow, it would be valuable for institutions to consider specific methodological approaches to co-creation, rather than always doing co-creation intuitively (i.e. jointly developing policies with an unstructured approach). The advantage of doing co-creation systematically, for instance using co-creation methods or other participatory research methods (including approaches outlined in 20), is that group dynamics can be accounted for, and reaching conclusions too quickly can be prevented (17 (Chapter 5)). Furthermore, co-creation can occur at different levels and in different ways; some co-creation approaches involve stakeholders already at the stage of defining the problem and developing a solution, while other co-creation approaches only involve users once a solution has already been developed (i.e. at the stage of implementation) (20). There are significant advantages of involving stakeholders as much as possible in setting the policy agenda, developing and implementing guidelines. These include increased ownership and commitment of stakeholders towards policies (21). Institutions interested in creating RI policies that are sensitive to their context – rather than merely applying guidelines in a top-down manner – therefore, have a range of options available to jointly develop policies with stakeholders.

3. Box-checking is a danger

In every empirical step taken in this PhD trajectory, one of the key concerns that stakeholders expressed about RI policies is the danger of creating a 'check-box' mentality – a mentality concerned with checking boxes regarding meeting compliance to set requirements (2 (Chapter 4),3 (Chapter 6),5 (Chapter 3)). While check-boxing is an immediate concern of many RI stakeholders, there is little explicit attention to it in the RI literature (although it has been mentioned as a problem in the research ethics literature, e.g. 22). At the level of the researcher, a 'check-box' mentality entails individuals not reflecting on their practices carefully, but rather being focused on meeting legal and procedural requirements related to RI, and thereby potentially engaging in questionable behaviors not addressed in formal procedures (23). When considering the development and implementation of institutional RI policies, a check-box mentality is particularly a danger considering the risk that RI policies, such as mandatory education for various target groups, overburden individual researchers (5 (Chapter 3),24). If policies are seen solely as a burden, as opposed to a support for researchers, they can just become a box-ticking exercise. However, there are various approaches that institutions, ideally together with other stakeholders (such as funders) can use to reduce the burden that RI might impose on researchers. For instance, some RI educational events may be integrated into existing courses, workshops, and other events focused on research methods or research ethics (3 (Chapter 6)). In other cases, it could be that the burden imposed by RI on researchers is actually helpful in reducing other types of burdens (24). For example, taking time to plan a PhD trajectory, make a publication plan, consider collaboration, or write meticulous study and data-analysis protocols (which are then preregistered) can be crucial in preventing conflicts, delays,

methodological problems, and confusion later. While my own case is not sufficient to make any conclusions, it can definitely be said that engaging in RI 'burdens' at the start of this PhD trajectory was invaluable in ensuring a smooth PhD trajectory. Nonetheless, it cannot be denied that RI policies will, to some extent, inevitably increase burdens on researchers. So long as this burden is rooted in network processes – or in other words, is supported by agreements in the research community about what entails good research practice – and serves to further the research community's goals, such a burden can be justified and can allow for commitment from researchers (4 (Chapter 7)).

Although I had anticipated the concern about creating a check-box mentality at the level of individual researchers, it became apparent that check-boxing is a danger both at the individual researcher level and organizational level (2 (Chapter 4),5 (Chapter 3),25). At an organizational level, a 'check-box' mentality entails organizations ticking off that they have some policies in place (e.g. to make themselves appear eligible for funding) without actually considering how these policies can interact and influence research (26). Together with the stakeholders involved in this PhD trajectory, as well as my supervisors and collaborators, I gained various insights while working on the chapters of this PhD thesis on how to prevent a check-box mentality at the institutional level when fostering RI. First, co-creation of RI policies with stakeholders (discussed earlier as also important for tailoring policies to the context at hand) is important in ensuring that policies are meaningful to all institutional stakeholders and actually contribute towards better research practices. Secondly, institutions can focus on the values and goals that RI policies are aimed at, rather than allowing the resulting rules and procedures to become ends in themselves (4 (Chapter 7),25). When a specific rule does not help to strengthen the value that it was originally intended to represent, there can be flexibility in adhering to this rule (4 (Chapter 7),26,27). For instance, institutions that promote open science practices in their policies can take into account the consequences of open science policies on different data types and disciplinary fields, rather than be strict on researchers who may have legitimate reasons for not following a specific accepted open science practice (28). Thirdly, institutions can alter reward and incentive systems so as to reflect the values and goals that RI rules and procedures aim at (4 (Chapter 7),23). For instance, when mandating RI training, institutions can provide appropriate incentives to researchers to ensure that they are actually motivated to actively take part in RI training (2 (Chapter 4),3 (Chapter 6)). Fourthly, institutions can evaluate policies together with affected stakeholders and adjust them continuously, to ensure that they help with the aim of fostering RI, rather than just serving as a box-checking exercise (4 (Chapter 7)).

4.   RI is a journey

I was not surprised to hear from my research participants that RI is complex in that it is influenced by various stakeholders and factors, as well as their interactions, and that it has no simple solutions. This was indeed one of the assumptions I started out with at the beginning of this PhD trajectory. However, by talking to various stakeholders I learned that RI has no simple solutions because it is not an acute problem that can be solved permanently at once, but is rather a continuous journey that both individual

researchers and research institutions can take. What does this mean for individual researchers? Stakeholders explained that researchers are faced with various challenges, moral 'gray areas' (situations with no clear right or wrong answers), and dilemmas in every phase of their careers (2 (Chapter 4),29). Therefore, the effort to engage in RRPs and avoid QRPs requires continuity throughout a researcher's career. The RI education guidelines I co-created with stakeholders highlight this continuous feature of RI for researchers, by emphasizing the importance of not approaching RI education as a single training event, but rather a continuous process that students, junior, and senior researchers can engage in (in target group relevant and appropriate ways) to reflect on their research practice and deal with challenges they face (3 (Chapter 6)).

