Title
Academic misconduct, misrepresentation and gaming: A reassessment

Journal
Research Policy, 48(2)

ISSN
0048-7333

Authors
Biagioli, Mario
Kenney, Martin
Martin, Ben R
et al.

Publication Date
2019-03-01

DOI
10.1016/j.respol.2018.10.025

Peer reviewed

Forthcoming in *Research Policy* 2019

Academic Misconduct, Misrepresentation and Gaming: A Reassessment

Mario Biagioli

Center for Science & Innovation Studies

University of California, Davis

Davis, CA 95616

Martin Kenney*

Department of Human Ecology

University of California, Davis

Davis, CA 95616

Ben R. Martin

Science Policy Research Unit (SPRU)

University of Sussex, Brighton, UK

and

John P. Walsh

School of Public Policy

Georgia Institute of Technology

Atlanta, Georgia

* Corresponding author: mfkenney@ucdavis.edu

Abstract

The motivation for this Special Issue is increasing concern not only with academic misconduct but also with less easily defined forms of misrepresentation and gaming. In an era of intense emphasis on measuring academic performance, there has been a proliferation of scandals, questionable behaviors and devious stratagems involving not just individuals but also organizations, including universities, editors and reviewers, journal publishers, and conference organizers. This introduction first reviews the literature on the prevalence of academic misconduct, misrepresentation and gaming (MMG). The core of the article is organized around a life-cycle model of the production and dissemination of research results. We synthesize the findings in the MMG literature at the level of the investigator or research team, emphasizing that misbehavior extends well beyond fabrication and falsification to include behaviors designed to exaggerate or to mislead readers as to the significance of research findings. MMG is next explored in the post-research review, publication, and post-publication realms. Moving from the individual researcher to the organizational level, we examine how MMG can be engaged in by either journals or organizations employing or funding the researchers. The changing institutional environment including the growth of research assessment exercises, increased quantitative output measurement and greater pressure to publish may all encourage MMG. In the final section, we summarize the main conclusions and offer suggestions both on how we might best address the problems and on topics for future research.

Keywords: academic misconduct; research fraud; academic misrepresentation; gaming; predatory journals; pseudo conferences

1. Introduction

The motivation for this Special Issue is increasing concern with academic misconduct (and perhaps also a growing incidence of misconduct) along with a broadening of that concern to include less easily defined forms of misrepresentation and gaming. The sheer enormity, diversity, and complexity of the academic research enterprise ensure that any overview of academic misconduct, misrepresentation and gaming (MMG), however long, may struggle to capture all the facets of the issue. Academic MMG has an extensive history (Merton, 1973; Steneck, 1994). Indeed, as Merton, (1973, p.323) pointed out in his discussion of the importance of priority, the science system itself is potentially pathogenic given the high premium placed on originality and the difficulty of producing careful, original research. He also warned of the gaming that was bound to emerge around metrics indicators due to “goal displacement” (Csiszar, 2019). In recent years, besides clearly forbidden forms of misconduct (particularly fabrication, falsification, and plagiarism), we have witnessed a proliferation of other forms of gaming of the academic system, some serious enough to be considered as detrimental research practices or even as formal misconduct, while other liminal practices are raising concerns but also provoking disagreement about whether they constitute inappropriate behavior or not (cf. Mohliver, 2018). Should one view these new questionable practices as isolated developments or instead as symptoms of an ongoing and long-term expansion of misconduct beyond the traditional focus on fabrication, falsification, and plagiarism? Are increasing pressures for improved performance and a belief that performance can and should be measured through easily applied metrics (Edwards and Roy, 2017) creating incentives to push the boundaries of what is acceptable behavior in order to secure grants, publish in leading journals and secure more citations (Bedeian et al., 2010)? Is science itself suffering, particularly in fields such as the biomedical sciences where the rewards are especially large due to commercial opportunities and interests, but where the risks of dubious research results are also great in terms of potential patient harm (Ioannidis, 2012)?

Technological change has opened up new opportunities for misconduct and gaming. For example, manipulating figures, selectively mining data and copying other author's texts are all greatly facilitated by software tools. In the pre-computer era, researchers had to copy and manipulate by hand. Now, cheap and simple digital tools offer more sophisticated and productive ways of gaming the system, making it trivially easy to recycle large blocks of text (yours or someone else's) to increase the rate of article production.[1] Similarly, digital technologies greatly facilitate the methods for falsifying or misrepresenting results. For example, with the diffusion of Photoshop and similar tools, Western blots can be altered to fit desired research results, such as adding bands as needed, and even recycling the same blots in different articles to support different claims. The result is the accelerated production of fraudulent papers. If in the past fraud was artisanal, new digital tools are almost turning it into a routine or automated process (Biagioli, 2019). Hence, technological developments as well as changes in the governance, metrics of evaluation, and resource environment of science may all be contributing to increasing opportunistic if not devious behavior among researchers.

This Special Issue solicited articles dealing with any issue in the broad area of academic misconduct (for a recent review, see Macfarlane et al., 2014) or what might be termed 'academic misrepresentation' – i.e. new professional practices that many academics would see as unethical but which have not as yet been (and indeed may never be) categorized as misconduct (see the discussion in Jeremy Hall and Ben Martin, this issue). In an era characterized by an emphasis on measuring academic performance, there has been a proliferation of scandals, questionable behaviors and devious stratagems involving not only individuals but also organizations, whether it be universities 'fine-tuning' data to score well on global university rankings, or editors and reviewers engaging in coercive citation and other abuses, or journal publishers creating so-called

[1] At the same time, detecting such misbehaviors has been be facilitated by computers, such as plagiarism detection software and statistical techniques that track unlikely results, raising the possibility that some of the increased attention is simply a result of the enhanced ability to spot violations or of the ready availability of internet avenues to 'name and shame' those caught.

‘predatory’ or ‘pseudo’ journals, or conference organizers offering ‘pay-to-play’ conferences and guaranteed publication of the presenters’ abstracts. Given the importance for the research community of confidence and trust in the academic system, it is remarkable that there have not been more studies exploring the scale of MMG and the rapidity with which such practices are evolving.

