Fostering a Research Integrity Culture:

Actionable Advice for Institutions

Steven De Peuter[1] 0000-0003-4137-4431

Stijn Conix[2] 0000-0002-1487-0213

[1] Methods, Individual and Cultural Differences, Affect and Social Behavior (MICAS), KU Leuven, Belgium

[2] Centre for Logic and Philosophy of Science, Institute of Philosophy, KU Leuven, 3000 Leuven, Belgium

Research institutions' research culture is increasingly recognized as a central driver of research integrity. Institutions are urged to develop research integrity promotion plans to foster a culture of research integrity. A host of guidelines and associated initiatives have been issued but specific, actionable recommendations for institutions are generally lacking. Based on a broad literature review, in the current paper some practical advice for institutions is suggested, grouped into 1) policies, procedures, and processes, 2) dealing with breaches of research integrity, 3) education and training, and 4) monitoring and evaluation. From the review, it is clear that more research is necessary to gather additional knowledge on what works and what does not work, knowledge that is ideally shared openly. This will allow institutions to learn from one another, facilitate harmonization of policies, and may possibly create a virtuous cycle that empowers the shift towards a culture of collective openness in science.

Keywords: Research integrity; Research culture; Research misconduct; Responsible conduct of research; Review, tenure and promotion

Address for Correspondence Steven De Peuter, Tiensestraat 102 box 3727, 3000 Leuven, Belgium. steven.depeuter@kuleuven.be

Author note Steven De Peuter expresses his gratitude to Gert Storms for his constructive feedback on a previous version of this text and for his incessant moral righteousness. The authors have no conflicts of interest to declare. The ideas expressed in this paper are those of the authors and not necessarily those of KU Leuven.

Funding This work was supported by the Research Foundation -- Flanders (FWO) [3H200026 to S.C].

CRediT Author Contribution Statement Steven De Peuter: Conceptualization; Investigation; Writing – original draft; Writing – review & editing. Stijn Conix: Investigation; Writing – review & editing.

Introduction

There has been a surge of interest in scientific integrity over de past decade, covering grave cases of fraud, falsification and plagiarism (FFP) as well as other less blatant practices known as 'questionable research practices' (QRPs; Fanelli 2009; Xie et al. 2021). This has led to empirical work on research integrity as well as a host of guidelines and recommendations for improving it. With different papers targeting different stakeholders, such as institutions, funders, governments and researchers, as well as different kinds of breaches of integrity and the causes of such breaches, keeping track of the available methods of improving research integrity is no easy task. The aim of this paper is to collect and order specific, actionable recommendations for institutions.

Initial research into the drivers of research misconduct investigated personality factors – in particular the 'Dark triad of personality' (narcissism, psychopathy and Machiavellianism) and individual differences in (dis)honesty. However, none was a strong driver of questionable behaviour (Mazar and Ariely 2015; Tijdink et al. 2016). In contrast, *systemic* factors such as publication pressure, perverse evaluation mechanisms and organizational climate[1] have been shown to be related to research misconduct and QRPs (Bouter 2015; Crain et al. 2013; DuBois et al. 2013; Gopalakrishna, ter Riet, et al. 2021; Haven et al. 2021).

The historical response has been one of codification and compliance from the level of professional associations to the (supra)national level (ALLEA 2017; Desmond and Dierickx 2021; Godecharle et al. 2013; Hastings et al. 2022; Komić et al. 2015; Resnik, Neal, et al. 2015; Resnik, Rasmussen, et al. 2015). In recent years, a shift has taken place from repression towards a positive, constructive and aspirational *systems* approach of 'how to behave well': the *Responsible Conduct of Research* (RCR) or *research integrity* (e.g., Casadevall and Fang 2012; Editorial 2019; Fang and Casadevall 2012; Forsberg et al. 2018). Both the Institute of Medicine (National Research Council 2002) and Aubert Bonn and Pinxten (2021a) have proposed a comprehensive model in which skills, personality and history of the individual researcher interact with characteristics of the research institution and of related stakeholders to define both successes (Aubert Bonn and Pinxten 2021b) and problems (Aubert Bonn and Pinxten 2021a). Similarly, UK Research and Innovation describes a research ecosystem with incentives acting upon 'research activities', 'institutions and employment', 'funding and policy instruments', or their respective intersections (Vitae 2020). Despite the growing agreement about systemic factors as the main drivers of research integrity, however, most empirical work still seems to target individual researchers (Aubert Bonn and Pinxten 2019). Because institutions control many of the main incentives that structure academic research, they are ideally placed to instigate the shift from individual-level solutions to system-level solutions.

A host of associated initiatives has been launched, targeting the scientific process and the different stakeholders and actors, with the common goal of promoting rigorous, honest, authentic, reliable, valuable and trustworthy research[2]. We will refer to some of those initiatives throughout the paper.

What Institutions Can Do

Already in 2002, The National Research Council (2002) encouraged institutions to promote and foster a culture in which high ethical standards are the norm and ongoing professional development is encouraged. Since then, several guiding documents have focused on institutional responsibilities and organizational measures to strengthen research integrity (e.g., Committee on Responsible Science et al. 2017; Forsberg et al. 2018; Lerouge and Hol 2020). Research Integrity Promotion Plans have been

strongly encouraged (Labib et al. 2021; Lerouge and Hol 2020; Mejlgaard et al. 2020) and may become a contractual obligation within the next EU framework program 'Horizon Europe' (Bouter 2020). However, which *specific* actions research institutions can implement – or where to start remains unclear (Mejlgaard et al. 2020).

In the following, we review actionable suggestions from the literature. Roughly, these can be divided into 1) processes, policies, and procedures, 2) dealing with breaches of research integrity, 3) education and training, and 4) monitoring and evaluation, which are also the main sections of the paper. We conclude each section with the most important recommendations.

Processes, Policies, and Procedures

Rules, Codes of Conduct, and Guidelines

Rules, codes of conduct and guidelines may be a, if not *the*, major way for institutions to target systemic causes of integrity breaches. For example, there is general consensus that institutions should implement clear and fair authorship guidelines and obligations. A transition from authorship to contributorship may even alleviate some of the pressure associated with authorship and publishing by permitting assigning 'official' credit to contributors without granting authorship. The CRediT-taxonomy (CRediT n.d.) was developed specifically for this purpose and is increasingly adopted by scientific journals and institutions. Being a machine-readable standard, it can be mined from meta-data of published papers (Holcombe 2019), creating visibility of individual contributions that would otherwise be stacked away in acknowledgements.

Similarly, institutions could adopt clear guidelines and rules for mentoring. Poor mentoring is perceived as a highly impactful research misbehaviour (Bouter et al. 2016; Gopalakrishna, ter Riet, et al. 2021; Haven et al. 2019) whereas *responsible* mentoring is a central factor in *fostering* research integrity (Anderson et al. 2007; Forsberg et al. 2018; Gopalakrishna, Wicherts, et al. 2021; Labib et al. 2021; Sørensen et al. 2021). Nevertheless, junior academics often perceive their supervision as less than optimal (Haven et al. 2019; Wells et al. 2014). Therefore, there is an urgent need to better support and train supervisors in mentoring junior researchers, and possibly also to *evaluate* supervisors' guidance (ALLEA 2017; Lerouge and Hol 2020; Mejlgaard et al. 2020; Titus and Ballou 2014).

Besides specifying what is *not* allowed, operationally defining the desired situation in terms of expected behaviours can make abstract ideas and aspirations tangible and enable institutions to translate their ambitions into specific goals and actions – which can then be monitored and evaluated (Labib et al. 2021; Lerouge and Hol 2020; Valkenburg et al. 2021).

Whenever possible, adopting existing codes and guidelines is recommended, preferably those supported by a legislative framework when violations need to be penalized. 'As generic as possible, as (discipline-)specific as necessary' avoids confusion in the increasingly interdisciplinary contemporary research landscape (Abdi, Pizzolato, et al. 2021; Desmond and Dierickx 2021; Roje et al. 2021). Adopting widely-used guidelines promotes harmonization, streamlines expectations, may be perceived as fairer and may avoid legal uncertainties (and appeal procedures; Desmond and Dierickx 2021). The EU-sponsored SOPs4RI project is developing a freely available online toolbox with *Standard Operating Procedures* for research integrity, which can assist institutions in developing their own (available from www.sops4ri.eu/toolbox/). Finally, yearly evaluations and, if necessary, updates, are recommended (Abdi, Pizzolato, et al. 2021; Forsberg et al. 2018).

Of course, such guidelines are only effective if researchers are aware of them. Currently, however, a significant proportion of researchers is uncertain whether their department has written research integrity guidelines or do not know who to address with related questions (Hofmann and Holm 2019; Yeo-Teh and Tang 2021). Department heads, who are more likely to know the guidelines, often tend to think that their local context is void of research integrity problems (Degn 2020). Furthermore, PhD students and postdocs are even less aware of RCR resources than faculty (Haven et al. 2019; Wells et al. 2014). The latter may be due to faculty's representation in councils, boards and administrative bodies where policies are discussed (Haven et al. 2019) but also indicates that this knowledge does not trickle down. Researchers who are unaware of, for example, the authorship policy of their institution may look elsewhere and rely on policies that are in fact conflicting with their institution's (Yeo-Teh and Tang 2021).

Thus, just as important as adopting clear guidelines is communicating those guidelines to the relevant population (Forsberg et al. 2018; Lerouge and Hol 2020). Supported by a governance and support structure that is sufficiently resourced, dedicated persons should repeatedly emphasize the existence and importance of integrity guidelines. Their use, application and compliance should be repeatedly encouraged, framed in the context of the greater goals, long-term vision, and aspirations of the institute (ALLEA 2017; DuBois et al. 2013; Forsberg et al. 2018; Glerup et al. 2017; Lerouge and Hol 2020; Roje et al. 2021). Furthermore, they should be embedded in a comprehensive institutional policy that stretches as far as managerial support and ethical leadership (Desmond and Dierickx 2021; Glerup et al. 2017; Valkenburg et al. 2021).