Stakeholders also alluded to the fact that it is important to approach RI as a continuous journey for research institutions. Creating comprehensive RI policies and tailoring these to the specific context of the institution (through co-creation), while beneficial, also entails a significant burden in terms of time, and human and material resources (3 (Chapter 6),9). By seeing RI as a journey, institutions can implement a step-wise approach to fostering RI, where implementation of various policies is spread out over a number of years (4 (Chapter 7)). Such a slow approach to the implementation of RI policies is necessary to ensure that policies are well thought through and soundly developed. A slow and step-wise approach is also important for gaining acceptance for policies among the research community, and thereby contributing to a robust culture of RI (4 (Chapter 7)). Furthermore, institutions should continuously monitor, evaluate and update policies to ensure that they are meaningful and actually contribute to RI (4 (Chapter 7)). This is also essential because of the higher turnover rate of the research workforce per year (30); continuously updating RI policies together with the research workforce – and thereby ensuring that new members' perspectives are included – is therefore needed to ensure that the research workforce feels intrinsically motivated to engage in RRPs.

Institutions do not have to approach RI as a journey they must take on their own, but can look at other institutions' journeys for inspiration, and even work together with them to address hurdles (e.g. related to resources) in their path towards RI. Institutions with less developed RI policies can look to – and even pair up and collaborate with – institutions with a more developed RI infrastructure for insight on how to build RI policies (1,3 (Chapter 6)). While tailoring interventions to the local context is important, not all interventions need to be reinvented, and resource sharing can be valuable in addressing time and resource limitations (3 (Chapter 6),31). For instance, there is no need for each RI trainer to create a new RI curriculum from scratch, when there are excellent online and offline appropriate resources that can be used, adjusted and modified for use in different training programs (e.g. 18,32). Institutions can therefore benefit largely from supporting their RI stakeholders, such as RI officers and ombudspersons, to engage in national or international networking activities, so as to learn from others and find ways to partner up with other institutions, for their RI journey (3 (Chapter 6)).

<u>Strengths and limitations</u>

Strengths

The work produced in this PhD thesis is based on an iterative process consisting of various empirical cycles including a scoping review, Delphi study, focus groups, co-creation workshops, and case analysis, which built on one another. The combined methods allowed me to obtain an in-depth understanding of various stakeholders' needs and perspectives regarding RI, and allowed for triangulation of data to continually develop and refine my main findings (33). Since the PhD trajectory was part of the larger Horizon 2020 funded SOPs4RI project, the research was also informed by additional studies in the project falling outside of the scope of this PhD trajectory. These studies, including a survey on research integrity and pilot study with research institutions who used the resulting guidelines from this PhD trajectory, helped to further refine and develop the insights gained in the PhD research (1,19,34,35).

   I co-created RI education and training guidelines, which can be implemented by research institutions across Europe with lead users (i.e. those who will be implementing the guidelines themselves). I involved research stakeholders at different stages of the research process and in various ways in different studies. Since I engaged with various stakeholders across countries in Europe (and beyond), genders, ranks, and disciplines, I was able to achieve a large and broad outreach with this research. This was valuable to ensure that all relevant perspectives are included in the guidelines, as well as to engage stakeholders closely and thereby increase the likelihood of the later implementation of the guidelines.

   I used a broad definition of RI throughout the PhD trajectory; not only did I include the entire spectrum of research practices, including research misconduct, QRPs and RRPs in the definition, but I was also open to broadening the term beyond its traditional scope (5 (Chapter 3)). While there were discussions within my research group about the limits of what can be included as an RI phenomenon and we 'agreed to disagree' on some specific points, I was eager to keep this broad definition of RI, which allowed for a better understanding of different stakeholders' understanding of RI, as well as to incorporate the full range of factors that have an influence on research quality and relevance. This meant that when our research participants raised issues such as conflicts within the workplace, harassment, or diversity considerations, I categorized these issues as relevant for RI, regardless of whether there is consensus on this in the RI literature.

   As I mentioned in the introduction, this policy-oriented research had a focus on making a concrete impact on the actual RI policies of research institutions in Europe. It seems that the work is already having impact on policy on the ground, since a number of institutions have piloted the guidelines resulting from this PhD trajectory (34). Furthermore, considering that the European Commission is actively promoting the work resulting from this research in the context of the Horizon Europe framework program, I expect to be able to make a broader impact on the actual practices of research institutions towards RI with my work (1).

Limitations

The insights gained in this PhD trajectory, including the RI education guidelines, have been informed by stakeholder engagement. Therefore, they are experience-based. However, they are not based on empirical evidence about what kinds of interventions are most effective in promoting RRPs and preventing QRPs and research misconduct. This is because there is currently a lack of evidence about which RI interventions are effective. Even in the domain of RI education, which is one of the most intensively researched RI topics, there is little known about if RI education is effective, and if so, what makes it so (14,36).

Although I engaged with a broad range of stakeholders, I did not equally engage with all types of stakeholders in the different stages of the research. In particular, since my research question was focused on the responsibilities of research institutions, I engaged more closely with institutional leaders and RI policy staff (e.g. rectors, deans, RI officers, ombudspersons, RI trainers, policy makers) and senior researchers, compared to junior researchers. Considering the evidence that junior researchers might perceive the research climate differently from senior ones (37,38), and most likely also from leaders and policy staff, further research is needed to explore junior researchers' perspectives of our recommendations. In addition, I did not engage extensively with journals, learned societies, governments, and the general public, in my research. This is also reflected in the fact that the RI education guidelines I developed were focused on institutional RI stakeholders, whereas other RI stakeholders such as those from funding organizations and government bodies, could also be a valuable target for RI education.

Each individual study comprising this thesis was designed by myself, my PhD supervisors, and collaborators joining for a specific study (within the limits set by our European Commission funding grant agreement), and stakeholders could contribute to the study mainly in the 'data collection' phase. In other words, we did not involve stakeholders in the development of our research protocols, per chapter of this PhD thesis. Considering 1) the high number of researchers collaborating in this PhD trajectory; 2) the fact that the stakeholders we were targeting in this research matched the research profile of my collaborators and myself – of being involved in research in one way or another and being interested in RI –; and 3) that our team is embedded in a wider community of RI stakeholders who meet regularly and exchange ideas about research, I considered it appropriate to only burden additional stakeholders with being involved in the 'data collection' phase of the research.

While I intentionally chose to work with a broad definition of RI in this research, I did not include experts and knowledge from the closely related field of Responsible Research and Innovation (RRI) (39) in my work. While I did include some Research Ethics (RE) considerations into the work (e.g. by identifying research ethics review as a topic that research institutions should address in their RI policies), I did not elaborate on this in depth. Further research could look into how insights from the fields of RI, RRI, and RE can be implemented in a complementary way in research institutions, to ensure that relevant insights from the fields of RRI and RE which might also be relevant for RI, are included in RI policy discussions.