Literature on scientific fraud has often previously focused on individual high-profile cases (Broad et al., 1983; Kevles, 2000; Reich 2009). Things may be changing, however, following the introduction of performance metrics and research assessment exercises. What we are apparently witnessing now is far more metrics-oriented misconduct or questionable behavior aimed not so much at producing a high-profile publication but rather at incrementally increasing an individual’s or organization’s reputation, for example, through self-citation or departmental faculty members citing each other, sometimes with the connivance of their university managers. Individually, these may represent relatively minor transgressions, but in aggregate they could create significant advantage for some to the disadvantage of others.

Traditional definitions of misconduct focus on activities relating to the production of a publication, namely fabrication, falsification and plagiarism (Martin, 2007; NAS, 2017). However, there are also ‘pre-production’ forms of misconduct, such as the pirating of grant applications. In addition, there is considerable ‘post-production’ misconduct and misrepresentation focused not on the production of a publication but on the maximization of its impact and its metrics footprint (Biagioli, 2016). To organize our discussion of this wide-ranging set of deviant behaviors, we employ a life-cycle model of research.

We begin in Section 2 with a review of the literature on the prevalence of MMG. In Section 3, we examine where MMG at the level of the investigator or research team might occur at different stages in the life-cycle of the production and dissemination of research results. The enumeration of such behaviors extends far beyond fabrication and falsification (i.e. clearly false research findings) to include behaviors designed to exaggerate or even mislead readers as to the significance of research findings. We then examine the review and publication process as well as post-publication gaming. In Section 4, we shift the focus from the individual researcher to the organizational level, and examine how misconduct and gaming may be practiced by journals or by organizations employing or funding the researcher. In Section 5, we consider how the changing institutional environment characterized by performance metrics, research assessment exercises and ever greater pressures to publish may be affecting MMG. In Section 6, we discuss MMG detection, enforcement, and potential remedies. The final section summarizes the main conclusions, including suggestions on how to reduce the level of MMG and opportunities for further research.

2. How Common Are These Practices?

There have been various recent studies attempting to measure the rates of MMG (e.g. Fong and Wilhite, 2017). While recognizing the inherent limitations of such research, these studies provide some context on the types and incidence of various forms of MMG, as well as clues to the causes or possible interventions. For example, several studies have documented the rapid increase in the number of retracted papers over recent decades. In a study of what they term an "ongoing retraction epidemic", Fang et al. (2012, p.17028) found that for 2,047 biomedical retractions 67% were for reasons related to academic misconduct. Moreover, "the percentage of scientific articles retracted because of fraud has increased ~10-fold since 1975".[2] Using reports of misconduct, the NSF Inspector General found in 2013 that over the prior 10 years the office had observed a tripling of allegations received per year, with a concomitant increase in the number of findings of misconduct (Lerner, 2013).

One strategy to track MMG uses survey methods to ask about self-reported or 'observed' misconduct, misrepresentation or gaming. For example, Martinson et al. (2005), in a survey of NIH-funded researchers, found about one third self-reported engaging in at least one form of

[2] Van Noorden (2011) also reported a 10-fold increase in retractions over the first decade of the 21st Century.

misconduct or detrimental research practice.[3] They argued that the prevalence of such behavior suggests the need to move beyond individualist explanations and interventions and to focus on the broader systemic factors influencing scientific research practices, an issue discussed in detail by Hall and Martin (this issue). Similarly, Bergh et al. (2017) surveyed management scholars about their rates of engaging in misconduct and detrimental research practices. They find that self-reported fabrication is quite rare (less than 1%). However, 11% claim to have rounded off p values, and one-third reported selectively excluding data or including control variables to find statistically significant results. About half of the respondents state that they selectively reported hypotheses or engaged in HARKing (hypothesizing after the results are known – Kerr, 1998). A meta-analysis by Fanelli (2009) shows about 2% of respondents confessed to engaging in serious misconduct and about one-third to engaging in other detrimental research practices.[4] However, as Fanelli notes, given the sensitive nature of this topic, intrusive research methods such as surveys are likely to substantially under-estimate the scale of the problem.

Studies based on self-reports suggest that outright fabrication may be rare, but that a whole range of research practices that might affect the research results (and perhaps their placement in high profile journals) are not uncommon. In contrast, data based on second-hand knowledge suggests that such practices are comparatively widespread (Bedeian et al., 2010). Indirect experience with misconduct and misrepresentation may then affect one's willingness to engage in such behaviors, partly because they become normalized and partly because failing to do so may be seen as putting oneself at a significant competitive disadvantage.

---

[3] In this section, we sometimes refer to 'detrimental research practices' (as distinct from 'misconduct'), since this term has often been used in previous literature. However, in the rest of this article, we tend to adopt the term 'academic misrepresentation'.

[4] This is supported by a recent survey of biostatisticians, with around a quarter reporting that they had been requested by investigators in the studies in which they had been involved to engage in various forms of "inappropriate analysis and reporting", such as "removing or altering some data records to better support the research hypothesis" (24%) and "not reporting the presence of key missing data that might bias the results" (24%) (Wang et al., 2018).

Because of the limitations of the evidence provided by retractions or self-reported misconduct, a few studies attempt to measure misconduct or misrepresentation using analyses of the publications themselves. For example, using plagiarism detection software Honig and Bedi (2012) find 25% of papers accepted at the Academy of Management's International Management division had some evidence of plagiarism, with 14% exhibiting substantial plagiarism. Lewellyn et al. (2017) report that 34% of papers submitted to the Academy of Management IM section violated the rule against republishing or self-plagiarism. Park et al. (2017) conclude that about 20% of results in the National Institute of Standards and Technology (NIST) materials science database are statistically unlikely, although they are careful not to attribute causes or motives. Bergh et al. (2017) attempted to reproduce the findings from papers in *Strategic Management Journal*. They observe that about 70% of papers did not provide sufficient data to reproduce the findings, while for cases where replication was possible, about one-third had statistically significant hypothesis tests that could not be reproduced. Although there are some disagreements about the meaning and causes of the irreproducibility of research, there is considerable evidence of detrimental research practices and of journal articles consequently having significant problems in their results.