Assessing Researchers

While integrity guidelines and rules are a crucial step in fostering an institutional culture of research integrity, they are unlikely to be effective if they are not complemented by changes in how researchers are assessed. That is, they are far more likely to have systemic effects if hiring practices and review, tenure and promotion (RPT) criteria reward researchers that follow these guidelines – and punish researchers that do not.[3] Indeed, at the moment RPT criteria and hiring practices often hamper the prioritization of research integrity by researchers. More precisely, there is broad consensus that evaluation policies structured around the traditional RPT criteria suffer from three main problems (DORA n.d.; Hicks et al. 2015; Wilsdon 2015).

First, journal-level metrics and in particular the Journal Impact Factor (JIF) are inadequate for assessing researchers or individual publications (DORA n.d.; Seglen 1997; Waltman 2016). Because only a few papers are typically responsible for a journal's impact, the metric says little about the average paper in the journal. Still, the JIF is commonly used in RPT criteria, strongly incentivizing researchers to publish in high-impact journals even at the cost of the quality of their work (McKiernan et al. 2019; Rice et al. 2020).

Second, and more generally, traditional RPT criteria tend to value *quantity* over *quality*. In other words, they only track outputs measured in publications or citations and are blind to the content and quality of the research. This incentivizes authorship inflation, cutting corners in research, and cutting research up into *least publishable units*. It also disincentivizes efforts to publish negative results, replication studies and highly innovative research. We – and others – emphasize that this does not imply that publication-metrics should not be used. Rather, they should be used responsibly, and in addition to other criteria for evaluation (Hicks et al. 2015). For example, Ioannidis and Khoury (2014) recommend tracking productivity by counting publications that rank among the best cited for a particular field and

year. Similarly, a variety of new productivity and impact indicators has been proposed (e.g., Hutchins et al. 2016).

Third, traditional RPT criteria are very narrow in the outputs they are willing to acknowledge. Various reviews show that peer reviewed publications and successful grant applications are by far the most popular RPT criteria in use, across disciplines and countries (Alperin et al. 2019; Rice et al. 2020; Snider et al. 2021). Other common criteria are authorship order and journal impact factor, and, to a lesser extent, citation counts (Rice et al. 2020). Moreover, whenever RPT policies do include criteria other than publications and grants, they are often vaguely worded – and therefore harder to assess or enforce (Alperin et al. 2019). While (journal) publications, grants and citations are undoubtedly important, it is clear that academic research also produces valuable outcomes such as societal relevance, communicating to a non-scientific audience, or advisory roles (cf. below). By narrowing down RPT criteria to a small subset of outcomes researchers are incentivized to allocate a disproportionate share of their time and efforts to these outcomes even if it would be better for both society and science if they did not.

A broader assessment scope and alternative criteria that effectively encourage desirable behaviours are necessary –although productivity will likely continue to play an important role in RPT criteria (and probably should). A wide range of proposals has been made in this context (see Moher et al. 2018 for a selection of important sources). Below, we summarize the main recommendations. Before we do, it is important to emphasize that appropriate dedicated human resources, services and (digital) infrastructure supporting researchers in the new behaviours that are expected of them should complement any change in assessment criteria (Forsberg et al. 2018; Lerouge and Hol 2020; Mejlgaard et al. 2020). For example, if researchers are evaluated on the appropriate use of statistical methods, data management and pre-registration, there should be institutional services that provide free guidance and training on these topics (Abdi, Pizzolato, et al. 2021; ALLEA 2017; Bruton et al. 2020; Forsberg et al. 2018; Lerouge and Hol 2020; Van Calster et al. 2021). Furthermore, it should be emphasized that acknowledging a broader range of research outputs into what could be called a 'holistic evaluation' does not mandate every researcher regardless of discipline to pursue all objectives *on top of* a long publication list.

Value and encourage responsible research practices (RRPs): Over the past two decades it has become clear that many published findings cannot be replicated or may be subject to undesirable biases (Ioannidis 2005; Lundh et al. 2017; Open Science Collaboration 2015). Adopting RRPs is crucial to reverse this, including full reporting and publication of negative results, pre-registration of hypotheses, proper data management and storage, maximizing reproducibility of research, high-quality study designs, and proper use of statistical methods (Ioannidis 2014; Ioannidis and Khoury 2014; Kleinert and Horton 2014; Moher et al. 2020; Nosek et al. 2012). RPT criteria should encourage these behaviours, potentially through metrics similar to the ones currently in use (Bouter 2020). Ioannidis (2014), for example, recommends tracking the proportion of researchers' publications that meet one or more of such standards of responsible research.

Value and encourage openness and sharing: RRPs require openness about methods, analysis and hypotheses. These should go hand in hand with further openness: data should be openly accessible as much as possible to allow others to replicate results or to use the data for new research; outputs should be as much as possible open access to maximize societal impact and follow-up research; and software and code should be as much as possible publicly available for others to use and, again, replicate findings (Moher et al. 2018, 2020; Schekman and Patterson 2013; Wilsdon 2015). Currently, however, open access is a criterion in nearly none of the RPT documents surveyed in recent studies (Alperin et al. 2019; Rice et al. 2020). To encourage openness, institutions could make open access

publication mandatory (as is already the case in many institutions) or track the proportion of researchers' publications that come with open data sets, code, or other materials. In addition, they could track how often researchers' data is used in subsequent studies that do not include them as authors (Olfson et al. 2017).

Value and encourage a broad, diverse range of outputs: In addition to peer reviewed publications, (pre-)registrations and open data and code, researchers may contribute to peer review, policy reports, articles in popular press, membership in advisory committees and entrepreneurship. These are all valuable, but often erroneously not recognized in RPT criteria (Moher et al. 2020; Nuffield Council on Bioethics 2014). Furthermore, RPT criteria could recognize and reward service work like managerial obligations and organizational duties, mentoring and managing of students and other researchers, and community outreach (Benedictus et al. 2016; Titus and Ballou 2014). Recognizing these outputs requires ways of tracking and measuring them, which requires a system that enables researchers to record their various outputs (Sandmann et al. 2016). As types of outputs and their relative value differ between disciplines (Sørensen et al. 2021), institutions should be flexible, context-sensitive and inclusive with respect to outputs they include in their RPT criteria.

Value and encourage translation and societal impact: While researchers commonly express intrinsic motivation towards societal impact (O'Meara 2003), they often prioritize traditional academic output. Therefore, to track community engagement and societal outreach, institutions could use various kinds of altmetrics (Schekman and Patterson 2013), such as social media mentions, publications in news outlets, or YouTube videos. A large survey study of Chief Academic Officers suggests that including societal impact and engagement in RPT criteria leads to an increase of their mentioning in applications for promotion (O'Meara 2005). However, like all metrics, altmetrics have important limitations, and care should be taken to avoid that the use of these altmetrics leads to the same problems as traditional metrics (Bornmann 2013).

Evaluate quality rather than quantity: The influential Leiden Manifesto (Hicks et al. 2015) emphasizes that RPT criteria need to shift focus from the quantity of publications to qualitative evaluation by experts. Similarly, DORA (n.d.) emphasizes to focus on the content and quality of research rather than on metrics, for example by reading individual papers (Larivière et al. 2016), having researchers present their most important outputs in a portfolio (Alberts et al. 2015; Editorial 2016) or a narrative account of their main accomplishments (ACUMEN 2014; O'Meara 2003; 2008). Using proxies to assess quality like proportion-based indicators for openness and RRPs as described above is also an option.

Collective Openness

In addition to issuing policies and changing evaluation practices, it is crucial that institutional leaders at all levels exert visible ethical leadership and 'walk the talk' (Desmond and Dierickx 2021; Forsberg et al. 2018; Lerouge and Hol 2020; Valkenburg et al. 2021; Vie 2020; Winchester 2018). Even more important is installing a non-defensive, non-confrontational culture of deliberation in which there is openness to ask questions – about data, procedures and theories – as a *central part of scientific practice* (Forsberg et al. 2018; National Research Council 2002; Zwart and ter Meulen 2019). Gopalakrishna, ter Riet and colleagues (2021) recently suggested that researchers often may subscribe to the norms of scientific integrity but nevertheless digress from those norms due to dissonance in their research environment. Routine discussions of research methods in lab meetings, department or cross-department meetings to discuss manuscripts in preparation, or even 'friendly critique sessions' teaching researchers to defend their research without feeling defensive encourages RCR as a joint

responsibility (Kumar 2010). This philosophy of *collective openness* ensures that anyone from the most senior to the most junior level can voice their concerns even about counter-productive cultural beliefs and that helping each other to stay on track becomes standard practice (Abdi, Pizzolato, et al. 2021; Forsberg et al. 2018; Kumar 2010; Zwart and ter Meulen 2019). Most importantly, in such culture junior researchers can learn about RCR and *good research* from appropriate case models, normative peer pressure becomes positive, and opportunities for misconduct are minimized (ALLEA 2017).

The main recommendations about processes, policies, and procedures are summarized in *Box 1*.

Dealing with Integrity Breaches

Although dealing with breaches of integrity is strictly speaking part of processes, policies, and procedures, we discuss it separately due to its importance: research policy experts and institutional leaders consider effectively dealing with (suspected) breaches of research integrity one of the top 3 priorities of research performing organizations (Labib et al. 2021). In contrast, a significant proportion of researchers does not know where or how to report suspected breaches of integrity. Less than half would feel comfortable doing so without fear of retaliation or personal impact (Moran et al. 2020) and lack confidence in their institution's willingness or ability to investigate the case thoroughly and correctly, to take corrective action, or even to take the report seriously (Vie 2020).

