Research practices and assessment of research misconduct

Hartgerink, C.H.J.; Wicherts, J.M.

Published in:
ScienceOpen Research

DOI:
10.14293/S2199-1006.1.SOR-SOCSCI.ARYSBI.v1

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Research practices and assessment of research misconduct

Chris H.J. Hartgerink* and Jelte M. Wicherts

Department of Methodology and Statistics, Tilburg University, Tilburg, The Netherlands
*Corresponding author's e-mail address: c.h.j.hartgerink@tilburguniversity.edu

Published online: 2 August 2016 (version 1)
Cite as: Hartgerink and Wicherts. ScienceOpen Research 2016 (DOI: 10.14293/S2199-1006.1.SOR-SOCSCI.ARYSBI.v1)

Reviewing status: The review of this article was author-mediated. The ScienceOpen editorial office has verified and confirmed the identities of the reviewers. The endorsement statement and reviewers names are provided below. The article is also available for continuous post-publication peer review on its dedicated ScienceOpen homepage.

Peer Review by Endorsement Statement:
"*I have read this article, given feedback to the authors and now feel that it is of appropriate quality to be included in the scientific literature and be part of the open scientific discourse.*"

Prof. Dr. Bjoern Brembs
Professor of Neurogenetics, Institute of Zoology
Universität Regensburg, 2012-present

Dr. Brian Nosek
Professor, Department of Psychology, University of Virginia
Executive Director, Center for Open Science

Primary discipline: General social science
Secondary discipline: Applied ethics, General behavioral science, Ethics, Applied probability & Statistics
Keywords transparency, research misconduct, questionable research practices, responsible research, data fabrication

ABSTRACT

This article discusses the responsible conduct of research, questionable research practices, and research misconduct. Responsible conduct of research is often defined in terms of a set of abstract, normative principles, professional standards, and ethics in doing research. In order to accommodate the normative principles of scientific research, the professional standards, and a researcher's moral principles, transparent research practices can serve as a framework for responsible conduct of research. We suggest a "prune-and-add" project structure to enhance transparency and, by extension, responsible conduct of research. Questionable research practices are defined as practices that are detrimental to the research process. The prevalence of questionable research practices remains largely unknown, and reproducibility of findings has been shown to be problematic. Questionable practices are discouraged by transparent practices because practices that arise from them will become more apparent to scientific peers. Most effective might be preregistrations of research design, hypotheses, and analyses, which reduce particularism of results by providing an a priori research scheme. Research misconduct has been defined as fabrication, falsification, and plagiarism (FFP), which is clearly the worst type of research practice. Despite it being clearly wrong, it can be approached from a scientific and legal perspective. The legal perspective sees research misconduct as a form of white-collar crime. The scientific perspective seeks to answer the following question: "Were results invalidated because of the misconduct?" We review how misconduct is typically detected, how its detection can be improved, and how prevalent it might be. Institutions could facilitate detection of data fabrication and falsification by implementing data auditing. Nonetheless, the effect of misconduct is pervasive: many retracted articles are still cited after the retraction has been issued.

Main points

1. Researchers systematically evaluate their own conduct as more responsible than colleagues, but not as responsible as they would like.
2. Transparent practices, facilitated by the Open Science Framework, help embody scientific norms that promote responsible conduct.
3. Questionable research practices harm the research process and work counter to the generally accepted scientific norms, but are hard to detect.
4. Research misconduct requires active scrutiny of the research community because editors and peer-reviewers do not pay adequate attention to detecting this. Tips are given on how to improve your detection of potential problems.

INTRODUCTION

Research practices directly affect the epistemological pursuit of science: Responsible conduct of research affirms it; research misconduct undermines it. Typically, a responsible scientist is conceptualized as objective, meticulous, skeptical, rational, and not subject to external incentives such as prestige or social pressure. Research misconduct, on the other hand, is formally

defined (e.g., in regulatory documents) as three types of condemned, intentional behaviors: fabrication, falsification, and plagiarism (FFP; Office of Science and Technology Policy, 2000). Research practices that are neither conceptualized as responsible nor defined as research misconduct could be considered questionable research practices, which are practices that are detrimental to the research process (QRPs; Panel on Scientific Responsibility and the Conduct of Research, 1992; Steneck, 2006). For example, the misapplication of statistical methods can increase the number of false results and is therefore not responsible. At the same time, such misapplication can also not be deemed research misconduct because it falls outside the defined scope of FFP. Such undefined and potentially questionable research practices have been widely discussed in the field of psychology in recent years (e.g., John, Loewenstein, & Prelec, 2012; Nosek & Bar-Anan, 2012; Nosek, Spies, & Motyl, 2012; Open Science Collaboration, 2015; Simmons, Nelson, & Simonsohn, 2011).

This article discusses the responsible conduct of research, questionable research practices, and research misconduct. For each of these three, we extend on what it means, what researchers currently do, and how it can be facilitated (i.e., responsible conduct) or prevented (i.e., questionable practices and research misconduct). These research practices encompass the entire research practice spectrum proposed by Steneck (2006), where responsible conduct of research is the ideal behavior at one end, FFP the worst behavior on the other end, with (potentially) questionable practices in between.

RESPONSIBLE CONDUCT OF RESEARCH

What is it?

Responsible conduct of research is often defined in terms of a set of abstract, normative principles. One such set of norms of good science (Anderson, Ronning, Devries, & Martinson, 2010; Merton, 1942) is accompanied by a set of counternorms (Anderson et al., 2010; Mitroff, 1974) that promulgate irresponsible research. These six norms and counternorms can serve as a valuable framework to reflect on the behavior of a researcher and are included in Table 1.