One potential cause of deviant behaviors is disagreement about whether such behaviors are actually inappropriate (Lewellyn et al. 2017; Mohliver, 2018). Banks et al. (2015) conducted surveys of management scholars to gauge their views on various detrimental or questionable research practices. For many of those practices, a significant minority disagreed with the statement that the particular behavior was inappropriate. For example, 13% disagreed that rounding down the p-value of a significance test was inappropriate; 23% disagreed that selectively reporting hypotheses on the basis of whether they were statistically significant was inappropriate; 21% disagreed that excluding data after looking at the impact of doing so on the results was inappropriate; and 25% disagreed that HARKing was inappropriate. In fact, some of the increased scrutiny of findings in fields such as social psychology may derive from a shift in norms about the uses of statistical testing and data exploration (Bartlett, 2018).

While some might view such behaviors as harmless polishing of results, Simmons et al. (2011) show how susceptible research findings are to such practices. They use the term "researcher degrees of freedom" to describe the various decisions that researchers make about measures, sample sizes, models, outliers, variable transformations, etc. The disagreements described above regarding the appropriateness of particular behaviors suggest that drawing a clear line between acceptable and unacceptable practices may be impossible to do in a systematic fashion. Calls for more standards and stricter enforcement tend to emphasize an absolutist standard, but as the National Academy of Sciences (2017) points out, operationalizing such definitions can prove difficult. Discussions in the sociology of science echo this skepticism about the ability to draw clear distinctions between whether something was 'misconduct' or only a 'detrimental research practice', or maybe even an 'honest mistake' (Zuckerman, 1988; Hackett, 1994; Butler et al., 2017; Hall and Martin, this issue).

The term 'epistemic risk' has been used by Biddle (2018) to discuss the ways in which the exercise of researcher degrees of freedom can change the likelihood of either rejecting or failing to reject a null hypothesis. For example, when estimating the toxicity of GMOs on rats, is a 90-day observation cycle sufficient, or is a 2-year cycle more appropriate? Furthermore, Biddle argues that interested scientists (especially in cases of research related to industry regulation) can produce seemingly objective results in support of their interests by being 'strategic' in exercising these researcher degrees of freedom.

3. The Life Cycle of Research Misconduct, Misrepresentation and Gaming

3.1 Stage 1: The Research Process

We begin with a discussion of misconduct and misrepresentation occurring during the research process. The most obvious example is fabrication of data, material evidence, and research results. Such fabrications have a long history, including famous scientific frauds such as Piltdown Man (Oakley and Weiner, 1955). In recent high-profile cases, researchers have been caught fabricating results in numerous papers. Jan Hendrik Schön**,** Yoshihiro Sato and Diederick Stapel are examples, spanning physics, biomedical research, and social psychology. While such cases are dramatic, these are probably best understood as significant individual pathologies, although, as John Walsh, You-Na Lee and Li Tang (this issue) point out, this does raise the question of what aspects of the collaborations allowed such widespread fabrications to escape detection by the co-authors involved in many of the fabricated papers (see also Beasley et al., 2002). In the case of Paolo Macchiarini discussed by Berggren and Karabag (this issue), the numerous co-authors involved were apparently unaware of how the data in their articles was produced. The spatial organization of research may be a factor. Fraud seems to be only a minor worry in a field like particle physics, where scientists work together in close proximity at the same facility, taking turns at operating the same equipment. The problems exemplified by the Macchiarini case are unlikely to emerge in a research context where the tightness of the collaboration produces close mutual supervision.

Another form of serious misconduct is falsification. Here, however, the lines between acceptable and unacceptable behavior begin to blur. As noted above, practices such as selectively reporting results or trimming outliers are common and things that at least a substantial minority of researchers feel can be appropriate (Banks et al., 2015). One might view this through the lens of deception: is the activity designed to more clearly see a result, or to create a result that is not really there? Alternatively, are the scientists aware of the decision points in their research and actively making a choice in such a way as to reduce epistemic risk, or are they instead choosing to maximize the chances of publication, especially in high-impact journals (Bergh et al., 2017)? More worrying but relatively common is that researchers may fail to report experiments yielding results not in keeping with their hypotheses (with particularly adverse potential effects in the case of drug or pesticide trials). The decision not to report research results is even more serious when the reason is that the research sponsor (often a commercial entity) discourages reporting because it might cause that sponsor financial harm (Salandra, 2018).

While outright data falsification is clearly a serious violation of ethics, somewhat short of this are a range of activities aimed at 'puffing up' results to increase their impact such as p-hacking and HARKing. Simmons et al. (2011) give four examples where, taking advantage of 'researcher degrees of freedom', those involved may engage in practices that might help puff up a result, increasing its significance. A researcher could have two or more measures of the dependent variable, and report only on the one that 'worked'. A researcher could collect the data, and test for significant effects, and if they are not found, increase the sample size and test again, stopping when there is 'enough data to see the effect'. The researcher could also include or drop control variables, keeping those in the model that make the hypothesized effect significant. The researcher may also make choices about including certain interaction terms. As noted above, there is disagreement whether such behaviors are detrimental, or whether this is how good science is done. While each of these choices may seem minor and may be made with good reason, they can have the unintended (or perhaps intended?) effect of dramatically increasing the rates of false positives, especially when based on information from the given dataset (for example, testing for significance and then, if none is found, increasing the sample size and testing again). Simulations of experimental psychology results show that what may seem like innocuous choices can substantially increase the rates of false positives. Using an alpha of .05, Simmons et al. (2011) find some 10% of results in their simulation were false positives, with even higher rates of false positives (over 60%) when the various methodological decisions are combined. Yet such practices may be quite common. Some of the findings that cannot be replicated, and which may result in retractions, might be attributed to 'errors', with the implication that the authors lacked intent, and therefore culpability. However, as Zuckerman (1988) points out, failure to carefully check one's work and to replicate it in-house before publication can shift such errors from being 'reputable' to 'disreputable' errors.