Therefore, institutions should develop – or, to increase harmonization, *adopt* – and repeatedly communicate clear, specific and formal procedures to translate existing guidelines into appropriate, timely and prompt reactions to reports of suspected breaches of integrity (Abdi, Pizzolato, et al. 2021; ALLEA 2017; Anderson et al. 2007; Desmond and Dierickx 2021; Forsberg et al. 2018; Labib et al. 2021; Mejlgaard et al. 2020; Roje et al. 2021). An independent and impartial committee should be installed to investigate complaints (ENERI, ENRIO, and OeAWI 2019). Alternatively, that responsibility can be transferred to an overarching – for example, national, or federal – committee (ALLEA 2017; Forsberg et al. 2018; Mejlgaard et al. 2020). Every step in the procedure should be explicitly described, from the possibility to informally consult a research integrity confidant (strictly confidential but clearly separated from a formal investigation and preferably at the level of the faculty; Forsberg et al. 2018; Hesselmann and Reinhart 2021; Lerouge and Hol 2020; Mejlgaard et al. 2020), to the procedure to file a formal complaint, to how the investigation will be conducted and when and how all those involved will receive updates or be notified of the final outcome. Similarly, the rights and responsibilities of – and measures to protect – both the 'whistle-blower' and the 'accused' should be spelled out (protection can be restricted to good-faith whistle-blowers to protect researchers against unfounded accusations; ALLEA 2017; L. M. Bouter and Hendrix 2017; Forsberg et al. 2018)[4].

Research integrity confidants or ombudspersons, like research ethics advisers, can lower the threshold for potential whistle-blowers to come forward (Forsberg et al. 2018; Mejlgaard et al. 2020). These persons need to have a research background – preferably discipline-specific – and be embedded in researchers' day-to-day activities (Winchester 2018). *Multiple* confidants can avoid the uncomfortable situation that researchers have to approach someone from their own department[5].

Finally, to move beyond the current case-by-case approach and increase congruence in sanctions (Abdi, Nemery, et al. 2021; Hesselmann and Reinhart 2021), institutions should publicly publish anonymized reports about the results of research integrity investigations, including sanctions and measures against whistle-blower retaliation (Forsberg et al. 2018; Gunsalus 2019; Gunsalus et al. 2018; Lerouge and Hol 2020). This increases potential whistle-blowers' confidence in their institutions' ability to safely, promptly, professionally and satisfyingly investigate suspected breaches of integrity

(Hesselmann and Reinhart 2021; Vie 2020). Gunsalus and colleagues (2018) proposed a checklist for research integrity investigations, which also covers the appropriateness and completeness of post-hoc reporting.

Institutions may be inclined to cover up breaches of integrity of their researchers or urge potential whistle-blowers to drop their allegations. A formal investigation into the behaviour of a single researcher may slow down a whole research team, may bring reputational damage, or large financial impact when investigations trail or result in reduction or retraction of research funding. Nevertheless, leaders who fail to act on signals that research integrity is compromised elicit frustration and shy away potential whistle-blowers (Vie 2020). In contrast, a formal, effective response to allegations of breaches of integrity can potentially *prevent* future breaches of research integrity (Abdi, Pizzolato, et al. 2021; Kumar 2010), especially because it fits into a culture of collective openness. Kumar (2010 p. 58) even suggests shifting the focus from punishing the guilty towards repairing the damage done and looking inwards: 'the important issue is not who committed the error, but how the safe-guards failed' (see also Zwart and ter Meulen 2019).

The main recommendations about how institutions should deal with breaches in research integrity are summarized in *Box 2*.

Education and Training

Funding agencies are increasingly mandating formal training addressing breaches of integrity and promoting research integrity in (junior) researchers (Bruton et al. 2020; Forsberg et al. 2018; Kumar 2010; Watts, Medeiros, et al. 2017). In the US both the National Institutes of Health (NIH) and the National Science Foundation (NSF) require institutions receiving their funds to provide ethics training during every career stage (National Institutes of Health 2009; National Science Foundation 2009; Phillips et al. 2018) and minimally every four years (National Institutes of Health 2009). In Europe, the European Molecular Biology Organization (EMBO n.d.) requires its long-term fellows and young investigators to follow a course on research integrity as does, for example, the German Deutsche Forschungsgemeinschaft (2019). Similarly, the ALLEA code by LERU recommends 'Researchers across the entire career path, from junior to the most senior level, undertake training in ethics and research integrity' (2017 p. 5) but does not set it as a requirement, nor does it provide specifics about RCR instruction[6].

Two meta-analytic reviews (Antes et al. 2009; Watts, Medeiros, et al. 2017) concluded that RCR training has clearly become more effective in the past decade (although Marusic and colleagues (2016) argue that the available evidence is at best inconclusive and Anderson and colleagues (2007) reported little or no relationship between ethics training and self-reported breaches of integrity). A shift is taking place from *education* – the transfer of knowledge about codes – towards *training*: teaching skills to solve complex ethical dilemmas. Indeed, knowledge of the principles and codes of RCR is clearly beneficial and maybe even a prerequisite for successful training (Yeo-Teh and Tang 2021). However, rules and codes lack effective guidance to navigate the complexities of modern research settings and many situations are not covered by rules and codes (Anderson et al. 2007; Mulhearn et al. 2017; Tang and Lee 2020; Tomić et al. 2021; Watts, Mulhearn, et al. 2017; Yeo-Teh and Tang 2021). Particularly junior scientists may struggle to deal with dynamic emotional, interpersonal and situational dilemmas and need to develop, practice and train metacognitive reasoning strategies such as problem solving, analysing constraints, forecasting, and analysing other stakeholders' views (DuBois et al. 2013;

Mumford et al. 2008; Pennock and O'Rourke 2017; Tang and Lee 2020; Tomić et al. 2021; Watts, Medeiros, et al. 2017; Watts, Mulhearn, et al. 2017; Yeo-Teh and Tang 2021).

Importantly, other educational goals may be equally important. Laboratory management skills and stress and time management skills can alleviate the pressure that leads to corner cutting (Antes 2014). Interpersonal skills, intercultural sensitivity and conflict resolution skills can defuse conflicts before they become problematic, can facilitate collaboration, and can assist students in more collectively focused cultures where conflicts with authority figures have to be avoided to defend their rights not to be exploited (Tang and Lee 2020; Vie 2020; Yeo-Teh and Tang 2021). Furthermore, strategies to achieve peer support, counselling advice and awareness of the proper channels to report breaches of integrity may lower the threshold for potential whistle-blowers to come forward (Yeo-Teh and Tang 2021). Explicitly preparing PhD students for a career outside academia by teaching transferable soft skills may be a strong preventive measure in the light of the increasingly skewed distribution of PhD and faculty positions.

Recommendations for RCR Skill Training

RCR skill training is preferably delivered in a stand-alone fashion, not embedded in existing courses (Antes et al. 2009). Training content should *either* be discipline-specific *or* generic as mixed approaches possibly contribute to an unfocused, open-ended discussion of ethics and perceptions of irrelevance (Mulhearn et al. 2017; Watts, Medeiros, et al. 2017). However, e*xamples* from other disciplines may create awareness of differences in codes, aims and methods, which is essential in the light of increasing interdisciplinarity and collaborative spirit (Lerouge and Hol 2020; Mejlgaard et al. 2020; Pennock and O'Rourke 2017). Depending on the discipline, particular topics may lead to better learning: training about FFP resorts better effects in the biomedical sciences than training about authorship or data management, whereas in the social sciences moral philosophy and ethical guidelines resort the largest effects (Watts, Medeiros, et al. 2017). Importantly, regardless of course content, formulating the learning objectives in terms of professional development provides scope and purpose and keeps participants motivated (Antes 2014; Antes et al. 2009).

Although one could assume that mixed audiences facilitate awareness of perspectives beyond one's own (Antes 2014; Tang and Lee 2020), Watts, Medeiros and colleagues (2017) find substantially higher effect sizes for training delivered to field-specific audiences. Still, trainings with trainees from various backgrounds *within* a scientific discipline may be particularly fruitful as their differing stances and practicalities will enrich discussions (Pennock and O'Rourke 2017; Tang and Lee 2020).

Whether to mix experience levels is an open question. The different ideas and experiences researchers in different career stages contribute will stimulate the exchange of ideas, enrich discussions, and stimulate learning (Antes 2014; Pennock and O'Rourke 2017). Contrarily, in particular more junior researchers may be overwhelmed by senior researchers' complex problems and, obviously, different career stages require different skills. Furthermore, training junior researchers together with (their) superiors may create a power imbalance – and potential problems with confidentiality (Pennock and O'Rourke 2017).

Regarding the format, face-to-face training allows intense interaction between participants and between participants and trainers, as well as flexibly responding to what comes up during training. Furthermore, knowledge of more complex issues may be best facilitated in social interaction. Therefore, particularly training ethical problem-solving skills is preferably done in face-to-face workshops. On the other hand, online training allows more flexible processing in terms of planning and

speed of progress. Therefore, it is particularly suited to teach compliance-based content such as codes of conduct. Combining the advantages of both approaches into a mixed approach results in an attractive and effective program (Todd et al. 2017; Watts, Medeiros, et al. 2017). Importantly though, the NIH explicitly states that online-only training is insufficient for recipients of NIH funding (NIH 2009).

Finally, RCR training is most effective when it is delivered by multiple – preferably more than two – expert trainers that are present for all sessions (Todd et al. 2017; Watts, Medeiros, et al. 2017; Watts, Mulhearn, et al. 2017), spaced in time and not delivered in one massed session (Antes et al. 2009; Phillips et al. 2018). Frequent and highly interactive practice opportunities such as debates, small group discussion, role-plays and teaching others not only keep participants engaged but also make them better equipped to address problems in real life (Antes 2014; Antes et al. 2009; Berling et al. 2019; Koterwas et al. 2021; Tang and Lee 2020; Tomić et al. 2021; Watts, Medeiros, et al. 2017; Watts, Mulhearn, et al. 2017; Yeo-Teh and Tang 2021). Finally, voluntary training programs clearly demonstrate larger effects than mandatory programs (Antes et al. 2009; Watts, Medeiros, et al. 2017), but as already mentioned funders increasingly mandate RCR training.