Besides abiding by these norms, responsible conduct of research consists of both research integrity and research ethics (Shamoo & Resnik, 2009). Research integrity is the adherence to professional standards and rules that are well defined and uniform, such as the standards outlined by the American Psychological Association (2010). Research ethics, on the other hand, is "the critical study of the moral problems associated with or that arise in the course of pursuing research" (Steneck, 2006, p. 56), which is abstract and pluralistic. As such, research ethics is more fluid than research integrity and is supposed to fill in the gaps left by research integrity (Koppelman-White, 2006). For example, not fabricating data is the professional standard in research, but research ethics informs us on why it is wrong to fabricate data. This highlights that ethics and integrity are not the same, but rather two related constructs. Discussion or education should therefore not only reiterate the professional standards but also include training on developing ethical and moral principles that can guide researchers in their decision-making.

What do researchers do?

Even though most researchers subscribe to the aforementioned normative principles, fewer researchers actually adhere to them in practice and many researchers perceive their scientific peers to adhere to them even less. A survey of 3,247 researchers by Anderson, Martinson, and De Vries (2007) indicated that researchers subscribed to the norms more than they actually behaved in accordance to these norms. For instance, a researcher may be committed to sharing his or her data (the norm of communality), but might shy away from actually sharing data at an early stage out of a fear that of being scooped by other researchers. This result aligns with surveys showing that many researchers express a willingness to share data, but often fail to do so when asked (Krawczyk & Reuben, 2012; Savage & Vickers, 2009). Moreover, although researchers admit they do not adhere to the norms as much as they subscribe to them, they still regard themselves as adhering to the norms more so than their peers. For counternorms, this pattern reversed. These results indicate that researchers systematically evaluate their own conduct as more responsible than other researchers' conduct.

This gap between subscription and actual adherence to the normative principles is called normative dissonance and could potentially be due to substandard academic education or lack of open discussion on ethical issues. Anderson, Horn, et al. (2007) suggested that different types of mentoring affect the normative behavior by a researcher. Most importantly, ethics mentoring (e.g., discussing whether a mistake that does not affect conclusions should result in a corrigendum) might promote adherence to the norms, whereas survival mentoring (e.g., advising not to

Table 1. Six norms of responsible conduct of research and their respective counternorms (Anderson et al., 2010; Merton, 1942; Mitroff, 1974).

Norm	Description norm	Counternorm
Universalism	Evaluate results based on pre-established and non-personal criteria	Particularism
Communality	Freely and widely share findings	Secrecy
Disinterestedness	Results not corrupted by personal gains	Self-interestedness
Skepticism	Scrutinize all findings, including own	Dogmatism
Governance	Decision-making in science is done by researchers	Administration
Quality	Evaluate researchers based on the quality of their work	Quantity

submit a noncrucial corrigendum because it could be bad for your scientific reputation) might promote adherence to the counternorms. Ethics mentoring focuses on discussing ethical issues (Anderson, Horn, et al., 2007) that might facilitate higher adherence to norms due to increased self-reflection, whereas survival mentoring focuses on how to thrive in academia and focuses on building relationships and specific skills to increase the odds of being successful.

Improving responsible conduct

Increasing exposure to ethics education throughout the research career might improve responsible research conduct. Research indicated that weekly 15-minute ethics discussions facilitated confidence in recognizing ethical problems in a way that participants deemed both effective and enjoyable (Peiffer, Hugenschmidt, & Laurienti, 2011). Such forms of active education are fruitful because they teach researchers practical skills that can change their research conduct and improve prospective decision-making, where a researcher rapidly assesses the potential outcomes and ethical implications of the decision at hand, instead of in hindsight (Whitebeck, 2001). It is not to be expected that passive education on guidelines should be efficacious in producing behavioral change (Kornfeld, 2012), considering that participants rarely learn about useful skills or experience a change in attitudes as a consequence of such passive education (Plemmons, Brody, & Kalichman, 2006).

Moreover, in order to accommodate the normative principles of scientific research, the professional standards, and a researcher's moral principles, transparent research practices can serve as a framework for responsible conduct of research. Transparency in research embodies the normative principles of scientific research: universalism is promoted by improved documentation; communalism is promoted by publicly sharing research; disinterestedness is promoted by increasing accountability and exposure of potential conflicts of interest; skepticism is promoted by allowing for verification of results; governance is promoted by improved project management by researchers; and higher quality is promoted by the other norms. Professional standards also require transparency. For instance, the APA and publication contracts require researchers to share their data with other researchers (American Psychological Association, 2010). Even though authors often make their data available upon request, such requests frequently fail (Krawczyk & Reuben, 2012; Wicherts, Borsboom, Kats, & Molenaar, 2006), which results in a failure to adhere to professional standards. Openness regarding the choices made (e.g., on how to analyze the data) during the research process will promote active discussion of prospective ethics, increasing self-reflective capacities of both the individual researcher and the collective evaluation of the research (e.g., peer-reviewers).

In the remainder of this section, we outline a type of project management, founded on transparency, which seems apt to be the new standard within psychology (Nosek & Bar-Anan, 2012; Nosek et al., 2012). Transparency guidelines for journals have also been proposed (Nosek et al., 2015), and the outlined project management adheres to these guidelines from an author's perspective. The provided format focuses on empirical research and is certainly not the only way to apply transparency to adhere to responsible conduct of research principles.

Transparent project management

Research files can be easily managed by creating an online project at the Open Science Framework (OSF; osf.io). The OSF is free to use and provides extensive project management facilities to encourage transparent research. Project management via this tool has been tried and tested in, for example, the Many Labs project (osf.io/wx7ck; Klein et al., 2014) and the Reproducibility project (osf.io/ezcuj; Open Science Collaboration, 2015). Research files can be manually uploaded by the researcher or automatically synchronized (e.g., via Dropbox or Github). Using the OSF is easy and explained in-depth at osf.io/getting-started. The OSF provides the tools to manage a research project, but how to apply these tools still remains a question. Such online management of materials, information, and data is preferred above a more informal system lacking in transparency that often strongly rests on particular contributor's implicit knowledge.

As a way to organize a version-controlled project, we suggest a "prune-and-add" template, where the major elements of most research projects are included but which can be specified and extended for specific projects. This template includes folders as specified in Table 2, which covers many of the research stages. The template can be readily duplicated and adjusted on the OSF for practical use in similar projects (like replication studies; osf.io/4sdn3).