At the same time, one could argue that error (or even undetected malfeasance by a collaborator) may sometimes be a rational action for a productivity- and reputation-seeking scientist. The incentives for rapid publication may be so great that they overwhelm the more cautious strategy of double-checking experimental results to ensure accuracy. Indeed, such care may not be a rational strategy in the face of severe scientific competition, the potential rewards garnered by success, and the severe possible consequences of 'failure' (e.g. failing to publish first).[5] The variety and complexity of misconduct and misrepresentation suggest that addressing them must be tailored to understanding complex and sometime contradictory incentives to which the various actors are subject, including not only researchers but also universities and other research organizations, journals, and funding agencies.

3.2 Stage 2: The Review Process

Several studies have examined misconduct by referees during the peer review process (Frey, 2003; Shibayama and Baba, 2015; Bergh et al., 2017; De Silva and Vance, 2017). Peer review is predicated upon the assumption that researchers in the field of the submission are best able to judge its quality. Referees are usually uncompensated, undertaking reviews as part of their academic duty. There is a continuing concern about whether this is the best system for determining the quality and importance of the reported scientific research (Lee et al., 2013). Unfortunately, peer review also provides opportunities for gaming or fraud. Some studies have shown that peer review can be swayed by positive bias due to mentorship networks (Colussi 2018), or that referees may be unfairly harsh when reviewing their competitors (De Silva and Vance, 2017). A referee asking for an additional experiment, or for re-analyses with different controls to test the robustness, may be viewed as trying to ensure the quality of the published literature, including investigating any possible puffing-up or devious use of the researcher degrees of freedom. At the same time, it is possible that the referee is retarding the publication of

---

[5] We thank Richard Freeman for this suggestion.

a rival's paper, while working on their own publication on that topic. Given that competitors are likely to be chosen as reviewers based on their familiarity with the research topic, such conflicts of interest may be common, and unlikely to be detected. Another reviewing-related behavior involves requesting the author cite the referee's papers. While editors have some ability to limit such demands, such requests can often appear reasonable given the common research interests. At the same time, such requests can increase the citation counts for the referee – something that is crucially important in an increasingly metricized academic performance regime, an issue to which we return later. Reviewer misconduct in its various guises is sufficiently widespread for it to even have an internet meme, 'the Reviewer 2 problem'.[6]

Such behavior can move from unethically delaying rivals to inducing detrimental research practices. To estimate the frequency of these problems, Bergh et al. (2017) surveyed management researchers as to whether they had been encouraged (or instructed) to engage in detrimental practices by a reviewer or editor. They find only 1% of respondents had been asked to 'round off' a p-value (and none had been asked to falsify data). However, 40% reported they had been asked to selectively report a hypothesis, 33% to HARK, 14% to selectively include control variables, and 10% to exclude data *post hoc*. Bergh et al. argue that one possible motivation for such practices is to increase the impact of the published paper, and thus the ranking of the journal (see below).

This raises the question of why authors might comply with such requests. Frey (2003) argues that the current review system, at least in economics, gives referees near veto power over publication in top journals, forcing authors with a strong need for publication, especially young scholars, to engage in what he refers to as 'academic prostitution', i.e. being forced to modify the article or its results against one's better judgement in order to satisfy the reviewer's demands. Since referees have no direct stake in the outcome of the submission, they are free to express

[6] See e.g. https://www.facebook.com/groups/reviewer2/ (accessed 1 October 2018).

their individual preferences with regard to topic, method and style, possibly resulting in papers that are significantly distorted from the author's original intent. Bergh et al.'s (2017) findings suggest that such behavior by reviewers is far from rare.

Shibayama and Baba (2015) use the term "dishonest conformity" to describe authors revising a paper to satisfy referees against the authors' own scientific judgement. In other words, the authors are exercising their researcher degrees of freedom, but in ways they themselves feel distort the results. Shibayama and Baba surveyed Japanese life scientists about a recent paper receiving a 'Revise and resubmit'. They find that 48% reported they had had been asked to engage in research practices against their better judgement and 63% of those engaged in dishonest conformity (i.e. 'academic prostitution'). Moreover, 55% felt that the reviewers' suggestions either made no scientific difference (40%) or actually had a negative effect on the paper (15%). Yet even among those who felt the reviewer's comments had a negative effect on the paper, 47% still complied. Moreover, dishonest conformity is higher in basic fields than applied ones, greater when the author has more competitors, is of lower rank or has fewer citations, or when the paper is in a lower impact journal. As far as we are aware, there are no systematic data on the impact of such coercive reviewing on the validity of scientific findings, but the available evidence suggests that substantial distortions in the scientific record may be produced in this way.

An especially insidious form of reviewer misconduct involves grant reviewing. Here, referees have the ability to read the details of a proposed study, deny funding, then appropriate the idea, conduct similar research and publish before the original proposal can obtain funding in a subsequent round. The plagiarism of grant proposals is particularly hard to uncover because the reviewer lifts ideas rather than an extensive body of text, thus making plagiarism software unlikely to detect similarities between the original proposal and the plagiarized one. Additionally, grant proposals are typically held in closed databases and thus beyond the reach of the plagiarism software (Biagioli, 2012). This form of misconduct is particularly damaging to young researchers who are more grant-dependent than their senior colleagues, while the latter are more likely to function as referees and thus are better placed to plagiarize proposals.

There are also 'acceptance circles',[7] in which groups of authors implicitly or explicitly agree to favorably review each other's papers (and maybe harshly review outsiders), so publication is greatly facilitated by being a member of this circle. At first sight, this strategy would seem to be effective only when an editor accepts an author's suggestions for reviewers (Ferguson et al., 2014; Haug, 2015) or indeed is a member of this 'acceptance circle'. However, even without a reviewer suggestion system, if a 'school' of thought exists, its members can write, edit and review each other's papers, creating a stream of publications far greater than if their papers were put through more wide-ranging peer review. This can be even more powerful if the school is able to seize control of a journal's editorial board or establish their own journal.[8]

In addition to misbehavior by referees, there are cases where authors rigged the review process. For example, Hyung-In Moon at South Korea's Dong-A University managed to publish several articles in respected journals by falsifying the email addresses of referees he recommended to journal editors. Unaware editors sent their review requests to these false email addresses, which were duly completed by the author. Rigged peer review lies behind the retraction of more than 600 articles since 2012 (McCook, 2018). While Professor Moon may not have been accused of producing a fraudulent article and thus may not be in violation of the U.S. Federal misconduct definition, his rigging of peer review nevertheless certainly constitutes fraud in the eyes of colleagues. More insidiously, this is a form of misconduct that is virtually impossible to detect through peer review, given that the peer review process is itself hacked.