The Sensemaking Approach and Case-Based Training

A case-based approach has proven particularly effective to increase researchers' ethical knowledge and ethical decision-making (Antes et al. 2012; Johnson et al. 2012; Mumford et al. 2008; Watts, Medeiros, et al. 2017; Watts, Mulhearn, et al. 2017; the Appendix lists some examples of online resources with relevant cases). A host of studies has investigated which aspects make the case-based approach most beneficial. Especially longer cases focusing on process-oriented reflection about real-life professional ethical problems with low to moderate complexity and affectivity appear to support instructional effectiveness. Furthermore, cases with a positive outcome are preferred over cases with a negative outcome and mixed outcomes. The former lead to better identification of critical causes, resources and opportunities, better forecast quality, and increased application of the case material to other problems. Cases with a negative outcome may be (too) threatening and inhibit process-oriented reflection – as do cases with high emotional content (Antes et al. 2012; Johnson et al. 2012; Mulhearn et al. 2017; Watts, Medeiros, et al. 2017; Watts, Mulhearn, et al. 2017). However, as DuBois and colleagues (2013) point out, the most salient (i.e. mediatized) and engaging cases are ethics scandals that tend to have negative outcomes. Therefore, well-known cases can serve as examples – emphasizing the processes involved, but case-based training should proceed with less emotionally laden cases.

A Virtue-Based Approach

Recently, a virtue-based approach to RCR training has been introduced with substantial effects (Berling et al. 2019; Editorial 2019; Peels et al. 2019; Pennock and O'Rourke 2017; Todd et al. 2017). It focuses on developing good character traits or 'behavioural dispositions' that may predispose scientists to act responsibly and exemplary. Whereas the compliance approach promotes obeying externally imposed rules – for example, you should not fabricate data because getting caught will get you punished – the virtue-based approach treats research integrity as something intrinsic to sound scientific practice: the very idea of fabricating data violates what it means to be a scientist (Berling et al. 2019; Pennock and O'Rourke 2017).

The virtue-based approach is particularly valuable to teach scientists what to do (instead of merely being aware of the conflict) when two or more scientific norms come into conflict (Pennock and O'Rourke 2017; Tang and Lee 2020). Furthermore, it may cultivate sensitivity for, and understanding of, the different but equally reasonable decisions of other stakeholders and researchers from other disciplines or cultural backgrounds (Peels et al. 2019).

Virtues are best developed via socialization in which mentors demonstrate virtuous behaviour through their actions (cf. 'Collective openness' above; Tomić et al. 2021). Additionally, Pennock and O'Rourke (cf. also Berling et al. 2019; 2017) have developed a brief workshop in which particular scientific virtues are explored using prompts to stimulate thoughtful dialogue and discussion. It can be flexibly used with an audience of graduate students, postdocs, and faculty, is preferred over more traditional RCR training by participants, and can easily be implemented to augment standard RCR training (Berling et al. 2019). An 'open source' virtue-based RCR training was developed within the European Union's Horizon 2020 research program (available online at https://embassy.science/wiki/Training).

Ample Room for Improvement

There still is ample room for improvement. Less than half of the NSF ethics training policies of the 'very high research activity' institutes investigated by Phillips and colleagues (2018) incorporated even some of the best practices known at the time of implementation. Strikingly, they found that the majority of universities did not require *any* face-to-face interaction, four out of five had plans that could be entirely met by online-only training and none required involvement of PIs – although some encouraged the latter. In freely accessible research integrity educational resources from the US and Europe, discipline-specific resources are scarce and primarily tailored for the Biomedical sciences. Furthermore, proactive participation or active learning is required in less than half of those resources (Pizzolato et al. 2020). In their review of educational RCR material from European research universities, Abdi, Pizzolato and colleagues (2021) established that *all* universities provided once-only training, mostly less than eight hours. Surprisingly, almost all programs in their sample emphasized RCR guidelines, procedures addressing research integrity breaches, and designated RI officers, which apparently contradicts the deficient knowledge of those resources discussed above.

Research integrity experts agree that the current approach to RCR training should be updated to match new developments such as open science, reproducibility, and environmental responsibility (Lerouge and Hol 2020; Tomić et al. 2021). Preferably, a systematic approach is used, starting from a needs assessment determining what should be learned and incorporating the input of professionals with relevant knowledge of the research practice (Antes 2014; McIntosh et al. 2018). Needs should be formulated in terms of awareness, knowledge, skills, attitudes and/or behaviour using active verbs (e.g., 'describe', 'identify', 'analyse', 'reflect',… ; Antes 2014; Desmond and Dierickx 2021; Marusic et al. 2016; Pizzolato et al. 2020). Subsequently, learning materials and procedures to bridge the gap between the current and the desired situation should be planned and created (Antes 2014; McIntosh et al. 2018). To prevent institutions having to reinvent the wheel and to aid institutions with limited resources, creating repositories of trainings with proven effectiveness has been suggested (Abdi, Pizzolato, et al. 2021; European Science Foundation 2010) and various resources are already available online – such as those of the Online Ethics Center for Engineering and Science (Available from https://onlineethics.org/resources), the American Office of Research Integrity (n.d.), and the European Network of Research Ethics and Research Integrity and the European Network of Research Integrity Offices (ENERI-ENRIO n.d.). Pizzolato and colleagues (2020) assessed and categorized 237 freely available online RCR educational resources from across the world using 21 carefully designed criteria,

facilitating finding the right tool at the right time (their inventory, ranging from case studies to webinars to online trainings is available online: https://embassy.science/wiki/Resources).

The main recommendations about education and training are summarized in *Box 3*.

Knowledge Is Power

An important caveat to most of the recommendations in this paper is that there exists little or no supporting empirical evidence. In fact, while there is broad consensus that more empirical research on research integrity is necessary to effectively foster research integrity (e.g., Aubert Bonn and Pinxten 2019; Bouter 2020; National Research Council 2002; Rice et al. 2020), only a small proportion of papers currently published on the topic are empirical (Aubert Bonn and Pinxten 2019). Institutions, in particular those focused on research, are in an ideal position to perform such research, as they typically have easy access to data as well as the necessary in-house expertise. Hence, a final and crucial recommendation for institutions is to closely monitor and empirically investigate the effectiveness of their research integrity policies.

Monitoring

First, institutions should closely monitor their rules, guidelines and policies (Abdi, Pizzolato, et al. 2021; Forsberg et al. 2018). Not only is there high uncertainty about which measure are effective, effectiveness is also likely to be context-dependent and may differ between research fields. As with any policy – particularly those requiring behaviour change (Michie et al. 2011) – closely connecting a top-down research integrity policy with the everyday reality at the level of the lab or research group may be challenging (Bouter 2020; Moran et al. 2020; Zwart and ter Meulen 2019). Therefore, baseline conditions need to be established. Similarly, those aspects of the research climate in need of improvement should be specified in terms of awareness, knowledge and behaviour (in particular solid evidence on the uptake of RRPs is missing; Gopalakrishna, Wicherts, et al. 2021; but see below), but also in terms of attitudes and motivation (Bouter 2020; Gopalakrishna, ter Riet, et al. 2021; Martinson, Thrush, and Crain 2013). Otherwise, even interventions that empower researchers may be met with indifference or ritual compliance or encounter obstruction and hostility up to the level of management (Bruton et al. 2020; Degn 2020).

Collect and Share Data

Second, institutions should collect data related to research integrity, and make it available to the scientific community. Most institutions already have a range of rules, guidelines and policies to promote research integrity in place. This means that institutions can easily collect data to facilitate research on research integrity by tracking the prevalence of FFP and QRPs, but also by monitoring RCR such as how often researchers mention open access, pre-registrations, or open data sets in their tenure or promotion files, whether this differs between generations or stages seniority, and how research fields differ in the kinds of societal impact they report and in the outputs they generate.

Similarly, institutions can relatively easily collect and share assessments of their research culture, for example using The Survey of Research Organization Climate that assesses several dimensions of the research climate at the level of the institution and the researcher's primary department (SOuRCe;

Martinson et al. 2013; Wells et al. 2014). Its validity, reliability and sensitivity to disciplinary field and academic rank have been established in different settings and benchmark data for the US and The Netherlands have been published (Crain, Martinson, and Thrush 2013; Haven et al. 2019; Wells et al. 2014). In their large-scale, comprehensive study in The Netherlands, Gopalakrishna and colleagues (Gopalakrishna, ter Riet, et al. 2021; Gopalakrishna, Wicherts, et al. 2021) recently used twelve psychometrically tested scales to assess potential drivers of research integrity and related those to prevalence estimates of QRPs and RRPs. They observed – and emphasize the need to understand – important differences between scientific disciplines both in the prevalence of QRPs and RRPs and in their relationship with the explanatory variables, suggesting researchers from different disciplines may need to be rewarded differently (cf. '*Assessing researchers*' above).

Making these data and results available to the larger community will facilitate larger, comparative studies. However, no matter how important these results are in understanding how to foster research integrity, they remain abstract and lack practical actionability. Therefore, we may need to resort back to surveys, focus groups, and Delphi studies to gain detailed, in depth-knowledge (Bouter 2020; Labib et al. 2021; Mejlgaard et al. 2020). Two surveys have recently attempted this. The first attempted to gain insight in lab practices and cultures in the biomedical sciences (Van Noorden 2018; survey questions are available as supplemental material). The second is a large-scale survey targeting UK researchers' thoughts about and issues with the culture they work in (Moran et al. 2020; survey questions are available as extended data). Both surveys are considerably long – the UK survey included up to 70 questions depending on respondent route. Moreover, because application in different settings or institutes may necessitate adaptations, results can be difficult to compare (Moran et al. 2020). Nonetheless, precisely these surveys may result in a bottom-up (or 'crowd-sourced') inventory of easily implementable changes, best practices, and interventions that have proven their worth in the research setting under investigation (Gopalakrishna, ter Riet, et al. 2021; Gopalakrishna, Wicherts, et al. 2021; Lerouge and Hol 2020).