Since the stakeholder engagement in this PhD trajectory was mostly focused on Europe, I would be hesitant to promote the use of the findings in research contexts

outside Europe. Further research is needed to explore in which ways the findings may or may not be relevant in various other settings. I expect, for instance, that due to the existence of a more legalistic culture in the USA (40,41), our non-legalistic approach to RI – where we provide tools for institutions (and thereby also researchers) on how to foster RI, rather than setting recommendations for legal requirements – might be less appropriate there. Furthermore, considering the presence of colonial legacies in research (42), and to prevent 'ethical imperialism' – the imposition of ethical rules and procedures of one community on another (43) – I would be hesitant to recommend institutions outside Europe to directly make use of these findings. I would rather recommend that researchers outside of Europe investigate to what extent the findings may be relevant or valuable to their specific research context, something that is typical for qualitative research results anyway.

Some issues and directions for future research

After completing this PhD trajectory, I am left with several issues which need further work. These relate to 1) institutional, country and disciplinary differences, 2) implementation of RI policy, 3) evaluation of RI policy, and 4) diversity and inclusion. In this section, I will briefly discuss each of these points and provide some suggestions for future research on them.

1.   Institutional, country, and disciplinary differences

As described earlier, considering the many types of contextual differences that are relevant for RI policy, I refrained from providing specific recommendations for different contexts when developing guidelines on RI education, and rather provided more general recommendations that can be tailored to specific contexts. Nonetheless, it would be valuable to have more insight into how differences among institutions, disciplinary backgrounds and countries might influence how research institutions can foster RI. At the level of institutions, some major differences that can be important to consider are whether the institution 1) is large or small, 2) is academic or commercial, 3) only focuses on research or also provides education; as well as 4) the level of RI infrastructure already present. These factors will not only have an influence on the availability of resources for implementing RI policies, but will also determine how RI policies should be optimally shaped. For instance, a commercial research institution may have more resources in place to implement RI policies than an academic one.

Country differences also play an important role in how RI policies can be implemented in institutions. There is a large variety in Europe regarding research and development infrastructure (44). This means that some countries will have more national resources and infrastructures in place to help institutions foster RI, than others. For example, in the Netherlands, the National Body for Research Integrity (LOWI) provides institutions with a second opinion on their handling of allegations of research misconduct (45). Other countries might have no such support systems in place. Additionally, in some countries, policies concerning RI breaches are more legalized than in others. In Denmark, there is a law on RI, as well as a strong national oversight committee (46). This has consequences for the role of different RI bodies in

the institution, as well as the procedures for handling allegations. Furthermore, country-specific cultural differences will also be important to consider when developing and implementing institutional RI policies. We found that in some European countries, there is not even a specific equivalent to the term 'research integrity' in the local language (for instance in Estonia and Croatia where RI and research ethics are referred to with the same term, 5 (Chapter 3)), making it difficult to separate RI and research ethics policies. This will likely differ from the approach in other countries, such as the Netherlands, where there are specific bodies and procedures for RE as compared to RI, as the two concepts are approached differently – albeit that they are related and overlap to some extent (45,47).

Differences at the disciplinary level are particularly important to investigate further, considering that many existing RI efforts and guidance documents originate from the biomedical field and certain social sciences (such as psychology) (12 (Chapter 2),48). The other social sciences and the humanities are, in particular, underrepresented in RI efforts (12 (Chapter 1),48). Therefore, there is a danger of imposing standards of research that have been found important for the biomedical field into research in other fields. For instance, replication might be seen as a gold standard in biomedical, natural and some social sciences, but its feasibility and desirability in other social sciences and the humanities has generated rigorous debate (49,50). Institutions could work together with faculties (representing specific disciplines) within the institution, or national or international disciplinary learned societies, to jointly explore the appropriate standards to promote and teach (e.g. through RI education) for different fields of research, where possible. Furthermore, more attention is needed to develop an RI infrastructure that is appropriate for non-biomedical disciplinary fields. For example, while scholars in the humanities and qualitative researchers might value open science practices as much as those in other fields, current open science discussions are mostly developed with quantitative research fields in mind (51). Similarly, in many institutions there is currently little guidance and institutional support (e.g. in the form of a research ethics committee) available regarding research ethics for non-biomedical research (52).

2.   Implementation of RI policies

I investigated various potential implementation opportunities and challenges for RI policies in my work (e.g. related to availability of resources to implement policies). Furthermore, our partners in the SOPs4RI project have piloted our work with a number of volunteer institutions (34). However, since implementation takes time, it would be interesting to explore the actual considerations that institutions address in practice as they implement RI policies over time, using a years-long longitudinal study design. On a shorter timescale, it might also be valuable to create an assessment instrument indicating the various degrees of implementation that institutions can strive for regarding different RI policies (for instance, similar to 53,54,55, used in different contexts). Institutions could then use this instrument to benchmark their progress in fostering RI. An implementation assessment instrument will require a rigorous development and validation research process before it can be used in practice (54). Ideally, this would be done together with stakeholders, rather than in a top-down

manner. Internal audits, if applied for the purposes of communicative action – namely to learn and create awareness – can be suitable in evaluating where institutions lie on the implementation assessment instrument and helping them to move further (54).

3.   Evaluation of RI policies

At various events in which I discussed my PhD work with others, I was asked about the evidence behind the RI education guidelines developed. When I mentioned that the guidelines are experience-based and developed through empirical research engaging with stakeholders, rather than evidence-based in the sense of being backed up by research on the effectiveness of certain interventions, I received critical comments. Institutional RI staff and policy makers often want evidence showing the effectiveness of certain RI policies on researchers' practices, in order to be convinced of the need to implement such policies. This might particularly be the case for policies which are burdensome or risky in terms of possible harmful side-effects. At the moment, however, there is little insight about what RI policies are most effective. Therefore, research is needed to provide more insight about the effectiveness of different RI policies.

My approach to dealing with the little available evidence currently in the RI education guidelines was to recommend institutions to use process evaluations concerning specific educational events to obtain insight into how stakeholders experienced RI education events, for continuous updating and improvement (3 (Chapter 6)). Considering the view that RI education events – and other RI policies more generally – should be aimed at supporting researchers in their work and helping them to do more responsible work, researchers' perspectives on how they experience RI interventions is crucial for evaluations of the need for specific RI policies (3 (Chapter 6)). However, while not the responsibility of every institution, there is also a need for research on outcome evaluations of different educational events (3 (Chapter 6)). Furthermore, outcome evaluation studies are needed in particular to investigate the effectiveness of continuous RI education. There are currently various initiatives for conducting such outcome evaluations (e.g. 56,57). The challenge in conducting outcome evaluation studies on RI education – or any other RI policies and interventions – is that it is difficult to know what to measure and how.