3.3 Stage 3: Publication and Post-Publication Exaggeration

[7] We thank Mary Frank Fox for this term.

[8] In the case of the journal *Climate Research*, for example, climate change skeptic Chris de Freitas became one of the editors, and began accepting papers from other climate skeptics, presumably by sending them to like-minded referees (see https://www.theguardian.com/environment/2010/feb/09/peer-review-block-scientific-papers - accessed 1 October 2018).

After the research is complete, the next step is publication and, unsurprisingly, there is ample room here for MMG, ranging from deciding on who should be an author to reporting the claims in the article. Let us examine these in greater detail.

3.3.1. Authorship

With the rise of team science, article authorship has become negotiated and uncertain. In an investigation of papers in top medical journals, Wislar et al. (2011) concluded that 8% were ghost-written while 21% had guest authors (see also Flanagin et al., 1998). Jabbehdari and Walsh (2017) find that, across all fields, about one-third of publications have at least one author who made only a specialized contribution (potentially a guest author), including about 6% of papers with an author whose only contribution was to provide financial support (e.g. the lab head). Such specialized contributions are unlikely to satisfy the authorship requirements specified by the International Committee of Medical Journal Editors (ICMJE). Guest authorship is sometimes the result of an explicit agreement among colleagues in order to pump up the publication list for each. For example Yoshihiro Sato and Jun Iwamoto allegedly agreed to put each other's name on every paper they authored (Kupferschmidt, 2018). In addition, Jabbehdari and Walsh (2017) find that over half of papers had researchers who made a significant contribution to the study but were not included as authors (i.e. uncredited authors). Note that there is significant disagreement across and even within fields as to whether certain roles do or do not merit authorship (Wager, 2009).[9]

Hence, authorship problems include both adding undeserved authors and leaving off those who deserve authorship, both of which subvert the meaning of authorship and have implications for assigning credit and responsibility (Biagioli, 2003). For example, when a problem appears in a paper prompting discussion of potential misconduct, guest or specialist authors sometimes proclaim their innocence due to their limited participation. Thus, when Sato's

[9] For example, Jabbehdari and Walsh (2017) find significant variation by field in the rates of specialist (guest) and uncredited (ghost) authors and the roles that were included in each category.

papers came under scrutiny, Iwamoto claimed he knew nothing about the content of the papers, or even that his name had been added to many of them (Kupferschmidt, 2018). There were similar denials by Baltimore in the Imanishi-Kari case, by Green in the case of LaCour's fabrication, and by Schön's collaborators (see Walsh et al., this issue, for details).

3.3.2. Plagiarism, self-plagiarism and text-recycling

Plagiarism is one of the oldest forms of academic misconduct. Indeed, from Merton's discussion of priority disputes, one might almost view plagiarism as a victimless crime. From the societal point of view, one could even claim that plagiarism is beneficial in that it increases the circulation of scientific research. However, plagiarism looks very different from the point of view of the community of science. The results themselves are not necessarily suspect but rather the credit has been misappropriated, depriving the true author of the recognition s/he deserves while providing the plagiarist with undeserved benefits. These may include employment, promotion, grant funding or other benefits that result from having a longer-than-deserved list of publications. An example here is Hans Gottinger, who obtained several senior posts on the back of a long list of publications, many based on plagiarism (Martin et al. 2007; 2012).[10] Unlike more ambiguous forms of misconduct, plagiarism may meet the legal definition of fraud. Plagiarized publications are used to produce a fraudulent CV, a form of misrepresentation that the plagiarist uses to gain an unlawful advantage over others.

In addition to copying the work of others, there is the rather different problem of self-plagiarism. Because one cannot steal from oneself, the term self-plagiarism is something of a misnomer but the practices it refers to are nevertheless problematic. In many cases, early parts of a manuscript (e.g., literature review, methodology) adopt a 'boiler-plate' approach before getting to the important part, the new findings. Some journals have taken a hard line on self-plagiarism

[10] Despite having been publicly exposed as a serial plagiarist over ten years ago, this individual continues to produce large numbers of publications – see https://www.researchgate.net/profile/Hans_Werner_Gottinger (accessed on 12 October 2018).

(perhaps driven by automated flagging from plagiarism-detection software). In many cases it may be difficult to thoroughly report a research study without re-using some text describing the methods or baseline findings of the study (as in the case of multiple papers from the same survey or suite of experiments). Likewise, it may not be easy to avoid a certain amount of boiler-plate text if publishing outside one's normal disciplinary boundaries, addressing readers who need more background information. Yet while the line between acceptable text recycling and self-plagiarism may be blurred, the key issue is how original the contribution is.

In this SI, building from the controversy that surrounded the work of the noted Dutch economist Peter Nijkamp, who was accused of engaging in publication-recycling on an unprecedented scale, Serge Horbach and Willem Halffman examine text recycling in four research fields, finding it to be more common in economics and psychology than biochemistry and history. In addition, more productive authors are likelier to recycle text from previous publications. For this particular paper, we also publish a comment received from Jasper Lukkezen and a rejoinder by Horbach and Halffman. Essentially, Lukkezen argues that the 'Nijkamp effect' is not characteristic of Dutch economists as a whole, while Horbach and Halffman in their response reconfirm that recycling appears more common in economics, although it may have decreased after the Nijkamp controversy. Clearly, the problem of text recycling will continue – and continue to be hotly debated – as some academics pursue strategies to artificially inflate their publication totals, while others attempt to increase their efficiency in increasing publication counts. Here again, the issue often comes down to whether there is an intent to deceive the reader that what is being presented is original, when there are in fact one or more undeclared parallel papers making 'overlapping contributions' (Martin, 2013). Elsewhere in the SI, Hall and Martin examine the case of a serial self-plagiarist (Ulrich Lichtenthaler), who on the basis of just two studies (his PhD and his Habilitation) produced dozens of overlapping papers (many published in leading journals). This initially escaped the attention of reviewers and editors before the research community became aware of the problem and brought it to the attention of the journals involved. An extensive CV had meanwhile enabled this individual to advance in just a few years from PhD student to full professor, a post he then lost when large numbers of his papers were retracted (see the online Appendix to Hall and Martin, this issue).