Collect and Share Empirical Evidence

Finally, and related to the former two points, institutions are uniquely well placed to conduct experiments (A/B designs, pre-post designs, …) on research integrity guidelines and policies. Institutions could implement any of the broad range of recommendations from this paper and assess their efficacy before implementing them across the institution. For example, institutions could evaluate the implementation of new RPT criteria in selected departments and its effect on outputs, competitiveness and researchers' well-being. Ultimately, the results of these experiments should also be shared across the scientific community (Committee on Responsible Science et al. 2017; Lerouge and Hol 2020; Martinson, Thrush, and Crain 2013).

A prime example of experiments that institutions could conduct concerns the efficacy of RCR training (Antes 2014; Forsberg et al. 2018; Marusic et al. 2016; McIntosh et al. 2018; Mumford et al. 2015). Because institutions are, in most cases, organizing trainings themselves they are optimally positioned to design studies and collect empirical data about the effectiveness of trainings and about how they are evaluated by trainees. Nonetheless, although pre-post designs with control groups provide the most conclusive evidence for program effectiveness, a multitude of designs and instruments is used and often only one or two outcome criteria are used (e.g., Antes et al. 2009; Marusic et al. 2016; McIntosh et al. 2018; Mumford et al. 2015; Steele et al. 2016; Zollitsch et al. 2021). As a result, the available evidence is often unreliable, with a high risk of bias (Marusic et al. 2016; Zollitsch et al. 2021). Therefore, Mumford and colleagues (2015) proposed a 'multilevel approach' in which outcome

measures are tailored to the training objectives, i.e. *awareness, knowledge, skills, attitudes and/or behaviour* (see Abdi, Fieuws, et al. 2021 for a preliminary questionnaire assessing knowledge, attitudes, and behaviour; Marusic et al. 2016; McIntosh et al. 2018) and evidence of different strength is combined into a balanced, global evaluation. For example, student ratings or answers to open-ended questions fall short of hard evidence. Still, they can lead to simple adjustments or the detection of underperforming individual trainers and in particular ratings of content relevance and course satisfaction have been shown to be related to training effectiveness (McIntosh et al. 2018; Turner et al. 2018). Furthermore, Mumford and colleagues (2015) suggest monitoring institutional outcomes such as the number of questions to research confidants or research integrity officers, how soon (suspected) breaches of integrity are reported, and the duration of formal investigations (see also Marusic et al. 2016).

Again, some researchers have suggested that a database of evaluation studies would allow identifying the most effective RCR training interventions and their most effective target population and delivery methods through meta-analysis (Marusic et al. 2016; Mumford et al. 2015; Phillips et al. 2018; Steele et al. 2016; Zollitsch et al. 2021). Todd and colleagues (2017) add that multiple process-based variables have not yet been used and are hence not included in meta-analyses – so we do not even know whether the most indicative measures are used.

The main recommendations about monitoring and evaluation are summarized in *Box 4*.

Conclusion

Although there has clearly been a recent shift from punishing breaches of integrity towards the promotion of the responsible conduct of research across disciplines, it is clear that there is a lot we currently do not know. It may be reasonable to expect a substantial increase of knowledge in the near future, as the field of research-on-research is growing. Still, increased investments in educating and training researchers in research integrity are urgently needed, as are better assessments of the effectiveness of training – and of the specific components that generate the largest effects. Furthermore, sharing policies, procedures, and processes may drive their adoption and harmonization across institutes.

Evidently, institutions depend on government funding and university rankings, both of which often reinforce some of the incentives that, as we will discuss below, are detrimental to research integrity. Hence, the agency of institutions is limited because fostering research integrity also requires higher-level change involving governments, funders and other stakeholders. However, this should be no excuse for institutions not to improve those aspects within their control and at the very least universities should meet the same requirements they set for their employees when it comes to research integrity. Whether promoting research integrity regardless of a potential hit in the rankings is laudable, audacious or reckless can be subject to debate; it is a bold step nevertheless (Editorial 2017).

Furthermore, institutions can exert pressure on, or join hands with, publishers and scientific journals to loosen the focus on novel and positive results or correct the scientific record when breaches of integrity are detected (Aubert Bonn and Pinxten 2021a; Bruton et al. 2020; Gopalakrishna, ter Riet, et al. 2021; Gopalakrishna, Wicherts, et al. 2021; Labib et al. 2021; Wager and Kleinert 2021). They can challenge publishers directly to embrace open science or start publishing preregistered reports. They can exert pressure more indirectly by negotiating subscriptions or influencing faculty serving in editorial boards or as peer reviewers to promote responsible practices and guidelines. Using university-

based portals to publish negative results and replication studies may convince publishers of their added value (Lerouge and Hol 2020). Similarly, institutions can encourage funders to use responsible metrics or fair distribution mechanisms when selecting and evaluating researchers and research (De Peuter and Conix 2021), or to increase budgets for RCR training and open science (e.g., open access publishing fees; Editorial 2017). Additionally, institutions could seek to harmonize RCR training requirements in a joint effort with funders (Labib et al. 2021).

This could generate a virtuous cycle of collective openness, genuinely changing how science is performed and pushing a culture of research integrity to a global phenomenon in which the responsible conduct of research is no longer a choice, but the reality. Ideally, this will push other stakeholders – e.g., funders, publishers – towards more responsible practices as well.

ENDNOTES

1. Schneider and colleagues (2013 p. 361) offer a historical review of the concepts of climate and culture in organizational research, defining climate as 'the meanings people attach to interrelated bundles of experiences they have at work' and culture as 'the basic assumptions about the world and the values that guide life in organizations'. Nowadays, the distinction has faded and both terms are often used interchangeably. Also see Valkenburg and colleagues (2021) for an advanced discussion of *culture* versus *practice.*
2. Cf. also The National Academies of Sciences, Engineering, and Medicine. *On Being a Scientist: A Guide to Responsible Conduct in Research: Third Edition* (2009), available from https://www.nationalacademies.org/our-work/on-being-a-scientist-a-guide-to-responsible-conduct-in-research-third-edition; and The National Academies of Sciences, Engineering, and Medicine. *Fostering Integrity in Research*. 2017. Washington, DC: The National Academies Press. https://doi.org/10.17226/21896.
3. RPT criteria do not typically include assessment for research funding provided by institutions. However, our arguments about RPT criteria equally apply to institutional funding of research.
4. The new EU directive 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law. PE/78/2019/REV/1) extends the protection against retaliatory measures to facilitators: 'a natural person who assists a reporting person in the reporting process in a work-related context, and whose assistance should be confidential', which means that also confidants and members of committees should be protected.
5. The authors are grateful to Shila Abdi for this suggestion.
6. Although *ethics* and *integrity* can be differentiated (see e.g., Marusic et al. 2016; Valkenburg et al. 2021; Zollitsch et al. 2021) and may require separate training, there is substantial overlap and in particular in the early stages of *integrity* education and training development, the terms have been used interchangeably in the literature. We use 'RCR training' as an umbrella term.

REFERENCES

Abdi, S., Fieuws, S., Nemery, B., et al. (2021). 'Do We Achieve Anything by Teaching Research Integrity to Starting PhD Students?', *Humanities and Social Sciences Communications*, 8/1: 232. DOI: 10.1057/s41599-021-00908-5

Abdi, S., Nemery, B., and Dierickx, K. (2021). 'What Criteria are Used in the Investigation of Alleged Cases of Research Misconduct?', *Accountability in Research*, 1–23. DOI: 10.1080/08989621.2021.1973894

Abdi, S., Pizzolato, D., Nemery, B., et al. (2021). 'Educating PhD Students in Research Integrity in Europe', *Science and Engineering Ethics*, 27/1: 5. DOI: 10.1007/s11948-021-00290-0

ACUMEN. (2014). *Final Report*. <https://cordis.europa.eu/docs/results/266/266632/final1-acumen-final-report-29-april-2014.pdf> accessed 10 Dec 2021.

Alberts, B., Cicerone, R. J., Fienberg, S. E., et al. (2015). 'Self-Correction in Science at Work', *Science*, 348/6242: 1420–2. DOI: 10.1126/science.aab3847

ALLEA. (2017). *The European Code of Conduct for Research Integrity*. ALLEA – All European Academies. <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf> accessed 4 Jun 2020.