This suggested project structure also includes a folder to include preregistration files of hypotheses, analyses, and research design. The preregistration of these ensures that the researcher does not hypothesize after the results are known (Kerr, 1998), but also

Table 2. Project management folder structure, which can be pruned and added to in order to meet specific research needs. This folder structure can be duplicated as an OSF project at osf.io/4sdn3

Folder	Summary of contents
Analyses	Analyses scripts (e.g., as reported in the paper, exploratory files)
Archive	Outdated files or files not of direct value (e.g., unused code)
Bibliography	Reference library or related articles (e.g., Endnote library, PDF files)
Data	All data files used (e.g., raw data, processed data)
Figures	Figures included in the manuscript and code for figures
Functions	Custom functions used (e.g., SPSS macro, R scripts)
Materials	Research materials specified per study (e.g., survey questions, stimuli)
Preregister	Preregistered hypotheses, analysis plans, research designs
Submission	Manuscript, submissions per journal, and review rounds
Supplement	Files that supplement the research project (e.g., notes, codebooks)

ensures readers that the results presented as confirmatory were actually confirmatory (Chambers, 2015; Wagenmakers, Wetzels, Borsboom, Van der Maas, & Kievit, 2012). The preregistration of analyses also ensures that the statistical analysis chosen to test the hypothesis was not dependent on the result. Such preregistrations document the chronology of the research process and also ensure that researchers actively reflect on the decisions they make prior to running a study, such that the quality of the research might be improved.

Also available in this project template is a file to specify contributions to a research project. This is important for determining authorship, responsibility, and credit of the research project. With more collaborations occurring throughout science and increasing specialization, researchers cannot be expected to carry responsibility for the entirety of large multidisciplinary papers, but authorship does currently imply this. Consequently, authorship has become a too imprecise measure for specifying contributions to a research project and requires a more precise approach.

Besides structuring the project and documenting the contributions, responsible conduct encourages independent verification of the results to reduce particularism. A co-pilot model has been introduced previously (Veldkamp, Nuijten, Dominguez-Alvarez, Van Assen, & Wicherts, 2014; Wicherts, 2011), where at least two researchers independently run all analyses based on the raw data. Such verification of research results enables streamline reproduction of the results by outsiders (e.g., Are all files readily available? Are the files properly documented? Do the analyses work on someone else's computer?), helps find out potential errors (e.g., rounding errors; Bakker & Wicherts, 2011; Nuijten, Hartgerink, Van Assen, Epskamp, & Wicherts, 2015), and increases confidence in the results. We therefore encourage researchers to incorporate such a co-pilot model into all empirical research projects.

QUESTIONABLE RESEARCH PRACTICES

What is it?

Questionable research practices are defined as practices that are detrimental to the research process (Panel on Scientific Responsibility and the Conduct of Research, 1992). Examples include inadequate research documentation, failing to retain research data for a sufficient amount of time, and actively refusing access to published research materials. However, questionable research practices should not be confounded with questionable academic practices, such as academic power play, sexism, and scooping.

Attention for questionable practices in psychology has (re-) arisen in recent years, in light of the so-called replication crisis (e.g., Makel, Plucker, & Hegarty, 2012). Pinpointing which factors initiated doubts about the reproducibility of findings is difficult, but most notable seems an increased awareness of widely accepted practices as statistically and methodologically questionable.

Besides affecting the reproducibility of psychological science, questionable research practices align with the aforementioned counternorms in science. For instance, confirming prior beliefs by selectively reporting results is a form of dogmatism; skepticism and communalism are violated by not providing peers with research materials or details of the analysis; universalism is hindered by lack of research documentation; governance is deteriorated when the public loses its trust in the research system because of signs of the effects of questionable research practices (e.g., repeated failures to replicate) and politicians initiate new forms of oversight.

Suppose a researcher fails to find the (a priori) hypothesized effect, subsequently decides to inspect the effect for each gender, and finds an effect only for females. Such an ad hoc exploration of the data is perfectly fine if it were presented as an exploration (Wigboldus & Dotsch, 2016). However, if the subsequent publication only mentions the effect for females and presents it as confirmatory, instead of exploratory, this is questionable. The *p*-values should have been corrected for multiple testing (three hypotheses rather than one were tested), and the result is clearly not as convincing as one that would have been hypothesized a priori.

These biases occur in part because researchers, editors, and peer-reviewers are biased to believe that statistical significance has a bearing on the probability of a hypothesis being true. Such misinterpretation of the *p*-value is not uncommon (Cohen, 1994). The perception that statistical significance bears on the probability of a hypothesis reflects an essentialist view of *p*-values rather than a stochastic one; the belief that if an effect exists, the data will mirror this with a small *p*-value (Sijtsma, Veldkamp, & Wicherts, 2015). Such problematic beliefs enhance publication bias because researchers are less likely to believe in their results and are less likely submit their work for publication (Franco, Malhotra, & Simonovits, 2014). This enforces the counternorm of secrecy by keeping non-significant results in the file-drawer (Rosenthal, 1979), which in turn greatly biases the picture emerging from the literature.

What do researchers do?

Most questionable research practices are hard to retrospectively detect, but one questionable research practice, the misreporting of statistical significance, can be readily estimated and could provide some indication of how widespread questionable practices might be. Errors that result in the incorrect conclusion that a result is significant are often called gross errors, which indicates that the decision error had substantive effects. Large-scale research in psychology has indicated that 12.5–20% of sampled articles include at least one such gross error, with approximately 1% of all reported test results being affected by such gross errors (Bakker & Wicherts, 2011; Nuijten et al., 2015; Veldkamp et al., 2014).

Nonetheless, the prevalence of questionable research practices remains largely unknown, and reproducibility of findings has been shown to be problematic. In one large-scale project, only 36% of findings published in three main psychology journals in a given year could be replicated (Open Science Collaboration, 2015).

Effect sizes were smaller in the replication than in the original study in 80% of the studies, and it is quite possible that this low replication rate and decrease in effect sizes are mostly due to publication bias and the use of questionable research practices in the original studies.