3.3.3. Spinning results

The issue of misrepresentation or 'spinning' research findings is also coming under increased scrutiny. Here, one can distinguish between misrepresenting what was done in the research prior to writing the paper (discussed above), and misrepresenting what the paper actually says when presenting it to academic or public audiences (e.g. media interviews, Congressional testimonies). This spinning may involve misreporting the methods or misinterpreting the results. Boutron and Ravaud (2018) find that articles receiving media attention on average receive 75% more citations. This suggests that attracting media attention can increase impact, but conversely may encourage authors to spin their results to attract more media attention. Brandon Stell and Boris Barbour, the co-founders of the prominent fraud-detection platform PubPeer, link this phenomenon to attempts to publish in particularly high JIF outlets such as *Nature*, journals where the criteria for acceptance require papers to present "revolutionary concepts" (Stell and Barbour, 2019). In the biomedical field, this may have particularly egregious effects, as people often act upon media reports – for instance, in the stem cell field where scientist-created hype has led to a substantial market for unproven treatments administered by dubious clinics (Ryan et al. 2015).

4. Organizational Misconduct

The growing emphasis on auditing and metrics to improve research performance may paradoxically encourage potentially detrimental gaming efforts (Oravec, 2017). The spread of the New Public Management to universities (Martin, 2016b) has brought about changes in the governance of universities (especially in Europe) to emphasize performance and visible metrics (Whitley, 2007). Increasingly, university or department funding is closely linked to performance evaluations, pressuring universities to improve their rankings in what is often a zero-sum game (especially when New Public Management is coupled with austerity measures). In the U.S., scoring highly on private ranking systems, especially for faculties such as management schools that are heavily dependent on enrollments for their budgets, can encourage gaming or even outright fraud. Research on organizational behavior shows that incentive systems designed to improve performance do tend to increase misconduct (Hall and Martin, this issue). While one can probably do better in designing such systems, misconduct may be an unfortunate but inevitable consequence of performance-based incentive systems (Larkin and Pierce, 2015).

As impact factors, performance indicators and ranking systems have increasingly come to determine rewards, so the pressures to engage in gaming have escalated (Martin, 2016; Hall and Martin, this issue, Biagioli and Lippman, 2019). Because publications, especially those in high profile journals, are ever more important for departments and universities concerned with rankings, organizations have begun engaging in various practices, ranging from more to less legitimate, in order to improve their visibility and their ranking. Universities and even governments have started paying authors direct cash awards for articles published, and these can be substantial. In China for a publication in *Science* or *Nature*, awards of up to $165,000 have been reported (Quan et al., 2017), although China is certainly not alone in adopting this practice (Franzoni et al., 2011; Abritis et al., 2017).

In this Special Issue, Marco Seeber, Mattia Cattaneo, Michele Meoli and Paolo Malighetti find that the introduction in Italy of a regulation linking career advancement to the number of citations received led to a significant increase in self-citations among scientists able to benefit most from the increase. They conclude that, while “metrics are introduced to spur virtuous behaviors, when not properly designed they favor the usage of questionable practices.” Similarly, Franzoni et al. (2011) found that these direct rewards are often associated with ritualistic submission, flooding journals with low-quality submissions in the hope of getting lucky (i.e. a ‘lottery-ticket’ publication strategy). This behavior increases the burdens on reviewers and editors without necessarily benefitting either the scientists or the research system.

For organizations, fierce competition is resulting in some trying to game metrics-based regimes of academic evaluation. This pressure to game the system seems to have grown with demands for proof of impact stemming from the proliferation of research assessment exercises (Fang et al. 2012; Martin, 2013). With the new impact metrics, are we perhaps witnessing a shift from 'publish or perish' to 'impact or perish' (Biagioli, 2016)? Do new evaluation metrics create new incentives for misconduct? And can we still draw a reasonably clear line between gaming the metrics and engaging in misconduct? Traditional discourses and misconduct policies have been rooted in a dichotomy between truth and falsehood, right and wrong, honest mistake and fabrication. However, new metrics-driven misconduct or misrepresentation seems to be defined by the extent of the gaming involved – gaming that is conducted not just by individual scientists but also by research groups, journals and academic organizations (Biagioli and Lippman, 2019). The important point to note here is that the metrics may be re-aligning the incentives of universities and their staff to encourage such gaming (Berggren and Karabag, this issue).

Various ranking systems of universities have also increased the pressure on organizations to game the system. One way to improve an organization's ranking is to increase pressure on researchers to modify their behavior in ways that can improve rankings. However, organizations have developed other ingenious strategies for bolstering their metric-based performance. One is to 'rent vitas', hiring productive scientists (often for just a short period and for a small fraction of their time) just before the evaluation to increase the institution's perceived productivity and impact, a practice common in UK universities as a consequence of successive Research Assessment Exercises. While not technically misconduct, it suggests an attempt to 'puff up' the university's performance. Other games involve modifying the university's performance to specifically match the ranking criteria. One case involved Boston's Northeastern University where a new president felt the university would benefit from a higher score on the *U.S. News and World Report* academic ranking system[11], and embarked on changing university policies to

[11] The U.S. News ranking system is a private effort and entirely voluntary, and thus quite different from

achieve that goal (Kutner, 2014). This involved reverse-engineering the ranking system, then setting about improving the university's score on all the measures that were amenable to action (such as encouraging more applicants so that they could be rejected, thereby making the university appear more exclusive). While these initiatives were not illegal, they were part of a comprehensive effort to game the metrics. In other cases, the belief in the need to game the system can lead to blatant dishonesty, as in 2013 when Tulane University Business School submitted false information to the *U.S. News and World Report* in order to receive a higher ranking (Ellis, 2013). Similarly, University of Missouri, Kansas City was found to have provided falsified data (e.g. on enrolment figures, student clubs, and student start-ups) in submissions to the Princeton Review Board to inflate the ranking of its Bloch School of Management (see the online Appendix to Hall and Martin, this issue).