Previously, various outcome measures have been used to assess the effectiveness of single RI trainings, including knowledge of RI concepts, moral reasoning, and attitudes towards RI (58,59). Ultimately, however, what matters is whether the educational events lead to more RRPs and a more responsible research culture; but how can these outcomes be operationalized reliably and validly? Several strategies currently exist (59,60), including surveys asking researchers to self-report on their behaviors or their colleagues' behaviors (e.g. 59). However, this is not a direct measure of behavior, but rather researchers' perception of their own or others' behavior. The Survey of Organizational Research Climate (SoURCE), measures researchers' perceptions of the research climate and can also serve as an indirect measure of behavior (61). Recently, there have been institutional attempts to directly measure the quality of research produced at the institution. Both at the University of Edinburgh, as well as in Charité Berlin, for instance, there are efforts to map out the

quality of research by looking at reporting indicators (6,62). Such measures could be useful for conducting evaluation research on specific RI policies, as they are being introduced. Institutions can see whether the quality of the research produced at the institution, as measured by certain reporting indicators, improves after the introduction of novel RI policies. For instance, if the institution (or faculty or department within it) is implementing RI education for a certain target group for the first time, it can then use reporting indicators to check whether the education has had any influence on various aspects of research, such as open science practices.

Measuring the quality of research at the institution using certain indicators also allows for institutional benchmarking, as other institutions can use the same approach to compare their research to one another (6). While valuable for evaluation research, this could have large implications in the future. At the moment, universities as a whole (and likewise individual researchers) are ranked mainly on numbers of publications and citations (63), but efforts to change the reward and incentive system in academia might alter this. Other factors such as quality of teaching, altmetrics (such as ones focused on open science practices), and societal impact are now discussed as a potential complementary or alternative way of evaluating universities and individual researchers. While this might seem like a good move forward to reward transparency and openness in research, as opposed to flashy and quick results, such an approach comes with its own caveats.

For example, open science practices, while important for creating transparency and openness in research, are not by definition equivalent to RRPs. In certain cases, open science might actually conflict with RI. This is because RI entails more values and principles than just openness and transparency. For instance, the European Code of Conduct on Research Integrity also mentions "respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment" as a core RI principle (13). In some cases, respect for others might actually conflict with the value of openness; for instance, in cases of dual use research, or when researchers do not have sufficient funding to be able to publish in an open access journal. On the latter point, particularly considering how expensive open access fees can be, changing standards of good research practice to include open access publications will strongly disadvantage researchers from low and middle income countries (64,65). Therefore, equating open science with RI can be problematic; open science is a valuable way of making research more responsible, but careful reflection and consideration about the goals of specific practices are needed to assess whether open science practices are appropriate for the situation at hand.

The example of open science illustrates how solely shifting from one type of metric (i.e. impact factors) to a different type of metric (i.e. number of open access publications) is unlikely to be beneficial in actually improving the quality of research and moving institutions and researchers away from a 'box-checking mentality'. Instead, what is needed is increased awareness about the need to shift away from simple metrics altogether, and assess research policies and practices in a more holistic manner, combining quantitative and qualitative approaches. This is no easy task and will require further research, as well as potentially, a shift in paradigm about what counts as 'evidence'. Quantitative evidence is unlikely to provide the definite answers about what RI policies actually improve research practices as a whole (66). Such type of

evidence will need to be combined with qualitative procedures (66), aimed at involving stakeholders in defining outcome criteria and interpreting results, in order to provide a fuller and more informative picture of the effects of RI policies. This is already recognized by various institutions, as can be seen, for instance, with the introduction of researcher evaluations based on narrative curriculum vitae (CVs) rather than number of publications (67).

4. Diversity and inclusion

Aiming to provide insights to research institutions across Europe on fostering RI, I intended to be inclusive to different demographic factors when engaging with stakeholders. However, this inclusiveness – as was pointed out by some participants in the co-creation workshops I organized (17 (Chapter 5)) – fell short of important factors such as ethnicity, race, and cultural background. This was mainly because – despite there being a large minority of non-white people in Europe – of the lack of racial diversity among research stakeholders (and particularly those in higher ranking positions, and administrative and policy roles), which I am very much aware of, as a brown woman myself. It proved to be an overwhelming challenge to be inclusive regarding countries in Europe, disciplinary backgrounds, stakeholder types, institutions, career stages, gender, and ethnicity, race, and cultural background given all the practical challenges of stakeholder engagement (such as a finite number of participants in each study, and time and resources). Given my strong conviction that diversity and inclusiveness play an important role in research (in terms of the questions that are asked, the results that are developed, and the relevance for society), I see this as a large limitation of this research[5].

One way in which this may have affected the research results is that issues of diversity, inclusion, and equity remained marginal throughout the research process and barely discussed in the chapters of my thesis. While I had many discussions with my supervisors and collaborators to urge for the consideration of diversity and inclusion as something to address in RI policy (e.g. 68) – for instance by posing it as an institutional responsibility for RI in the Delphi study – the relevance of diversity and inclusion remained controversial throughout the research, with various collaborators and research participants doubting its relevance or importance for RI. I wonder whether this would have been the case had there been a larger diversity of stakeholders in the study. Still, I am happy that I remained persistent to give diversity and inclusion space within the larger project this PhD is part of, SOPs4RI (for instance as a crucial element of instilling a responsible research environment), even if this remained marginal.

Although the topic of diversity and inclusion remains controversial in the field of RI, lately, attention for it is growing. Namely, there is increasing acknowledgement that diversity and inclusion in research teams, as well as research samples, increases research quality and relevance and thus bears on trustworthiness and validity (48). Various efforts within the field of RI are now aiming at promoting diversity and inclusion more prominently, such as the 2022 World Conference on Research Integrity focused on 'Research in an unequal world' including the resulting Cape Town

---

[5] I did not discuss this in the limitations section, because I believe it deserves its own section as a larger discussion point.