How can it be prevented?

Counternorms such as self-interestedness, dogmatism, and particularism are discouraged by transparent practices because practices that arise from them will become more apparent to scientific peers.

Therefore, transparency guidelines have been proposed and signed by editors of over 500 journals (Nosek et al., 2015). To different degrees, signatories of these guidelines actively encourage, enforce, and reward data sharing, material sharing, preregistration of hypotheses or analyses, and independent verification of results. The effects of these guidelines are not yet known, considering their recent introduction. Nonetheless, they provide a strong indication that the awareness of problems is trickling down into systemic changes that prevent questionable practices.

Most effective might be preregistrations of research design, hypotheses, and analyses, which reduce particularism of results by providing an a priori research scheme. It also outs behaviors such as the aforementioned optional stopping, where extra participants are sampled until statistical significance is reached (Armitage, McPherson, & Rowe, 1969) or the dropping of conditions or outcome variables (Franco, Malhotra, & Simonovits, 2016). Knowing that researchers outlined their research process and seeing it adhered to help ensure readers that results are confirmatory—rather than exploratory of nature, when results are presented as confirmatory (Wagenmakers et al., 2012), ensuring researchers that questionable practices did not culminate in those results.

Moreover, use of transparent practices even allows for unpublished research to become discoverable, effectively eliminating publication bias. Eliminating publication bias would make the research system an estimated 30 times more efficient (Van Assen, Van Aert, Nuijten, & Wicherts, 2014). Considering that unpublished research is not indexed in the familiar peer-reviewed databases, infrastructures to search through repositories similar to the OSF are needed. One such infrastructure is being built by the Center for Open Science (SHARE; osf.io/share), which searches through repositories similar to the OSF (e.g., figshare, Dryad, arXiv).

RESEARCH MISCONDUCT

What is it?

As mentioned at the beginning of the article, research misconduct has been defined as fabrication, falsification, and plagiarism (FFP). However, it does not include "honest error or differences of opinion" (Office of Science and Technology Policy, 2000; Resnik & Stewart, 2012). Fabrication is the making up of datasets entirely. Falsification is the adjustment of a set of data points to ensure the wanted results. Plagiarism is the direct reproduction of other's creative work without properly attributing it. These behaviors are condemned by many institutions and organizations, including the American Psychological Association (2010).

Research misconduct is clearly the worst type of research practice, but despite it being clearly wrong, it can be approached from a scientific and legal perspective (Wicherts & Van Assen, 2012). The scientific perspective condemns research misconduct because it undermines the pursuit for knowledge. Fabricated or falsified data are scientifically useless because they do not add any knowledge that can be trusted. Use of fabricated or falsified data is detrimental to the research process and to knowledge building. It leads other researchers or practitioners astray, potentially leading to waste of research resources when pursuing false insights or unwarranted use of such false insights in professional or educational practice.

The legal perspective sees research misconduct as a form of white-collar crime, although in practice it is typically not subject to criminal law but rather to administrative or labor law. The legal perspective requires intention to commit research misconduct, whereas the scientific perspective requires data to be collected as described in a research report, regardless of intent. In other words, the legal perspective seeks to answer the following question: "Was misconduct committed with intent and by whom?"

The scientific perspective seeks to answer the following question: "Were results invalidated because of the misconduct?" For instance, a paper reporting data that could not have been collected with the materials used in the study (e.g., the reported means lie outside the possible values on the psychometric scale) is invalid scientifically. The impossible results could be due to research misconduct but also due to honest error.

Hence, a legal verdict of research misconduct requires proof that a certain researcher falsified or fabricated the data. The scientific assessment of the problems is often more straightforward than the legal assessment of research misconduct. The former can be done by peer-reviewers, whereas the latter involves regulations and a well-defined procedure allowing the accused to respond to the accusations.

Throughout this part of the article, we focus on data fabrication and falsification, which we will illustrate with examples from the Diederik Stapel case—a case we are deeply familiar with. His fraudulent activities resulted in 58 retractions (as of May, 2016), making this the largest known research misconduct case in the social sciences.

What do researchers do?

Given that research misconduct represents such a clear violation of the normative structure of science, it is difficult to study how many researchers commit research misconduct and why they do it. Estimates based on self-report surveys suggest that around 2% of researchers admit to having fabricated or

falsified data during their career (Fanelli, 2009). Although the number of retractions due to misconduct has risen in the last decades, both across the sciences in general (Fang, Steen, & Casadevall, 2012) and in psychology in particular (Margraf, 2015), this number still represents a fairly low number in comparison to the total number of articles in the literature (i.e., 31 retractions to 136,191 publications in PsycINFO for 2015; Wicherts, Hartgerink, & Grasman, 2016). Similarly, the number of researchers found guilty of research misconduct is relatively low, suggesting that many cases of misconduct go undetected; the actual rate of research misconduct is unknown. Little research has addressed why researchers fabricate or falsify data, but it is commonly accepted that they do so out of self-interest in order to obtain publications and further their career. What we know from some exposed cases, however, is that fabricated or falsified data are often quite extraordinary and so could sometimes be exposed as not being genuine.

Humans, including researchers, are quite bad in recognizing and fabricating probabilistic processes (Mosimann, Dahlberg, Davidian, & Krueger, 2002; Mosimann, Wiseman, & Edelman, 1995). For instance, humans frequently think that, after five coin flips that result in heads, the probability of the next coin flip is more likely to be tails than heads; the gambler's fallacy (Tversky & Kahneman, 1974). Inferential testing is based on sampling; by extension variables should be of probabilistic origin and have certain stochastic properties. Because humans have problems adhering to these probabilistic principles, fabricated data are likely to lead to data that does not properly adhere to the probabilistic origins at some level of the data (Haldane, 1948).