Alperin, J. P., Muñoz Nieves, C., Schimanski, L. A., et al. (2019). 'How Significant are the Public Dimensions of Faculty Work in Review, Promotion and Tenure Documents?', *eLife*, 8: e42254. DOI: 10.7554/eLife.42254

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*, 82/9: 853–60. DOI: 10.1097/ACM.0b013e31812f764c

Antes, A. L. (2014). 'A Systematic Approach to Instruction in Research Ethics', *Accountability in Research*, 21/1: 50–67. DOI: 10.1080/08989621.2013.822269

Antes, A. L., Murphy, S. T., Waples, E. P., et al. (2009). 'A Meta-Analysis of Ethics Instruction Effectiveness in the Sciences', *Ethics & Behavior*, 19/5: 379–402. DOI: 10.1080/10508420903035380

Antes, A. L., Thiel, C. E., Martin, L. E., et al. (2012). 'Applying Cases to Solve Ethical Problems: The Significance of Positive and Process-Oriented Reflection', *Ethics & Behavior*, 22/2: 113–30. DOI: 10.1080/10508422.2012.655646

Aubert Bonn, N. and Pinxten, W. (2019). 'A Decade of Empirical Research on Research Integrity: What Have We (Not) Looked At?', *Journal of Empirical Research on Human Research Ethics*, 14/4: 338–52. DOI: 10.1177/1556264619858534

Aubert Bonn, N. and Pinxten, W. (2021a). 'Rethinking Success, Integrity, and Culture in Research (Part 2) — a Multi-Actor Qualitative Study on Problems of Science', *Research Integrity and Peer Review*, 6: 3. DOI: 10.1186/s41073-020-00105-z

Aubert Bonn, N. and Pinxten, W. (2021b). 'Rethinking Success, Integrity, and Culture in Research (Part 1) — a Multi-Actor Qualitative Study on Problems of Science', *Research Integrity and Peer Review*, 6/1: 1. DOI: 10.1186/s41073-020-00104-0

Benedictus, R., Miedema, F., and Ferguson, M. W. J. (2016). 'Fewer Numbers, Better Science', *Nature*, 538/7626: 453–5. DOI: 10.1038/538453a

Berling, E., McLeskey, C., O'Rourke, M., et al. (2019). 'A New Method for a Virtue-Based Responsible Conduct of Research Curriculum: Pilot Test Results', *Science and Engineering Ethics*, 25/3: 899–910. DOI: 10.1007/s11948-017-9991-2

Bornmann, L. (2013). 'What is Societal Impact of Research and How Can it Be Assessed? A Literature Survey', *Journal of the American Society for Information Science and Technology*, 64/2: 217–33. DOI: 10.1002/asi.22803

Bouter, L. (2020). 'What Research Institutions Can Do to Foster Research Integrity', *Science and Engineering Ethics*, 26/4: 2363–9. DOI: 10.1007/s11948-020-00178-5

Bouter, L. M. (2015). 'Commentary: Perverse Incentives or Rotten Apples?', *Accountability in Research*, 22/3: 148–61. DOI: 10.1080/08989621.2014.950253

Bouter, L. M. and Hendrix, S. (2017). 'Both Whistleblowers and the Scientists They Accuse Are Vulnerable and Deserve Protection', *Accountability in Research*, 24/6: 359–66. DOI: 10.1080/08989621.2017.1327814

Bouter, L. M., Tijdink, J., Axelsen, N., et al. (2016). 'Ranking Major and Minor Research Misbehaviors: Results from a Survey among Participants of Four World Conferences on Research Integrity', *Research Integrity and Peer Review*, 1/1: 17. DOI: 10.1186/s41073-016-0024-5

Bruton, S. V., Medlin, M., Brown, M., et al. (2020). 'Personal Motivations and Systemic Incentives: Scientists on Questionable Research Practices', *Science and Engineering Ethics*, 26/3: 1531–47. DOI: 10.1007/s11948-020-00182-9

Casadevall, A. and Fang, F. C. (2012). 'Reforming Science: Methodological and Cultural Reforms', *Infection and Immunity*, 80/3: 891–6. DOI: 10.1128/IAI.06183-11

Committee on Responsible Science, Committee on Science, Engineering, Medicine, and Public Policy, Policy and Global Affairs, and National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering Integrity in Research*. Washington, D.C.: National Academies Press. DOI: 10.17226/21896

Crain, A. L., Martinson, B. C., and Thrush, C. R. (2013). 'Relationships Between the Survey of Organizational Research Climate (SORC) and Self-Reported Research Practices', *Science and Engineering Ethics*, 19/3: 835–50. DOI: 10.1007/s11948-012-9409-0

CRediT. (n.d.). *Contributor Roles Taxonomy*. <https://credit.niso.org> accessed 26 Oct 2021.

De Peuter, S. and Conix, S. (2021). 'The Modified Lottery: Formalizing the Intrinsic Randomness of Research Funding', *Accountability in Research*, 1–22. DOI: 10.1080/08989621.2021.1927727

Degn, L. (2020). 'Integrating Integrity: The Organizational Translation of Policies on Research Integrity', *Science and Engineering Ethics*, 26/6: 3167–82. DOI: 10.1007/s11948-020-00262-w

Desmond, H. and Dierickx, K. (2021). 'Research Integrity Codes of Conduct in Europe: Understanding the Divergences', *Bioethics*, 35/5: 414–28. DOI: 10.1111/bioe.12851

Deutsche Forschungsgemeinschaft. (2019). *Guidelines for Safeguarding Good Research Practice. Code of Conduct.*

<https://www.dfg.de/en/research_funding/principles_dfg_funding/good_scientific_practice/index.html> accessed 20 Dec 2021.

DORA. (n.d.). *San Francisco Declaration on Research Assessment*. <https://sfdora.org/read/> accessed 7 Jan 2022.

DuBois, J. M., Anderson, E. E., Chibnall, J., et al. (2013). 'Understanding Research Misconduct: A Comparative Analysis of 120 Cases of Professional Wrongdoing', *Accountability in Research*, 20/5–6: 320–38. DOI: 10.1080/08989621.2013.822248

Editorial. (2016). 'Time to Remodel the Journal Impact Factor', *Nature*, 535/7613: 466–466. DOI: 10.1038/535466a

Editorial. (2017). 'Integrity Starts with the Health of Research Groups', *Nature*, 545/7652: 5–6. DOI: 10.1038/545005b

Editorial. (2019). 'Research Integrity is Much More than Misconduct', *Nature*, 570/7759: 5–5. DOI: 10.1038/d41586-019-01727-0

EMBO. (n.d.). *Research Integrity*. <https://www.embo.org/science-policy/research-integrity> accessed 12 Feb 2020.

ENERI-ENRIO. (n.d.). *Training Resources*. <http://www.enrio.eu/resources/?cat=4> accessed 7 Aug 2021.

ENERI (European Network of Research Ethics and Research Integrity), ENRIO (European Network of Research Integrity Offices), and OeAWI (Austrian Agency for Research Integrity). (2019). *ENRIO Handbook. Recommendations for the Investigation of Research Misconduct.* <http://www.enrio.eu/newsactivities/resources/recommendations-for-the-investigation-of-research-misconduct-byenrio-2019/inv-handbook_enrio_web_final/> accessed 25 Apr 2021.

European Science Foundation. (2010). *Fostering Research Integrity in Europe. A report by the ESF Member Organisation Forum on Research Integrity*. <https://www.esf.org/fileadmin/user_upload/esf/ResearchIntegrity_Report2011.pdf> accessed 17 Mar 2020.

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

Fang, F. C. and Casadevall, A. (2012). 'Reforming Science: Structural Reforms', *Infection and Immunity*, 80/3: 897–901. DOI: 10.1128/IAI.06184-11

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*, 24/4: 1023–34. DOI: 10.1007/s11948-018-0034-4

Glerup, C., Davies, S. R., and Horst, M. (2017). '"Nothing Really Responsible Goes on Here": Scientists' Experience and Practice of Responsibility', *Journal of Responsible Innovation*, 4/3: 319–36. DOI: 10.1080/23299460.2017.1378462

Godecharle, S., Nemery, B., and Dierickx, K. (2013). 'Guidance on Research Integrity: no Union in Europe', *The Lancet*, 381/9872: 1097–8. DOI: 10.1016/S0140-6736(13)60759-X

Gopalakrishna, G., ter Riet, G., Vink, G., et al. (2021). 'Prevalence of Questionable Research Practices, Research Misconduct and their Potential Explanatory Factors: A Survey among Academic

Researchers in The Netherlands' (preprint), *MetaArXiv*. <https://osf.io/vk9yt> accessed 5 Nov 2021. DOI: 10.31222/osf.io/vk9yt

Gopalakrishna, G., Wicherts, J. M., Vink, G., et al. (2021). 'Prevalence of Responsible Research Practices and their Potential Explanatory Factors: A Survey among Academic Researchers in The Netherlands' (preprint), *MetaArXiv*. <https://osf.io/xsn94> accessed 5 Nov 2021. DOI: 10.31222/osf.io/xsn94

Gunsalus, C. K. (2019). 'Make Reports of Research Misconduct Public', *Nature*, 570/7759: 7–7. DOI: 10.1038/d41586-019-01728-z

Gunsalus, C. K., Marcus, A. R., and Oransky, I. (2018). 'Institutional Research Misconduct Reports Need More Credibility', *JAMA*, 319/13: 1315. DOI: 10.1001/jama.2018.0358

Hastings, R. C., Labib, K., Lechner, I., et al. (2022). 'Guidance on Research Integrity Provided by European Discipline-Specific Learned Societies: A Scoping Review' (preprint), *MetaArXiv*. <https://osf.io/preprints/metaarxiv/kn5y9/> accessed 15 Jan 2022. DOI: 10.31222/osf.io/kn5y9

Haven, T. L., 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*, 14/1: e0210599. DOI: 10.1371/journal.pone.0210599

Haven, T., Tijdink, J., Martinson, B., et al. (2021). 'Explaining Variance in Perceived Research Misbehavior: Results from a Survey among Academic Researchers in Amsterdam', *Research Integrity and Peer Review*, 6/1: 7. DOI: 10.1186/s41073-021-00110-w

Hesselmann, F. and Reinhart, M. (2021). 'Cycles of invisibility: The limits of transparency in dealing with scientific misconduct', *Social Studies of Science*, 51/3: 414–38. DOI: 10.1177/0306312720975201

Hicks, D., Wouters, P., Waltman, L., et al. (2015). 'Bibliometrics: The Leiden Manifesto for Research Metrics', *Nature*, 520/7548: 429–31. DOI: 10.1038/520429a

Hofmann, B. and Holm, S. (2019). 'Research Integrity: Environment, Experience, or Ethos?', *Research Ethics*, 15/3–4: 1–13. DOI: 10.1177/1747016119880844