Statement on Research Collaborations (69). As the topics of diversity, inclusion, and equity are gaining increased attention, there are efforts – particularly in the UK and USA – to create institutional policies aimed at improving diversity, inclusion, and equity in research (70). While this shift in attention to diversity and inclusion is welcome, there is a danger of approaching the topic from a shallow perspective merely focused on increasing the representation of women – or in some cases minority researchers – in the research endeavor without taking an in-depth look into how to increase the diversity of research ideas, approaches, methodologies, and designs themselves. In other words, diversity and inclusion policies that increase the number of diverse researchers in an institution – by asking them to integrate into a white-centric institution, rather than giving them the opportunity to express their diversity through their research ideas, approaches, methodologies and designs – do not really increase diversity and are not actually inclusive.

Furthermore, in discussions about diversity and inclusion and RI, there is a need to differentiate between increasing diversity and inclusion within European institutions as compared to research more generally. Addressing the latter requires more than the implementation of institutional policies; it requires having a hard and painful look at the history of knowledge production and the power structures that have privileged certain ways and types of knowledge over others, and thereby biased our way of understanding the entire world in favor of Eurocentric ideals and standards (42). In fact, I believe that in both cases, RI researchers and policy stakeholders should address the coloniality of research more seriously, for instance by acknowledging them in codes of conduct for RI, and working together with anti-colonial scholars to decolonize research practices. This is particularly important to prevent ethical imperialism of research – the imposition of European ideals on research done in other contexts and settings – or at least minimize it, given that it is already prevalent (43). Considering that there is already resistance to diversifying research in a more shallow sense (e.g. appointing more women in research) currently, it is likely that pursuing such an approach to RI will be a difficult and long path.

Conclusion

If RI is about producing better research – and 'better' might mean different things in various contexts – then only when research institutions work together with other research stakeholders to co-create meaningful RI policies that address the various interrelated factors influencing research and researchers, such as awareness about research practices, rewards and incentives, and the research culture – is it possible to further the research community's goals and produce 'better' research. RI is a journey. In the words of Schaller-Demers (23):

"[RI] is an ongoing conversation. It does not begin and end with mandated legislation, taking a course or attending a workshop... There is a plethora of valuable online resources, conferences, and organizations that are devoted to [RI]-related topics. These are helpful and convenient tools – but they are only effective when incorporated into a culture of responsibility that has been embedded into the organizational structure."

Contributions

KL drafted and revised this chapter. LB, GW, NE and JT reviewed drafts and accepted the final version of the chapter.

References

1. *Mejlgaard N, Bouter L, Gaskell G, Kavouras P, Allum N, Bendtsen AK, et al. Research integrity: nine ways to move from talk to walk. Nature. 2020;586:358–60.*
2. *Labib K, Evans N, Roje R, Kavouras P, Reyes Elizondo A, Kaltenbrunner W, et al. Education and training policies for research integrity: Insights from a focus group study. Sci Public Policy. 2021;49(2):scab077. DOI: https://doi.org/10.1093/scipol/scab077*
3. *Labib K, Evans N, Pizzolato D, Aubert Bonn N, Widdershoven G, Bouter L, et al. Co-creating research integrity education guidelines for research institutions. MetaArXiV Prepr. 2022; DOI: 10.31222/osf.io/gh4cn*
4. *Labib K, Tijdink JK, Sijtsma K, Bouter L, Evans N, Widdershoven G. How to combine rules and commitment in fostering research integrity? MetaArXiV Prepr. 2022; DOI: 10.31222/osf.io/sx58q*
5. *Labib K, Roje R, Bouter L, Widdershoven G, Evans N, Marušić A, et al. Important topics for fostering research integrity by research performing and research funding organizations: a Delphi consensus study. Sci Eng Ethics. 2021;27(47):1–22. DOI: 10.1007/s11948-021-00322-9*
6. *Macleod M. Want research integrity? Stop the blame game. Nature. 2021;599:533.*
7. *Research integrity is much more than misconduct. Nature. 2019;570(5). DOI: 10.1038/d41586-019-01727-0*
8. *Zwart H, Ter Meulen R. Addressing research integrity challenges: from penalising individual perpetrators to fostering research ecosystem quality care. Life Sci Soc Policy. 2019;15(5).*
9. *Bouter L. Fostering responsible research practices is a shared responsibility of multiple stakeholders. J Clin Epidemiol. 2018;96:143–6.*
10. *Lerouge I, Hol T. Towards a research integrity culture at universities: From recommendations to implementation [Internet]. League of European Research Universities; 2020. Available from: https://www.leru.org/files/Towards-a-Research-Integrity-Culture-at-Universities-full-paper.pdf*
11. *Titus S, Bosch X. Tie funding to research integrity. Nature. 2010;466(7305):436–7.*
12. *Ščepanović R, Labib K, Buljan I, Tijdink J, Marušić A. Practices for Research Integrity Promotion in Research Performing Organisations and Research Funding Organisations: A Scoping Review. Sci Eng Ethics. 2021;27(4):1–20. DOI: 10.1007/s11948-021-00281-1*
13. *ALLEA. European Code of Conduct for Research Integrity [Internet]. Berlin; 2017. Available from: https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf*
14. *Fanelli D. MLE on Research Integrity – Training & Education – Thematic Report No 4 [Internet]. 2019. Available from: https://rio.jrc.ec.europa.eu/sites/default/files/report/MLE RI No4 - Training and education.pdf*
15. *Aubert Bonn N, Pinxten W. Rethinking success, integrity, and culture in research (Part 2)—A multi-actor qualitative study on problems of science. Res Integr Peer Rev. 2021;6(1):1–18.*
16. *Aubert Bonn N, Bouter L. Research assessments should recognize responsible research practices—Narrative review of a lively debate and promising developments. OSF Preprints. 2021. DOI: 10.31222/osf.io/82rmj*
17. *Labib K, Pizzolato D, Stappers PJ, Evans N, Lechner I, Widdershoven G, et al. Using co-creation for guideline development - How why and when? OSF Prepr. 2021; DOI: 10.31219/osf.io/cg3rw*

18.   Simon C, Beerman RW, Ariansen JL, Kessler D, Sanchez AM, King K, et al. Implementation of a responsible conduct of research education program at Duke University School of Medicine. Account Res. 2019;26(5):288–310.

19.   Roje R, Reyes Elizondo A, Kaltenbrunner W, Buljan I, Marušić A. Factors influencing the promotion and implementation of research integrity in research performing and research funding organizations: a scoping review. Account Res. 2022; DOI: https://doi.org/10.1080/08989621.2022.2073819

20.   Bergold J, Thomas S. Participatory research methods: A methodological approach in motion. Hist Soc Res Sozialforsch. 2012;191–222.