Exemplary of this lack of fabricating probabilistic processes is a table in a now retracted paper from the Stapel case (Ret, 2012; Ruys & Stapel, 2008). In the original Table 1, reproduced here as Table 3, 32 means and standard deviations are presented. *Fifteen* of these cells are duplicates of another cell (e.g., "0.87 (0.74)" occurs three times). Finding exact duplicates is extremely rare for even one case if the variables are a result of probabilistic processes as in sampling theory.

Why reviewers and editors did not detect this remains a mystery, but it seems that they simply do not pay attention to potential indicators of misconduct in the publication process (Bornmann, Nast, & Daniel, 2008). Similar issues with blatantly problematic results in papers that were later found to be due to misconduct have been noted in the medical sciences (Stewart & Feder, 1987). Science has been regarded as a self-correcting system based on trust. This aligns with the idea that misconduct occurs because of "bad apples" (i.e., individual factors) and not because of a "bad barrel" (i.e., systemic factors), increasing trust in the scientific enterprise. However, the self-correcting system has been called a myth (Stroebe, Postmes, & Spears, 2012) and an assumption that instigates complacency (Hettinger, 2010); if reviewers and editors have no criteria that pertain to fabrication and falsification (Bornmann et al., 2008), this implies that the current publication process is not always functioning properly as a self-correcting mechanism. Moreover, trust in research as a self-correcting system can be accompanied with complacency by colleagues in the research process.

The most frequent way data fabrication is detected is by those researchers who are scrutinous, which ultimately results in whistleblowing. For example, Stapel's misdeeds were detected by young researchers who were brave enough to blow the whistle. Although many regulations include clauses that help protect the whistleblowers, whistleblowing is known to represent a risk (Lubalin, Ardini, & Matheson, 1995), not only because of potential backlash but also because the perpetrator is often closely associated with the whistleblower, potentially leading to negative career outcomes such as retracted articles on which one is co-author. This could explain why whistleblowers remain anonymous in only an estimated 8% of the cases (Price, 1998). Negative actions as a result of loss of anonymity include not only potential loss of a position but also social and mental health problems (Allen & Dowell, 2013; Lubalin & Matheson, 1999). It seems plausible to assume that therefore not all suspicions are reported.

How often data fabrication and falsification occur is an important question that can be answered in different ways; it can be approached as incidence or as prevalence. Incidence refers

Table 3. Reproduction of Table 1 from the retracted Ruys and Stapel (2008) paper. The table shows 32 cells with "M (SD)," of which 15 are direct duplicates of one of the other cells. The original version with highlighted duplicates can be found at osf.io/89mcn.

	Prime emotion			
Exposure duration and fragment type	**Disgust**	**Fear**	**Anger**	**Neutral**
Quick (120 ms)				
Disgust fragments	2.33 (0.62)	1.20 (0.94)	1.20 (0.68)	1.53 (0.74)
Fear fragments	0.80 (0.78)	1.87 (0.92)	1.13 (0.92)	1.00 (0.93)
Anger fragments	0.93 (0.70)	0.93 (0.70)	1.80 (0.86)	0.80 (0.78)
Negative fragments	2.27 (0.46)	2.33 (0.82)	2.20 (0.41)	1.33 (0.98)
Super-quick (40 ms)				
Disgust fragments	1.27 (0.96)	1.07 (0.80)	1.27 (0.96)	1.33 (0.72)
Fear fragments	1.07 (0.59)	0.87 (0.74)	1.07 (0.59)	1.00 (0.66)
Anger fragments	0.87 (0.74)	1.07 (0.80)	0.87 (0.74)	0.87 (0.83)
Negative fragments	1.80 (0.56)	2.07 (0.80)	2.27 (0.46)	0.93 (0.88)

to new cases in a certain timeframe, whereas prevalence refers to all cases in the population at a certain time point. Misconduct cases are often widely publicized, which might create the image that more cases occur, but the number of cases seems relatively stable (Rhoades, 2004). Prevalence of research misconduct is of great interest and, as aforementioned, a meta-analysis indicated that around 2% of surveyed researchers admit to fabricating or falsifying research at least once (Fanelli, 2009).

The prevalence that is of greatest interest is that of how many research papers contain data that have been fabricated or falsified. Systematic data on this are unavailable because papers are not evaluated to this end in an active manner (Bornmann et al., 2008). Only one case study exists: *The Journal of Cell Biology* evaluates all research papers for cell image manipulation (e.g., Western blots; see also Bik, Casadevall, & Fang, 2016; Rossner & Yamada, 2004), a form of data fabrication/falsification. They have found that approximately 1% of all research papers that passed peer-review (out of total of over 3000 submissions) were not published because of the detection of image manipulation (*The Journal of Cell Biology,* 2015).

How can it be prevented?

Notwithstanding discussion about reconciliation of researchers who have been found guilty of research misconduct (Cressey, 2013), these researchers typically leave science after having been exposed. Hence, improving the chances of detecting misconduct may help not only in the correction of the scientific record but also in the prevention of research misconduct. In this section, we discuss how the detection of fabrication and falsification might be improved and what to do when misconduct is detected.

When research is suspect of data fabrication or falsification, whistleblowers can report these suspicions to institutions, professional associations, and journals. For example, institutions can launch investigations via their integrity offices. Typically, a complaint is submitted to the research integrity officer, who subsequently decides whether there are sufficient grounds for further investigation. In the United States, integrity officers have the possibility to sequester, that is, to retrieve, all data of the person in question. If there is sufficient evidence, a formal misconduct investigation or even a federal misconduct investigation by the Office of Research Integrity might be started. Professional associations can also launch some sort of investigation if the complaint is made to the association and the respondent is a member of that association. Journals are also confronted with complaints about specific research papers, and those affiliated with the Committee on Publication Ethics have a protocol for dealing with these kinds of allegations (see publicationethics.org/resources for details). The best way to improve detection of data fabrication directly is to further investigate suspicions and report them to your research integrity office, albeit the potential negative consequences should be kept in mind when reporting the suspicions, such that it is best to report anonymously and via analog mail (digital files contain metadata with identifying information).