Holcombe, A. O. (2019). 'Contributorship, Not Authorship: Use CRediT to Indicate Who Did What', *Publications*, 7/3: 48. DOI: 10.3390/publications7030048

Hutchins, B. I., Yuan, X., Anderson, J. M., et al. (2016). 'Relative Citation Ratio (RCR): A New Metric That Uses Citation Rates to Measure Influence at the Article Level', *PLOS Biology*, 14/9: e1002541. DOI: 10.1371/journal.pbio.1002541

Ioannidis, J. P. A. (2005). 'Why Most Published Research Findings Are False', *PLoS Medicine*, 2/8: e124. DOI: 10.1371/journal.pmed.0020124

Ioannidis, J. P. A. (2014). 'How to Make More Published Research True', *PLoS Medicine*, 11/10: e1001747. DOI: 10.1371/journal.pmed.1001747

Ioannidis, J. P. A. and Khoury, M. J. (2014). 'Assessing Value in Biomedical Research: The PQRST of Appraisal and Reward', *JAMA*, 312/5: 483. DOI: 10.1001/jama.2014.6932

Johnson, J. F., Bagdasarov, Z., Connelly, S., et al. (2012). 'Case-Based Ethics Education: The Impact of Cause Complexity and Outcome Favorability on Ethicality', *Journal of Empirical Research on Human Research Ethics*, 7/3: 63–77. DOI: 10.1525/jer.2012.7.3.63

Kleinert, S. and Horton, R. (2014). 'How Should Medical Science Change?', *The Lancet*, 383/9913: 197–8. DOI: 10.1016/S0140-6736(13)62678-1

Komić, D., Marušić, S. L., and 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. DOI: 10.1371/journal.pone.0133662

Koterwas, A., Dwojak-Matras, A., and Kalinowska, K. (2021). 'Dialogical Teaching of Research Integrity: An Overview of Selected Methods', *FACETS*, 6: 2138–54. DOI: 10.1139/facets-2021-0045

Kumar, M. N. (2010). 'A Theoretical Comparison of the Models of Prevention of Research Misconduct', *Accountability in Research*, 17/2: 51–66. DOI: 10.1080/08989621003641132

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*, 27/4: 47. DOI: 10.1007/s11948-021-00322-9

Larivière, V., Kiermer, V., MacCallum, C. J., et al. (2016). 'A Simple Proposal for the Publication of Journal Citation Distributions' (preprint), *bioRxiv*. <http://biorxiv.org/lookup/doi/10.1101/062109> accessed 10 Dec 2020. DOI: 10.1101/062109

Lerouge, I. and Hol, T. (2020). *Towards a Research Integrity Culture at Universities: From Recommendations to Implementation* (Advice paper). <https://www.leru.org/files/Towards-a-Research-Integrity-Culture-at-Universities-full-paper.pdf> accessed 20 Feb 2020.

Lundh, A., Lexchin, J., Mintzes, B., et al. (2017). 'Industry Sponsorship and Research Outcome', *Cochrane Database of Systematic Reviews*, 2017/2. DOI: 10.1002/14651858.MR000033.pub3

Martinson, B. C., Thrush, C. R., and Crain, L. A. (2013). 'Development and Validation of the Survey of Organizational Research Climate (SORC)', *Science and Engineering Ethics*, 19/3: 813–34. DOI: 10.1007/s11948-012-9410-7

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*. DOI: 10.1002/14651858.MR000038.pub2

Mazar, N. and Ariely, D. (2015). 'Dishonesty in Scientific Research', *Journal of Clinical Investigation*, 125/11: 3993–6. DOI: 10.1172/JCI84722

McIntosh, T., Higgs, C., Mumford, M., et al. (2018). 'Continuous Evaluation in Ethics Education: A Case Study', *Science and Engineering Ethics*, 24: 727–54. DOI: 10.1007/s11948-017-9927-x

McKiernan, E. C., Schimanski, L. A., Muñoz Nieves, C., et al. (2019). 'Use of the Journal Impact Factor in Academic Review, Promotion, and Tenure Evaluations', *eLife*, 8: e47338. DOI: 10.7554/eLife.47338

Mejlgaard, N., Bouter, L. M., Gaskell, G., et al. (2020). 'Research Integrity: Nine Ways to Move from Talk to Walk', *Nature*, 586/7829: 358–60. DOI: 10.1038/d41586-020-02847-8

Michie, S., van Stralen, M. M., and West, R. (2011). 'The Behaviour Change Wheel: A New Method for Characterising and Designing Behaviour Change Interventions', *Implementation Science*, 6/1: 42. DOI: 10.1186/1748-5908-6-42

Moher, D., Bouter, L., Kleinert, S., et al. (2020). 'The Hong Kong Principles for Assessing Researchers: Fostering Research Integrity', *PLOS Biology*, 18/7: e3000737. DOI: 10.1371/journal.pbio.3000737

Moher, D., Naudet, F., Cristea, I. A., et al. (2018). 'Assessing Scientists for Hiring, Promotion, and Tenure', *PLOS Biology*, 16/3: e2004089. DOI: 10.1371/journal.pbio.2004089

Moran, H., Karlin, L., Lauchlan, E., et al. (2020). 'Understanding Research Culture: What Researchers Think about the Culture they Work in', *Wellcome Open Research*, 5: 201. DOI: 10.12688/wellcomeopenres.15832.1

Mulhearn, T. J., Watts, L. L., Torrence, B. S., et al. (2017). 'Cross-Field Comparison of Ethics Education: Golden Rules and Particulars', *Accountability in Research*, 24/4: 211–24. DOI: 10.1080/08989621.2016.1274974

Mumford, M. D., Connelly, S., Brown, R. P., et al. (2008). 'A Sensemaking Approach to Ethics Training for Scientists: Preliminary Evidence of Training Effectiveness', *Ethics & Behavior*, 18/4: 315–39. DOI: 10.1080/10508420802487815

Mumford, M. D., Steele, L., and Watts, L. L. (2015). 'Evaluating Ethics Education Programs: A Multilevel Approach', *Ethics & Behavior*, 25/1: 37–60. DOI: 10.1080/10508422.2014.917417

National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering Integrity in Research*. Washington, DC: The National Academies Press. DOI: 10.17226/21896

National Institutes of Health. (2009). *Update on the requirement for instruction in responsible conduct of research.* <https://grants.nih.gov/grants/guide/notice-files/not-od-10-019.html> accessed 19 Jun 2020.

National Research Council. (2002). *Integrity in Scientific Research: Creating an Environment That Promotes Responsible Conduct*. Washington, D.C.: National Academies Press. DOI: 10.17226/10430

National Science Foundation. (2009). 'Responsible conduct of research', *Federal Register,* 74/160: 42126-8. https://www.govinfo.gov/content/pkg/FR-2009-08-20/html/E9-19930.htm

Nosek, B. A., Spies, J. R., and Motyl, M. (2012). 'Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability', *Perspectives on Psychological Science*, 7/6: 615–31. DOI: 10.1177/1745691612459058

Nuffield Council on Bioethics. (2014). *The findings of a series of engagement activities exploring the culture of scientific research in the UK*. <https://www.nuffieldbioethics.org/assets/pdfs/The-culture-of-scientific-research-report.pdf> accessed 4 Jun 2020.

Office of Research Integrity. (n.d.). *RCR Resources*. <https://ori.hhs.gov> accessed 3 Sep 2020.

Olfson, M., Wall, M. M., and Blanco, C. (2017). 'Incentivizing Data Sharing and Collaboration in Medical Research—The S-Index', *JAMA Psychiatry*, 74/1: 5. DOI: 10.1001/jamapsychiatry.2016.2610

O'Meara, K. (2003). 'Reframing Incentives and Rewards for Community Service-Learning and Academic Outreach', *Journal of Higher Education Outreach and Engagement*, 8/2: 201–20.

O'Meara, K. (2005). 'Encouraging Multiple Forms of Scholarship in Faculty Reward Systems: Does It Make a Difference?', *Research in Higher Education*, 46/5: 479–510. DOI: 10.1007/s11162-005-3362-6

O'Meara, K. (2008). 'Motivation for Faculty Community Engagement: Learning from Exemplars', *Journal of Higher Education Outreach and Engagement*, 12/1: 7–30.

Open Science Collaboration. (2015). 'Estimating the Reproducibility of Psychological Science', *Science*, 349/6251: aac4716. DOI: 10.1126/science.aac4716

Peels, R., de Ridder, J., Haven, T., et al. (2019). 'Value Pluralism in Research Integrity', *Research Integrity and Peer Review*, 4/1: 18. DOI: 10.1186/s41073-019-0076-4

Pennock, R. T. and O'Rourke, M. (2017). 'Developing a Scientific Virtue-Based Approach to Science Ethics Training', *Science and Engineering Ethics*, 23/1: 243–62. DOI: 10.1007/s11948-016-9757-2

Phillips, T., Nestor, F., Beach, G., et al. (2018). 'America COMPETES at 5 Years: An Analysis of Research-Intensive Universities' RCR Training Plans', *Science and Engineering Ethics*, 24/1: 227–49. DOI: 10.1007/s11948-017-9883-5

Pizzolato, D., Abdi, S., and Dierickx, K. (2020). 'Collecting and Characterizing Existing and Freely Accessible Research Integrity Educational Resources', *Accountability in Research*, 27/4: 195–211. DOI: 10.1080/08989621.2020.1736571

Resnik, D. B., Neal, T., Raymond, A., et al. (2015). 'Research Misconduct Definitions Adopted by U.S. Research Institutions', *Accountability in Research*, 22/1: 14–21. DOI: 10.1080/08989621.2014.891943

Resnik, D. B., Rasmussen, L. M., and Kissling, G. E. (2015). 'An International Study of Research Misconduct Policies', *Accountability in Research*, 22/5: 249–66. DOI: 10.1080/08989621.2014.958218

Rice, D. B., Raffoul, H., Ioannidis, J. P. A., et al. (2020). 'Academic Criteria for Promotion and Tenure in Biomedical Sciences Faculties: Cross Sectional Analysis of International Sample of Universities', *BMJ*, m2081. DOI: 10.1136/bmj.m2081

Roje, R., Tomić, V., Buljan, I., et al. (2021). 'Development and Implementation of Research Integrity Guidance Documents: Explorative Interviews with Research Integrity Experts', *Accountability in Research*, 1–38. DOI: 10.1080/08989621.2021.1989676

Sandmann, L., Saltmarsh, J., and O'Meara, K. (2016). 'An Integrated Model for Advancing the Scholarship of Engagement: Creating Academic Homes for the Engaged Scholar', *Journal of Higher Education Outreach and Engagement*, 20/1: 157–74.