21.   Cargo M, Mercer SL. The value and challenges of participatory research: strengthening its practice. Annu Rev Public Heal. 2008;29:325–50.

22.   Landi MS, Shriver AJ, Mueller A. Consideration and checkboxes: incorporating ethics and science into the 3Rs. J Am Assoc Lab Anim Sci. 2015;54(2):224–30.

23.   Schaller-Demers DS. Responsible Conduct of Research: Not Just for Researchers. J Res Adm. 2015;46(1):63–76.

24.   Barnes BE, Friedman CP, Rosenberg JL, Russell J, Beedle A, Levine AS. Creating an infrastructure for training in the responsible conduct of research: the University of Pittsburgh's experience. Acad Med. 2006;81(2):119–27.

25.   Sørensen MP, Ravn T, Marušić A, Elizondo AR, Kavouras P, Tijdink JK, et al. Strengthening research integrity: which topic areas should organisations focus on? Humanit Soc Sci Commun. 2021;8(1):1–15.

26.   Anand N. 'Checkbox Diversity' Must Be Left Behind for DEI Efforts to Succeed. Stanford Social Innovation Review [Internet]. 2019; Available from: https://ssir.org/articles/entry/checkbox_diversity_must_be_left_behind_for_dei_ef forts_to_succeed DOI: https://doi.org/10.48558/ZVYC-EK13

27.   Seth Sikkema CPA. The Harmonization of Accounting Standards. ACBSP Annu Ed. 2012;3:71–7.

28.   University of Eastern Finland. Basics of Open and Responsible Science: Open Access Publishing and Research Data [Internet]. [cited 2022 Jan 14]. Available from: https://blogs.uef.fi/ueflibrary-bors/benefits-and-challenges-of-open-science/

29.   Goddiksen MP, Gjerris M. Teaching phronesis in a research integrity course. FACETS. 2022;7(1):139–52.

30.   Figueron O. The influences impacting staff turnover in higher education. J Mgmt Sustain. 2015;5:86.

31.   van den Hoven MA, Theunissen B. Taking RCR seriously in teaching scientific integrity. In: Bioethica Forum. Midwest Bioethics Center; 2021. p. 89–93.

32.   Evans N, Marusic A, Foeger N, Lofstrom E, van Hoof M, Vrijhoef-Welten S, et al. Virtue-based ethics and integrity of research: train-the-trainer programme for upholding the principles and practices of the European Code of Conduct for Research Integrity (VIRT2UE). Res Ideas Outcomes. 2021;7:e68258.

33.   Carter N, Bryant-Lukosius D, DiCenso A, BlytheA J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurse Forum. 2014;41(5):545–7.

34.   Konach T, Horbach, Serge Rochambeau M, Taraj B. D7.2: Report on Pilot Studies [Internet]. 2022. Available from: https://sops4ri.eu/wp-content/uploads/D7.2_Report-on-Pilot-Studies.pdf

35.   Roje R, Tomić V, Buljan I, Marušić A. Development and implementation of research integrity guidance documents: Explorative interviews with research integrity experts. Account Res. 2021;1–38.

36.   Antes AL, Murphy ST, Waples EP, Mumford MD, Brown RP, Connelly S, et al. A meta-analysis of ethics instruction effectiveness in the sciences. Ethics Behav. 2009;19(5):379–402.

37.   Haven T, Tijdink J, Martinson BC, Bouter L. Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. PLoS One. 2019;14(1):e0210599.

38.   Armond ACV, Kakuk P. Perceptions of Research Integrity Climate in Hungarian Universities: Results from A Survey among Academic Researchers. Sci Eng Ethics. 2022;28(4):1–12.

39. *Owen R, Macnaghten P, Stilgoe J. Responsible research and innovation: From science in society to science for society, with society. Sci Public Policy. 2012;39(6):751–60.*

40. *Kelemen RD, Sibbitt EC. The globalization of American law. Int Organ. 2004;58(1):103–36.*

41. *Kagan RA. Should Europe worry about adversarial legalism. Oxford J Leg Stud. 1997;17:165.*

42. *Yako L. Decolonizing Knowledge Production: A Practical Guide [Internet]. CounterPunch. 2021. Available from: https://www.counterpunch.org/2021/04/09/decolonizing-knowledge-production-a-practical-guide/*

43. *Israel M. Ethical imperialism? Exporting research ethics to the global south. Sage Handb Qual Res ethics. 2017;89–102.*

44. *Godecharle S, Nemery B, Dierickx K. Heterogeneity in European research integrity guidance: Relying on values or norms? J Empir Res Hum Res Ethics. 2014;9(3):79–90.*

45. *Fernhout R. Research integrity procedures in the Netherlands. In: 3rd Interuniversity symposium on Research Integrity [Internet]. Brussels; 2017. Available from: https://www.vcwi.be/sites/default/files/symposiumWI2017_RIpcoeduresinNL_Fernhout.pdf*

46. *Danish National Research Foundation. Worth knowing: Research integrity from a Danish perspective [Internet]. 2021. Available from: https://dg.dk/en/research-integrity/*

47. *Kenter MJH. Regulating human participants protection in medical research and the accreditation of medical research ethics committees in the Netherlands. J Acad Ethics. 2009;7(1):33–43.*

48. *Hastings RC, Labib K, Lechner I, Bouter L, Widdershoven G, Evans N. Guidance on research integrity provided by European discipline-specific learned societies: A scoping review. MetaArXiV Prepr. 2022; DOI: 10.31222/osf.io/kn5y9*

49. *Penders B, Holbrook JB, De Rijcke S. Rinse and repeat: Understanding the value of replication across different ways of knowing. Publications. 2019;7(3):52.*

50. *Peels R, Bouter L. The possibility and desirability of replication in the humanities. Palgrave Commun. 2018;4(1):1–4.*

51. *Giglia E. OPERAS: bringing the long tail of Social Sciences and Humanities into Open Science. JLISIt. 2019;10(1):140–56.*

52. *De Wet K. The importance of ethical appraisal in social science research: reviewing a faculty of humanities' research ethics committee. J Acad Ethics. 2010;8(4):301–14.*

53. *Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, et al. Promoting an open research culture. Science. 2015;348(6242):1422–5.*