More indirectly, statistical tools can be applied to evaluate the veracity of research papers and raw data (Carlisle, Dexter, Pandit, Shafer, & Yentis, 2015; Peeters, Klaassen, & van de Wiel, 2015), which helps detect potential lapses of conduct. Statistical tools have been successfully applied in data fabrication cases, for instance, the Stapel case (Levelt Committee, Drenth Committee, & Noort, Committee, 2012), the Fuji case (Carlisle, 2012), and in the cases of Smeesters and Sanna (Simonsohn, 2013). Interested readers are referred to Buyse et al. (1999) for a review of statistical methods to detect potential data fabrication.

Besides using statistics to monitor for potential problems, authors and principal investigators are responsible for results in the paper and therefore should invest in verification of results, which improves earlier detection of problems even if these problems are the result of mere sloppiness or honest error. Even though it is not feasible for all authors to verify all results, ideally results should be verified by at least one co-author. As mentioned earlier, peer-review does not weed out all major problems (Bornmann et al., 2008) and should not be trusted blindly.

Institutions could facilitate detection of data fabrication and falsification by implementing data auditing. Data auditing is the independent verification of research results published in a paper (Shamoo, 2006). This goes hand-in-hand with co-authors verifying results, but this is done by a researcher not directly affiliated with the research project. Auditing data is common practice in research that is subject to governmental oversight, for instance, drug trials that are audited by the Food and Drug Administration (Seife, 2015).

Papers that report fabricated or falsified data are typically retracted. The decision to retract is often (albeit not necessarily) made after the completion of a formal inquiry and/or investigation of research misconduct by the academic institution, employer, funding organization, and/or oversight body. Because much of the academic work is done for hire, the employer can request a retraction from the publisher of the journal in which the article appeared. Often, the publisher then consults with the editor (and sometimes also with proprietary organizations like the professional society that owns the journal title) to decide on whether to retract. Such processes can be legally complex if the researcher who was guilty of research misconduct opposes the retraction. The retraction notice ideally should provide readers with the main reasons for the retraction, although quite often the notices lack necessary information (Van Noorden, 2011). The popular blog Retraction Watch normally reports on retractions and often provides additional information on the reasons for retraction that other parties involved in the process (co-authors, whistleblowers, the accused researcher, the [former] employer, and the publisher) are sometimes reluctant to provide (Marcus & Oransky, 2014). In some cases, the editors

of a journal may decide to publish an editorial expression of concern if there are sufficient grounds to doubt the data in a paper that is being subjected to a formal investigation of research misconduct.

Many retracted articles are still cited after the retraction has been issued (Bornemann-Cimenti, Szilagyi, & Sandner-Kiesling, 2015; Pfeifer & Snodgrass, 1990). Additionally, retractions might be issued following a misconduct investigation, but not completed by journals, that the original content is simply deleted, or that legal threats resulted in not retracting the work (Elia, Wager, & Tramèer, 2014). If retractions do not occur even though they have been issued, their negative effects, for instance, decreased author citations (Lu, Jin, Uzzi, & Jones, 2013), are nullified, reducing the costs of committing misconduct.

CONCLUSION

This article provides an overview of the research practice spectrum, where on the one end there is responsible conduct of research and with research misconduct on the other end. In sum, transparent research practices are proposed to embody scientific norms and a way to deal with both questionable research practices and research misconduct, inducing better research practices. This would improve not only the documentation and verification of research results; it also helps create a more open environment for researchers to actively discuss ethical problems and handle problems in a responsible manner, promoting good research practices. This might help reduce both questionable research practices and research misconduct.

REFERENCES

(2012). Retraction of "the secret life of emotions" and "emotion elicitor or emotion messenger? Subliminal priming reveals two faces of facial expressions." *Psychological Science*, *23*(7), 828, doi:10.1177/0956797612453137

Allen, M., & Dowell, R. (2013). Retrospective reflections of a whistleblower: Opinions on misconduct responses. *Accountability in Research, 20*(5–6), 339–348. doi:10.1080/08989621.2013.822249

American Psychological Association. (2010). *Ethical principles of psychologists and code of conduct*. Retrieved from http://www.apa.org/ethics/code/principles.pdf

Anderson, M. S., Horn, A. S., Risbey, K. R., Ronning, E. A., De Vries, R., & Martinson, B. C. (2007a). 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: Journal of the Association of American Medical Colleges*, *82*(9), 853. doi:10.1097/ACM.0b013e31812f764c

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

Anderson, M. S., Ronning, E. A., Devries, R., & Martinson, B. C. (2010). Extending the Mertonian norms: Scientists' subscription to norms of research. *The Journal of Higher Education*, *81*(3), 366–393. doi:10.1353/jhe.0.0095

Armitage, P., McPherson, C. K., & Rowe, B. C. (1969). Repeated significance tests on accumulating data. *Journal of the Royal Statistical Society. Series A*, *132*(2), 235–244. doi:10.2307/2343787

Bakker, M., & Wicherts, J. M. (2011). The (mis)reporting of statistical results in psychology journals. *Behavior Research Methods*, *43*(3), 666–678. doi:10.3758/s13428-011-0089-5

Bik, E. M., Casadevall, A., & Fang, F. C. (2016). The prevalence of inappropriate image duplication in biomedical research publications. *MBio*, *7*(7), e00809-16. doi:10.1128/mbio.00809-16

Bornemann-Cimenti, H., Szilagyi, I. S., & Sandner-Kiesling, A. (2015). Perpetuation of retracted publications using the example of the Scott S. Reuben case: Incidences, reasons and possible improvements. *Science and Engineering Ethics,* 1–10. doi:10.1007/s11948-015-9680-y

Bornmann, L., Nast, I., & Daniel, H.-D. (2008). Do editors and referees look for signs of scientific misconduct when reviewing manuscripts? A quantitative content analysis of studies that examined review criteria and reasons for accepting and rejecting manuscripts for publication. *Scientometrics*, *77*(3), 415–432. doi:10.1007/s11192-007-1950-2