Organizational gaming of ranking is not confined to developed nations. To boost its global research ranking, the Saudi Arabian King Abdul-Aziz University hired highly cited international professors as faculty members with very limited duties (two weeks or less per year), but handsome compensation for this limited involvement. As a result, the University was able to dramatically improve its position in the global rankings (Bhattacharjee, 2011), ranking #7 for mathematics, apparently ahead of MIT and the University of Cambridge (Pachter, 2014). In another case, the Dean of the School of Engineering at the University of Malaya instructed his faculty to cite at least three papers by their colleagues, thereby boosting the School's citation profile (McCook, 2017). In China, university ranking systems are at play, in particular the widely used Shanghai Jiaotong University Academic Ranking of World Universities, which was initially funded by the Chinese government to provide 'objective' measures of its universities' performance (Saisana et al., 2011). Goodhart's Law, which postulates that any performance metric will attract gaming and rapidly lose its value as an indicator (Goodhart, 1984), suggests

government-mandated systems such as the UK Research Assessment Exercise (now retitled Research Excellence Framework).

that with the increasing importance of such ranking systems for access to resources, we will experience ever more elaborate gaming schemes. The result is a Mertonian goal-displacement (Csizsar, 2019), shifting the goal of science from producing important and verifiable findings to producing high scores on metrics. Generating large numbers of publications (irrespective of whether these are mediocre or even fraudulent) and citations may come to dominate over producing truly novel but perhaps less immediately recognized work.

4.1. Coercive citation and other JIF gaming

Many journals and their editors also feel entrapped in a race to improve their metrics, in particular, the Journal Impact factor (JIF). While some of the measures to achieve this (e.g. publishing more review articles) might be considered legitimate, others involve more dubious activities, for example, insisting on a certain number of citations to the journal as a condition of acceptance for publication in that journal. Wilhite and Fong (2012) found that 20% of authors in social sciences had been asked to add irrelevant references, and of those, more than half complied. Because the JIF is calculated using a two-year time window, some artful editors have taken quiet advantage of this by publishing online version of articles but then delaying full publication (i.e. with a volume and page number) by one or even two years in order to very substantially inflate their two-year JIF (Tort et al., 2012; Martin, 2016). Similarly, a group of Brazilian scholars developed a scheme to cross-cite other Brazilian journals, thereby collectively raising their JIFs (Van Noorden, 2013), while a year later no less than six business journals had their JIFs suspended for engaging in the same practice of 'citation stacking' (Davis, 2014).

4.2 The Emergence of 'Pseudo' Journals and Conferences

Recently, a new group of organizational actors has emerged, namely firms established to mimic traditional journals and conferences but without the same level of content and oversight. This has been driven by the fact that over the decades, research journals and conferences have become an enormous industry fueled by academics with discretionary funds and under pressure to increase their visibility in order to improve their organizations' rankings. Ten years ago, it was estimated that the total system costs of conducting and communicating the research published in journals was £175 billion, comprising £116 billion for the costs of the research itself, £25 billion for publication, distribution and access to the articles, and £34 billion for reading them (Cambridge Economic Policy Associates, 2008, p.26). These figures are surely much higher today (Buranyi, 2017). The potential for profits in this vast enterprise has generated an industrial niche that produces what in its most egregious form might be termed 'pseudo journals'.

4.2.1. Relationship between 'Open Access' and 'Pseudo Journals'

The decreased costs of online publishing have allowed the creation of two relatively distinct genres of journals. The first, generally referred to as 'open access' journals, retain peer review and the editorial function, but no longer publish physical copies of the journal. Most importantly, they shift the cost of publication from the reader or subscribing library to the author (Shen and Bjork, 2015), typically charging more than $1,000 per article. These journals have disintermediated not only the traditional journals that were compensated through subscriptions (most often from libraries) but effectively also the library. More negatively, the low costs of online publishing have allowed the creation of another genre of what Beall (2012) has described as 'predatory journals', though we prefer the term 'pseudo journals'.[12] These differ from the legitimate open-access journals in that, while charging submitting authors for publication, they make little pretense at exercising peer review or editorial oversight, operating primarily as vehicles for generating income for the publisher. The result for the author, of course, is a 'publication'. While denounced in the more established research organizations, some have argued that these journals may be particularly valuable to researchers from the global South who otherwise find it difficult to publish in top journals (Frandsen, 2017). Besides being review-light, these journals charge publishing fees that are a fraction of those demanded by their more

[12] As is evident from the criticisms levelled at Beall's list of 'predatory journals', there is no simple line delineating reputable journals from 'predatory' or 'pseudo' journals but rather a continuous spectrum. We use the term 'pseudo journals' to label the disreputable end of this spectrum, while recognizing that in between the two extremes lie a range of 'questionable journals'.

legitimate cousins, thus offering a publication venue for scientists whose universities can only provide them with marginal resources while still expecting them to publish in ‘international’ journals. Given the rise of metrics-based evaluations, such ‘publications’ may have some value for the author, at least if the metrics are not discerning about which publications count.

These low-quality journals often find potential authors through mass solicitation emails. Jeffrey Beall, a librarian at the University of Colorado, previously maintained a list of what he termed ‘predatory journals’, but in 2017 he discontinued the list following pressure from publishers on his employer (Basken, 2017). While criticized for the opacity of its classification criteria, Beall’s list nevertheless became remarkably influential by filling an information vacuum. Universities did not (and still do not) know how to quickly assess the legitimacy of all the outlets in which their faculty publish. Beall’s list provided a quick answer, though not necessarily a universally agreed one. Its demise was largely a consequence of its success, illustrating the simultaneous demand for a line to be drawn between legitimate and illegitimate publications, while struggling to make explicit the criteria for that demarcation.