54. *Van Melle AL, Voskes Y, de Vet HCW, Van der Meijs J, Mulder CL, Widdershoven GAM. High and intensive care in psychiatry: validating the HIC monitor as a tool for assessing the quality of psychiatric intensive care units. Adm Policy Ment Heal Ment Heal Serv Res. 2019;46(1):34–43.*

55. *Peels R, Van Woudenberg R, de Ridder J, Bouter L. Academia's Big Five: a normative taxonomy for the epistemic responsibilities of universities. F1000Research. 2019;8(886).*

56. *Zollitsch L, Wilder N, Priess-Buchheit J. The development of a four-tier test to evaluate training in responsible conduct of research. Open Res Eur. 2022;1(90).*

57. *Katsarov J, Andorno R, Krom A, van den Hoven M. Effective Strategies for Research Integrity Training—a Meta-analysis. Educ Psychol Rev. 2022;34:935–955. DOI: https://doi.org/10.1007/s10648-021-09630-9*

58. *Kalichman M. Rescuing responsible conduct of research (RCR) education. Account Res. 2014;21(1):68–83.*

59. *Antes AL, Wang X, Mumford MD, Brown RP, Connelly S, Devenport LD. Evaluating the effects that existing instruction on responsible conduct of research has on ethical decision making. Acad Med J Assoc Am Med Coll. 2010;85(3):519.*

60.   Steele LM, Mulhearn TJ, Medeiros KE, Watts LL, Connelly S, Mumford MD. How do
        we know what works? A review and critique of current practices in ethics training
        evaluation. Account Res. 2016;23(6):319–50.

61.   Wells JA, Thrush CR, Martinson BC, May TA, Stickler M, Callahan EC, et al. Survey of
        organizational research climates in three research intensive, doctoral granting
        universities. J Empir Res Hum Res Ethics. 2014;9(5):72–88.

62.   Haven TL, Holst MR, Strech D. Stakeholders' views on an institutional dashboard
        with metrics for responsible research. PLoS One. 2022;17(6):e0269492.

63.   Stergiou KI, Lessenich S. On impact factors and university rankings: from birth to
        boycott. Ethics Sci Environ Polit. 2014;13(2):101–11.

64.   Pettorelli N, Barlow J, Nuñez MA, Rader R, Stephens PA, Pinfield T, et al. How
        international journals can support ecology from the Global South. Vol. 58, Journal
        of Applied Ecology. 2021. p. 4–8.

65.   Serwadda D, Ndebele P, Grabowski MK, Bajunirwe F, Wanyenze RK. Open data
        sharing and the Global South—Who benefits? Science. 2018;359(6376):642–3.

66.   Wilsdon J. We need a measured approach to metrics. Nat News. 2015;523(7559):129.

67.   Hatch A, Curry S. Research culture: Changing how we evaluate research is difficult,
        but not impossible. Elife. 2020;9:e58654.

68.   Labib K, Evans N. Gender, diversity, and the responsible assessment of researchers.
        PLoS Biol. 2021;19(4):e3001036.

69.   Horn L, Alba S, Blom F, Faure M, Flack-Davison E, Gopalakrishna G, et al. Fostering
        Research Integrity through the promotion of fairness, equity and diversity in
        research collaborations and contexts: Towards a Cape Town Statement (pre-
        conference discussion paper). OSF Prepr. 2022; DOI: 10.31219/osf.io/bf286

70.   Khan MS, Lakha F, Tan MMJ, Singh SR, Quek RYC, Han E, et al. More talk than
        action: gender and ethnic diversity in leading public health universities. Lancet.
        2019;393(10171):594–600.

List of publications

List of publications

Publications labelled with an asterisk (*) are included as chapters of this thesis.

Published

Hastings RC, Labib K, Lechner I, Bouter L, Widdershoven G, Evans N. and on behalf of the EnTIRE Consortium. Guidance on research integrity provided by European discipline-specific learned societies: A scoping review. Science and Public Policy. 2022. 00:1-18 DOI: 10.1093/scipol/scac067

Horbach SP, Bouter LM, Gaskell G, Hiney M, Kavouras P, Mejlgaard N, Allum N, Aubert Bonn N, Bendtsen AK, Charitidis CA, Claesen N, Dierickx K, Domaradzka A, Reyes Elizondo A, Foeger N, Kaltenbrunner W, Konach T, Labib, K, Marusic A, Pizzolato D, Ravn T, Roje R, Sorensen MP, Taraj B, Veltri G, Tijdink JK. Designing and implementing a research integrity promotion plan: Recommendations for research funders. PloS Biology. 2022 Aug 19;20(8):e3001773. DOI: https://doi.org/10.1371/journal.pbio.3001773

Labib K, Evans N. Gender, diversity, and the responsible assessment of researchers. PLoS Biology. 2021 Apr 27;19(4):e3001036.DOI: https://doi.org/10.1371/journal.pbio.3001036

*Labib K, Evans N, Roje R, Kavouras P, Reyes Elizondo A, Kaltenbrunner W, Buljan I, Ravn T, Widdershoven G, Bouter L, Charitidis C., Sørensen MP, Tijdink J. Education and training policies for research integrity: Insights from a focus group study. Science and Public Policy. 2022 Apr;49(2):246-66.DOI: https://doi.org/10.1093/scipol/scab077

*Labib K, Pizzolato D, Stappers PJ, Evans N, Lechner I, Widdershoven G, Bouter L, Dierickx K, Bergema K, Tijdink J. Using co-creation for RI guideline development-How, what, why and when. Accountability in Research. 2022. DOI: 10.1080/08989621.2022.2154154

*Labib K, Roje R, Bouter L, Widdershoven G, Evans N, Marušić A, Mokkink L, Tijdink J. Important topics for fostering research integrity by research performing and research funding organizations: A Delphi consensus study. Science and

Engineering Ethics. 2021 Aug;27(4):1-22. DOI: : https://doi.org/10.1007/s11948-021-00322-9

Mejlgaard N, Bouter LM, Gaskell G, Kavouras P, Allum N, Bendtsen AK, Charitidis CA, Claesen N, Dierickx K, Domaradzka A, Reyes Elizondo A, Foeger N, Hiney M, Kaltenbrunner W, Labib K, Marusic A, Sorensen MP, Ravn T, Scepanovic R, Tijdink JK, Veltri GA. Research integrity: nine ways to move from talk to walk. Nature. 2019 Oct; 586: 358-60. DOI: https://doi.org/10.1038/d41586-020-02847-8