Buyse, M., George, S. L., Evans, S., Geller, N. L., Ranstam, J., Scherrer, B., et al. (1999). The role of biostatistics in the prevention, detection and treatment of fraud in clinical trials. *Statistics in Medicine*, *18*(24), 3435–3451. doi:10.1002/(SICI)1097-0258(19991230)18:24<3435::AID-SIM365>3.0.CO;2-O

Carlisle, J. B. (2012). The analysis of 168 randomised controlled trials to test data integrity. *Anaesthesia*, *67*(5), 521–537. doi:10.1111/j.1365-2044.2012.07128.x

Carlisle, J. B., Dexter, F., Pandit, J. J., Shafer, S. L., & Yentis, S. M. (2015). Calculating the probability of random sampling for continuous variables in submitted or published randomised controlled trials. *Anaesthesia*, *70*(7), 848–858. doi:10.1111/anae.13126

Chambers, C. D. (2015). Ten reasons why journals must review manuscripts before results are known. *Addiction*, *110*(1), 10–11. doi:10.1111/add.12728

Cohen, J. (1994). The earth is round (p<.05). *American Psychologist*, *49*(12), 997–1003. doi:10.1037/0003-066X.49.12.997

Cressey, D. (2013). 'Rehab' helps errant researchers return to the lab. *Nature News*, *493*(7431), 147. doi:10.1038/493147a

Elia, N., Wager, E., & Tramèer, M. R. (2014). Fate of articles that warranted retraction due to ethical concerns: a descriptive cross-sectional study. *PLoS One, 9*(1), e85846. doi:10.1371/journal.pone.0085846

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., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(42), 17028–17033. doi:10.1073/pnas.1212247109

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505. doi:10.1126/science.1255484

Franco, A., Malhotra, N., & Simonovits, G. (2016). Underreporting in psychology experiments: Evidence from a study registry. *Social Psychological and Personality Science*, *7*(1), 8–12. doi:10.1177/1948550615598377

Haldane, J. B. S. (1948). The faking of genetical results. *Eureka*, *6*, 21–28.

Hettinger, T. P. (2010). Misconduct: Don't assume science is self-correcting. *Nature*, *466*(7310), 1040. doi:10.1038/4661040b

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524–532. doi:10.1177/0956797611430953

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196–217. doi:10.1207/s15327957pspr0203_4

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B, Jr., Bahník, S., Bernstein, M. J., et al. (2014). Investigating variation in replicability. *Social Psychology*, *45*(3), 142–152. doi:10.1027/1864-9335/a000178

Koppelman-White, E. (2006). Research misconduct and the scientific process: Continuing quality improvement. *Accountability in Research*, *13*(3), 225–246. doi:10.1080/08989620600848611

Kornfeld, D. S. (2012). Research misconduct: The search for a remedy. *Academic Medicine: Journal of the Association of American Medical Colleges*, *87*(7), 877–882. doi:10.1097/ACM.0b013e318257ee6a

Krawczyk, M., & Reuben, E. (2012). (Un)available upon request: Field experiment on researchers' willingness to share supplementary materials. *Accountability in Research*, *19*(3), 175–186.

Levelt Committee, Drenth Committee, and Noort, Committee. (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel.* Technical report. Tilburg (the Netherlands).

Lu, S. F., Jin, G. Z., Uzzi, B., & Jones, B. (2013). The retraction penalty: Evidence from the web of science. *Scientific Reports*, *3*, 3146. doi:10.1038/srep03146

Lubalin, J. S., Ardini, M.-A. E., & Matheson, J. L. (1995). *Consequences of whistleblowing for the whistleblower in misconduct in science cases.* Technical report. Washington, DC: Research Triangle Institute.

Lubalin, J. S., & Matheson, J. L. (1999). The fallout: What happens to whistleblowers and those accused but exonerated of scientific misconduct? *Science and Engineering Ethics*, *5*(2), 229–250. doi:10.1007/s11948-999-0014-9

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on psychological science: a journal of the Association for Psychological Science*, *7*(6), 537–542. doi:10.1177/1745691612460688

Marcus, A., & Oransky, I. (2014). What studies of retractions tell us. *Journal of Microbiology & Biology Education*, *15*(2), 151–154. doi:10.1128/jmbe.v15i2.855

Margraf, J. (2015). Zur lage der psychologie. *Psychologische Rundschau; Ueberblick uber die Fortschritte der Psychologie in Deutschland, Oesterreich, und der Schweiz*, *66*(1), 1–30. doi:10.1026/0033-3042/a000247

Merton, R. K. (1942). A note on science and democracy. *Journal of Legal and Political Sociology, 1,* 115.

Mitroff, I. I. (1974). Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists. *American Sociological Review*, *39*(4), 579–595. doi:10.2307/2094423

Mosimann, J., Dahlberg, J., Davidian, N., & Krueger, J. (2002). Terminal digits and the examination of questioned data. *Accountability in Research*, *9*(2), 75–92. doi:10.1080/08989620212969

Mosimann, J. E., Wiseman, C. V., & Edelman, R. E. (1995). Data fabrication: Can people generate random digits? *Accountability in Research*, *4*(1), 31–55. doi:10.1080/08989629508573866

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., et al. (2015). Promoting an open research culture, *Science*, *348* (6242), 1422–1425. doi:10.1126/science.aab2374

Nosek, B. A., & Bar-Anan, Y. (2012). Scientific utopia: I. opening scientific communication. *Psychological Inquiry*, *23*(3), 217–243. doi:10.1080/1047840X.2012.692215

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(6), 615–631. doi:10.1177/1745691612459058

Nuijten, M. B., Hartgerink, C. H. J., Van Assen, M. A. L. M., Epskamp, S., & Wicherts, J. M. (2015). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, 1–22. doi:10.3758/s13428-015-0664-2

Office of Science and Technology Policy. (2000). *Federal policy on research misconduct.* Retrieved September 10, 2015, from https://web.archive.org/web/20150910131244/https://www.federalregister.gov/articles/2000/12/06/00-30852/executive-office-of-the-president-federal-policy-on-research-misconduct-preamble-for-research#h-16.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.