Schekman, R. and Patterson, M. (2013). 'Reforming Research Assessment', *eLife*, 2: e00855. DOI: 10.7554/eLife.00855

Schneider, B., Ehrhart, M. G., and Macey, W. H. (2013). 'Organizational Climate and Culture', *Annual Review of Psychology*, 64/1: 361–88. DOI: 10.1146/annurev-psych-113011-143809

Seglen, P. O. (1997). 'Why the Impact Factor of Journals Should not be Used for Evaluating Research', *BMJ*, 314/7079: 497–497. DOI: 10.1136/bmj.314.7079.497

Snider, A., Hight, K., Brunson, A., et al. (2021). 'Analysis of Research and Scholarship Criteria Within Promotion and Tenure Documents of US Pharmacy Schools', *American Journal of Pharmaceutical Education*, 85/3: 7983. DOI: 10.5688/ajpe7983

Sørensen, M. P., Ravn, T., Marušić, A., et al. (2021). 'Strengthening Research Integrity: Which Topic Areas Should Organisations Focus on?', *Humanities and Social Sciences Communications*, 8/1: 198. DOI: 10.1057/s41599-021-00874-y

Steele, L. M., Mulhearn, T. J., Medeiros, K. E., et al. (2016). 'How Do We Know What Works? A Review and Critique of Current Practices in Ethics Training Evaluation', *Accountability in Research*, 23/6: 319–50. DOI: 10.1080/08989621.2016.1186547

Tang, B. L. and Lee, J. S. C. (2020). 'A Reflective Account of a Research Ethics Course for an Interdisciplinary Cohort of Graduate Students', *Science and Engineering Ethics*, 26/2: 1089–105. DOI: 10.1007/s11948-020-00200-w

Tijdink, J. K., Bouter, L. M., Veldkamp, C. L. S., et al. (2016). 'Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study', *PLOS ONE*, 11/9: e0163251. DOI: 10.1371/journal.pone.0163251

Titus, S. L. and Ballou, J. M. (2014). 'Ensuring PhD Development of Responsible Conduct of Research Behaviors: Who's Responsible?', *Science and Engineering Ethics*, 20/1: 221–35. DOI: 10.1007/s11948-013-9437-4

Todd, E. M., Watts, L. L., Mulhearn, T. J., et al. 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: 1719–54. DOI: 10.1007/s11948-017-9869-3

Tomić, V., Buljan, I., and Marušić, A. (2021). 'Perspectives of Key Stakeholders on Essential Virtues for Good Scientific Practice in Research Areas', *Accountability in Research*, 1–32. DOI: 10.1080/08989621.2021.1900739

Turner, M. R., Watts, L. L., Steele, L. M., et al. (2018). 'How Did You Like This Course? The Advantages and Limitations of Reaction Criteria in Ethics Education', *Ethics & Behavior*, 28/6: 483–96. DOI: 10.1080/10508422.2017.1308193

Valkenburg, G., Dix, G., Tijdink, J., et al. (2021). 'Expanding Research Integrity: A Cultural-Practice Perspective', *Science and Engineering Ethics*, 27/1: 10. DOI: 10.1007/s11948-021-00291-z

Van Calster, B., Wynants, L., Riley, R. D., et al. (2021). 'Methodology over Metrics: Current Scientific Standards are a Disservice to Patients and Society', *Journal of Clinical Epidemiology*, 138: 219–26. DOI: 10.1016/j.jclinepi.2021.05.018

Van Noorden, R. (2018). 'Some Hard Numbers on Science's Leadership Problems', *Nature*, 557/7705: 294–6. DOI: 10.1038/d41586-018-05143-8

Vie, K. J. (2020). 'How Should Researchers Cope with the Ethical Demands of Discovering Research Misconduct? Going Beyond Reporting and Whistleblowing', *Life Sciences, Society and Policy*, 16/1: 6. DOI: 10.1186/s40504-020-00102-6

Vitae. (2020). *Research Integrity: A Landscape Study*. <https://www.vitae.ac.uk/vitaepublications/

reports/research-integrity-a-landscape-study> accessed 12 Jul 2020.

Wager, E. and Kleinert, S. (2021). 'Cooperation & Liaison between Universities & Editors (CLUE): Recommendations on Best Practice', *Research Integrity and Peer Review*, 6/1: 6. DOI: 10.1186/s41073-021-00109-3

Waltman, L. (2016). 'A Review of the Literature on Citation Impact Indicators', *Journal of Informetrics*, 10/2: 365–91. DOI: 10.1016/j.joi.2016.02.007

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*, 27/5: 351–84. DOI: 10.1080/10508422.2016.1182025

Watts, L. L., Mulhearn, T. J., Medeiros, K. E., et al. (2017). 'Modeling the Instructional Effectiveness of Responsible Conduct of Research Education: A Meta-Analytic Path-Analysis', *Ethics & Behavior*, 27/8: 632–50. DOI: 10.1080/10508422.2016.1247354

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*, 9/5: 72–88. DOI: 10.1177/1556264614552798

Wilsdon, J. (2015). *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. London: SAGE Publications Ltd. DOI: 10.4135/9781473978782

Winchester, C. (2018). 'Give Every Paper a Read for Reproducibility', *Nature*, 557/7705: 281. DOI: 10.1038/d41586-018-05140-x

Xie, Y., Wang, K., and Kong, Y. (2021). 'Prevalence of Research Misconduct and Questionable Research Practices: A Systematic Review and Meta-Analysis', *Science and Engineering Ethics*, 27/4: 41. DOI: 10.1007/s11948-021-00314-9

Yeo-Teh, N. S. L. and Tang, B. L. (2021). 'Research Ethics Courses as a Vaccination against a Toxic Research Environment or Culture', *Research Ethics*, 17/1: 55–65. DOI: 10.1177/1747016120926686

Zollitsch, L., Wilder, N., and Priess-Buchheit, J. (2021). 'The Development of a Four-Tier Test to Evaluate Research Integrity Training', *Open Research Europe*, 1: 90. DOI: 10.12688/openreseurope.13339.1

Zwart, H. and ter Meulen, R. (2019). 'Addressing Research Integrity Challenges: From Penalising Individual Perpetrators to Fostering Research Ecosystem Quality Care', *Life Sciences, Society and Policy*, 15/1: 5, s40504-019-0093–6. DOI: 10.1186/s40504-019-0093-6

BOXES

BOX 1 Processes, policies, and procedures

To promote a culture of research integrity, institutions should:

1) Develop – or adopt existing – clear and specific codes of conduct and guidelines that operationally define the desired situation in terms of expected behaviors

2) Communicate clearly about all codes of conduct and integrity guidelines, and check researchers' awareness of them

3) Provide support in the form of dedicated human resources and infrastructure

4) Promote a culture of collective openness

5) Use a broad range of assessment criteria that recognize various kinds of research outputs as well as responsible research practices, openness, societal impact, and mentoring and supervision

6) Avoid the use of journal-level metrics in the assessment of researchers

BOX 2 Dealing with breaches of research integrity

To promote a culture of research integrity, institutions should:

7) Develop – or adopt – specific procedures for dealing with suspected breaches of integrity, including measures to protect whistle-blowers and accused researchers

8) Communicate clearly about these procedures and check researchers' awareness of them

9) Publicly share anonymized reports of research misconduct investigations

BOX 3 Education and training

To promote a culture of research integrity, institutions should:

10) Develop and share – or adopt existing effective – trainings in the responsible conduct of research, preferably case-based

11) Develop and share – or adopt existing – trainings in mentoring and supervision

12) Combine the benefits of multiple interactive face-to-face sessions with that of online integrity training into a mixed approach

13) Train researchers throughout their careers

14) Remain up to date with the emerging empirical evidence to offer state-of-the-art training

BOX 4 Monitoring and evaluation

To promote a culture of research integrity, institutions should:

15) Permanently monitor and evaluate the effectiveness of research integrity policies (e.g., training sessions, authorship guidelines, training in mentoring and supervision), and adapt them when necessary

16) Collect empirical data about the effectiveness of research integrity policies, and makes these publicly available
17) Conduct experiments on research integrity policies, and make the results of these experiments publicly available
18) Remain up to date with the emerging empirical evidence to issue a state-of-the-art integrity policy

APPENDIX

There are several online resources with cases, some allow selecting cases based on target group. A few examples are:

- The Dilemma Game https://www.eur.nl/en/about-eur/policy-and-regulations/integrity/research-integrity/dilemma-game
- The COPE database https://publicationethics.org/guidance/Case
- The Embassy of Good Science https://embassy.science/wiki-wiki/index.php/Special:BrowseData/Resource?_search_Resource_Type%5B0%5D=Cases
- The video case studies of ORI https://ori.hhs.gov/videos/case-study and 'The Lab', their interactive movie on research misconduct https://ori.hhs.gov/TheLab/
- Association for Practical and Professional Ethics (APPE) – '*Research Ethics: Cases and Commentaries'*. https://ethics.unl.edu/ethics_resources/local/appe-case-studies.shtml#volume7