Pizzolato D, Labib K, Skoulikaris N, Evans N, Roje R, Kavouras P, Aubert Bonn N, Dierickx K, Tijdink J. How can research institutions support responsible supervision and leadership?. Accountability in Research. 2022 Aug 13. DOI: https://doi.org/10.1080/08989621.2022.2112033

*Ščepanović R, Labib K, Buljan I, Tijdink J, Marušić A. Practices for research integrity promotion in research performing organisations and research funding organisations: A scoping review. Science and engineering ethics. 2021 Feb;27(1):1-20.DOI: https://doi.org/10.1007/s11948-021-00281-1

Preprints

*Labib K, Evans N, Pizzolato D, Bonn NA, Widdershoven G, Bouter L, Konach T, Langendam M, Dierickx K, Tijdink JK. Co-creating research integrity education guidelines for research institutions. MetaArXiv Preprints. DOI: 10.31222/osf.io/gh4cn

*Labib K, Tijdink JK, Sijtsma K, Bouter L, Evans N, Widdershoven G. How to combine rules and commitment in fostering research integrity?. MetaArXiv Preprints. DOI: 10.31222/osf.io/sx58q

Acknowledgments

Acknowledgements

There are so many people who contributed towards my PhD journey, whom I would like to thank.

I have had the privilege of receiving guidance and support form the best supervisors a PhD student could ask for:

Lex, thank you so much for introducing me to the field of research integrity and showing me its allure. You have taught me to treat mistakes as learning moments, rather than failures, and grow from them. You have shown me what healthy discussions look like in an academic setting, and always urged me to voice my opinion – even when it differed from yours or others in higher ranked positions than myself – and also taught me how to do this in a constructive and respectful manner. I have learned a lot from you not just as a PhD student, but also as a teacher and a supervisor to students myself. Thank you for the world of wisdom you have given me.

Guy, thank you for showing me my strengths and making me feel valued. You have made continuous efforts to highlight what you appreciate in me as a researcher, which has been especially important in moments of doubts and roadblocks in my research journey. Whenever you have sensed these doubts, a phone call from you has been sufficient in getting me in good spirits and helping me move forward. With your sharp and witty insights and comments, you have also always urged me, and given me the necessary theoretical, methodological, and professional tools, to improve myself further and become a better researcher. Thank you for being an inspiration for me, with your brilliance and kindness.

Natalie, thank you for constantly keeping an eye out for me – not just in terms of supporting me with my PhD journey – but also by giving me opportunities outside it to make sure that I am building my possibilities for the future. Your extensive and varied experiences with qualitative research have been a continuous source of insight and motivation for me. You have also always made it clear to me, despite your many responsibilities as a senior researcher, that I can count on you not just for input about larger methodological issues in my research, but also with the more mundane questions and tasks that I struggle with. I will dearly miss interrupting you with – sometimes silly – questions, to end up having long chats and laughter. Thank you for being not only a great supervisor, but also a friend. Thank you also for being a role model, in many ways.

Joeri, thank you for listening and caring. Your abundant jokes, attention to my well-being, willingness to spend meeting time on just reflecting together on life and social justice, and always great advice, have made this PhD journey a breeze for me. Your motivation and passion in your work has helped to keep me focused and engaged, even during challenging moments. I am always inspired by your ideas and genuine drive to improve research. You have taught me the importance of focusing on collaboration,

and not just zooming into tasks, for doing successful and fun projects. I have learned so much from you regarding how to think of research ideas, design a study, facilitate sessions, reflect on the implications, and communicate this with others. I have also always admired how you are able to both create a supervisor-supervisee relationship that is open, two sided and non-hierarchical, but also look out for me and protect me when I need it. I know that I am truly spoilt to have had you as my day-to-day PhD supervisor, and I will dearly miss that experience.

I like to thank my current and former amazing colleagues at the department of Ethics, Law and Humanities for providing a wonderful environment in which to work:

Giulia, thank you for being there from before the start and encouraging me, both as a teacher and peer. It has been reassuring throughout my journey to know that I can count on you.

Noémie, I am so happy that we could work together in the past two years, because you are such a constructive and thoughtful colleague, and a joy to work with.

Iris, you have been such a great collaborator and friend, particularly with the co-creation workshops in which we worked more closely together, and I could not have done everything without your help.

I would also like to thank Karl, Sabina, Anne, Effy, Ben, Suzanne, Malene, Charlotte, Sylvia, Laura vM, Laura H, Lieke, Sisco, Janine, Maartje, Saskia, Wieke, Margreet, Gowri, Fenneke, Bert and many others at the department who have helped me in my research journey.

Working within the SOPs4RI consortium, I did not just have my department in Amsterdam as part of my research family, but also my SOPs4RI collaborators all over Europe. Thank you all for the opportunity to take part in such a motivating and impactful project. Thank you especially to Rea, Borana, Daniel, Mads, Wolfgang, Andrea, Nik, Kris, Ana, Teodora, Tine, and Panagiotis for the fun and wonderful collaborations we had.

There are numerous other collaborators who have supported me in different and important ways throughout these years, who I am grateful to, including Wieneke, Dorien, Pieter Jan, Katinka, Klaas, Rosie, Niko, Tamarinde, and many others.

At various stages of my PhD trajectory, I was lucky to receive input from the always insightful members of my PhD steering committee. Thank you Sonja, Henk, Rita, Klaas, and Ton.

I am really glad to have met Monika during the last years of my PhD. *Monika*, although not directly involved in my PhD trajectory, due to your work and everything you stand for, you have become an important role model and inspiration for me. Thank you for that.

Thank you especially to all the research participants and co-creators in this project. The research was made possible by you. Thank you for giving me your time.

Finally, I would like to thank my dear friends, my (sweet immediate, as well as warm and resilient extended) family, and my incredible husband, for the emotional support, listening to me go on about my research without complaining, and giving me the love, inspiration, and strength needed to continue and finish.

Research integrity is necessary to safeguard the trustworthiness of research.

What are research institutions currently doing to foster research integrity and raise awareness about it? What should they be doing?

What are researchers' and other stakeholders' perspectives on this? How can guidelines on research integrity for research institutions be co-created with stakeholders? What should these guidelines contain?

How can research institutons combine the implementation of policies based on these guidelines with researcher commitment to foster research integrity?

This PhD thesis explores these questions to address institutional responsibilities towards research integrity.