Panel on Scientific Responsibility and the Conduct of Research. (1992). *Responsible science, volume I: Ensuring the integrity of the research process.* Washington, DC: National Academies Press.

Peeters, C. A. W., Klaassen, C. A. J., & Van de Wiel, M. A. (2015). *Meta-response to public discussions of the investigation into publications by Dr. Förster*. Technical report. Amsterdam (the Netherlands): University of Amsterdam.

Peiffer, A. M., Hugenschmidt, C. E., & Laurienti, P. J. (2011). Ethics in 15 min per week. *Science and Engineering Ethics*, *17*(2), 289–297. doi:10.1007/s11948-010-9197-3

Pfeifer, M. P., & Snodgrass, G. L. (1990). The continued use of retracted, invalid scientific literature. *JAMA*, *263*(10), 1420–1423. doi:10.1001/jama.1990.03440100140020

Plemmons, D. K., Brody, S. A., & Kalichman, M. W. (2006). Student perceptions of the effectiveness of education in the responsible conduct of research. *Science and Engineering Ethics*, *12*(3), 571–582. doi:10.1007/s11948-006-0055-2

Price, A. R. (1998). Anonymity and pseudonymity in whistleblowing to the U.S. office of research integrity. *Academic Medicine: Journal of the Association of American Medical Colleges*, *73*(5), 467–472. doi:10.1097/00001888-199805000-00009

Resnik, D. B., & Stewart, C. N., Jr. (2012). Misconduct versus honest error and scientific disagreement. *Accountability in Research*, *19*(1), 56–63.

Rhoades, L. J. (2004). *ORI closed investigations into misconduct allegations involving research supported by the public health service: 1994–2003*. Technical report. Washington, DC: Office of Research Integrity.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638. doi:10.1037/0033-2909.86.3.638

Rossner, M., & Yamada, K. M. (2004). What's in a picture? The temptation of image manipulation. *The Journal of Cell Biology*, *166* (1), 11–15. doi:10.1083/jcb.200406019

Ruys, K. I., & Stapel, D. A. (2008). Emotion elicitor or emotion messenger?: Subliminal priming reveals two faces of facial expressions [retracted]. *Psychological Science*, *19*(6), 593–600. doi:10.1111/j.1467-9280.2008.02128.x

Savage, C. J., & Vickers, A. J. (2009). Empirical study of data sharing by authors publishing in PLoS journals. *PLoS One, 4*(9), e7078. doi:10.1371/journal.pone.0007078

Seife, C. (2015). Research misconduct identified by the US food and drug administration: out of sight, out of mind, out of the peer-reviewed literature. *JAMA Internal Medicine*, *175*(4), 567–577. doi:10.1001/jamainternmed.2014.7774

Shamoo, A. E. (2006). Data audit would reduce unethical behaviour. *Nature*, *439*(7078), 784. doi:10.1038/439784c

Shamoo, A. E., & Resnik, D. B. (2009). *Responsible conduct of research*. New York, NY: Oxford University Press.

Sijtsma, K., Veldkamp, C. L. S., & Wicherts, J. M. (2015). Improving the conduct and reporting of statistical analysis in psychology. *Psychometrika, 81*, 33–38.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. doi:10.1177/0956797611417632

Simonsohn, U. (2013). Just post it: The lesson from two cases of fabricated data detected by statistics alone. *Psychological Science*, *24*(10), 1875–1888. doi:10.1177/0956797613480366

Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, *12*(1), 53–74. doi:10.1007/s11948-006-0006-y

Stewart, W. W., & Feder, N. (1987). The integrity of the scientific literature. *Nature*, *325*(6101), 207–214. doi:10.1038/325207a0

Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific misconduct and the myth of self-correction in science. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(6), 670–688. doi:10.1177/1745691612460687

The Journal of Cell Biology. (2015). *About the journal.* Retrieved September 10, 2015, from https://web.archive.org/web/20150911132421/http://jcb.rupress.org/site/misc/about.xhtml.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. doi:10.1126/science.185.4157.1124

Van Assen, M. A. L. M., Van Aert, R. C. M., Nuijten, M. B., & Wicherts, J. M. (2014). Why publishing everything is more effective than selective publishing of statistically significant results. *PLoS One, 9*(1), e84896. doi:10.1371/journal.pone.0084896

Van Noorden, R. (2011). Science publishing: The trouble with retractions. *Nature, 478*(7367), 26–28. doi:10.1038/478026a

Veldkamp, C. L. S., Nuijten, M. B., Dominguez-Alvarez, L., Van Assen, M. A. L. M., & Wicherts, J. M. (2014). Statistical reporting errors and collaboration on statistical analyses in psychological science, *PLoS One, 9*(12), e114876. doi:10.1371/journal.pone.0114876

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *7*(6), 632–638. doi:10.1177/1745691612463078

Whitebeck, C. (2001). Group mentoring to foster the responsible conduct of research. *Science and Engineering Ethics*, *7*(4), 541–558, doi:10.1007/s11948-001-0012-z

Wicherts, J. M. (2011). Psychology must learn a lesson from fraud case. *Nature*, *480*(7375), 7. doi:10.1038/480007a

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *The American Psychologist*, *61*(7), 726–728. doi:10.1037/0003-066X.61.7.726

Wicherts, J. M., & Van Assen, M. A. L. M. (2012). Research fraud: Speed up reviews of misconduct. *Nature*, *488*(7413), 591. doi:10.1038/488591b

Wicherts, J. M., Hartgerink, C. H. J., & Grasman, R. P. P. P. (2016). *The growth of psychology and its corrective mechanisms: A bibliometric analysis (1950–2015).* Manuscript in preparation. Retrieved from osf.io/ah8k7

Wigboldus, D. H. J., & Dotsch, R. (2016). Encourage playing with data and discourage questionable reporting practices. *Psychometrika*, *81*(1), 27–32. doi:10.1007/s11336-015-9445-1

COMPETING INTERESTS

The authors declare no competing interests.

PUBLISHING NOTES