Besides editorials and other opinion pieces about the spread of pseudo or other questionable journals, there have been various field experiments in which pseudo journals have been shown to be willing to commit egregious, almost comical, violations of scientific norms. For example, Bohannon (2013) reports submitting a bogus article to 304 open-access journals and receiving acceptances from over 50%. In another exercise, researchers concocted a professor, Anna O. Szust (Oszust is Polish for ‘fraud’), who was equipped with a bogus and highly unqualified CV. They offered her services to 360 journals, 120 with an official impact factor, 120 appearing in the Directory of Open Access Journals (DOAJ), and 120 on Beall’s list of predatory journals. While none of the journals with an impact factor considered the application, 40 of the pseudo journals and eight DOAJ journals duly appointed her to their editorial board (Pisanski et al., 2017). Moreover, the lack of an impact factor does not seem to faze these journals, since they can now readily purchase an impact factor from fake JIF-certifying companies (Jalalian, 2015).

In this issue, Manuel Bagues, Mauro Sylos Labini, and Natalia Zinovyeva report a survey of Italian academicians who have published in journals appearing on Beall's list of predatory journals (some of these have since managed to become included in citation indexes such as Scopus and are therefore perhaps better characterized as 'questionable journals'). They find that these journals are often viewed positively by evaluation committee members who lack research expertise in the particular area, indicating that there are often critical information asymmetries in personnel evaluations. This creates an environment in which there is may be significant rewards for publishing in questionable or even pseudo journals.

Along with pseudo journals, another phenomenon gaining attention is the pseudo conference (Cress, 2017; Sorooshian, 2017). These are evidently proliferating to judge from the number of invitations appearing in academic inboxes, reflecting the ease of purchasing email lists of academicians and then soliciting their attendance. A recent exposé by Korean and German news organizations examined these conferences in detail. The journalists showed that a bogus paper generated, in one second, by the automated paper-generator SCIgen was accepted, and then presented, at a World Academy of Science, Engineering and Technology (WASET) organized conference in Venice.[13] At that conference, there were numerous other presenters, giving more or less serious papers on a hodgepodge of topics. The Korean journalists presented their paper with a straight face, and no one seemed any the wiser. The German journalists submitted a nonsense paper to another WASET conference, and even won a 'best-presenter' award for their efforts! Many of the presenters were students or post-docs, young scholars who admitted to a need to demonstrate their productivity and ability to present at international conferences. Other presenters noted the need to spend down their fiscal year budgets through

[13] https://news.kcij.org/67 (accessed 1 October 2018).

attending a conference. In addition, for some, publishing in WASET journals counted as a scientific accomplishment in their annual evaluations. Moreover, WASET is by no means the only such organization.

In short, we are observing very low quality or even fake research being presented at pseudo conferences and published in pseudo journals. Yet given the poor attendance and limited feedback on presentations at many 'legitimate' conferences, WASET and other similar organizations might be seen as trying to tap into some of the legitimacy of these conferences, making it more difficult to clearly draw the line between pseudo and legitimate avenues for disseminating scientific findings (Negro et al., 2014). While one might denounce such conference organizers as fraudsters taking advantage of naïve scholars, it might be more reasonable to view them as fulfilling a market niche for satisfying an increasingly misplaced goal of demonstrating one's productivity, engaging in an elaborate performance enacting a mere simulacrum of science (Baudrillard, 1988). Lastly, while it is difficult to resist the comic effect of exposing the desperately low standards of these conferences and journals by submitting computer-generated spoof articles to them, it should be noted that numerous bogus articles produced by the SCIgen software – more than 120 according to one report – have been accepted and published by top publishers like Springer and the Institute of Electrical and Electronic Engineers (IEEE) (Van Noorden, 2014).

In summary, there is a great variety of questionable practices occurring in the publication process, some affecting the research results themselves (dishonest conformity, unethical reviewing), others seemingly designed primarily to provide venues for those who are largely excluded from more reputable outlets. And while it may be convenient to cast pseudo journals as the main culprits, the problems are much more widespread and insidious, affecting the top journals as well as lower-tier ones. There is also no evidence that research based on detrimental research practices is especially targeted by more dubious outlets, but because there is no serious review process, the amount of flawed research published in these journals is likely to be far higher. Hence, it might be better to think of these as parallel developments that are helping to undermine the credibility of the research enterprise globally.

5. Environmental Context for Misconduct

Over its long history, the nature and types MMG have evolved, reflecting changing incentives, assessment regimes and technologies. In this section, we examine certain contextual issues that may affect rates of misconduct of various forms, including scientific competition and the impingement of increased economic interest in the outcomes of university research (see Tourish and Craig, 2018; Hall and Martin, this issue).

5.1. Scientific competition: Do 'Hot' Fields Attract More Outright Fraud?

A longstanding concern in the study of misconduct is that it is driven by the pressures of scientific competition (Merton, 1973). While Merton and Zuckerman (1971) argued that competition came paired with aggressive review and criticism, thus providing science with a strong self-correcting mechanism, unfortunately this has often proved not to be the case. Peer review is supposed to render science self-policing, but there is evidence that external forces such as strong public interest or government pressure may encourage more unethical scientific behavior. A case in point has been the enormous public interest in stem cells and regenerative medicine that led to large increases in funding and new programs intended to accelerate the introduction of cures (Kenney and Patton, 2018). Technological hype has often coalesced with national scientific pride, with countries investing in stem cell research as a way to bolster the international visibility of their science. In a number of cases, this has led to fraud. For instance, the 2006 *Science* article by Seoul National University researcher, Hwang Woo Suk, who claimed to have created human embryonic stem cells by cloning, was later exposed as fraudulent, destroying his career (Kakuk, 2009). Not only were several experiments not reproducible, but Hwang had also apparently obtained human oocytes without proper authorization. In another case, in 2014 Haruko Obokata, then of the Japanese RIKEN research institute, published two papers in *Nature* claiming her team had developed a way to convert any mammalian adult cell into a stem cell. Although these were relatively quickly debunked and Obokata was dismissed, episodes like this raise the question of whether 'hot' scientific fields are more attractive for fraud, or whether such stories simply attract more media attention (Lancaster, 2016).

In this special issue, Christian Berggren and Filiz Karabag describe how the Swedish Karolinska Institute hired a famous surgeon, Paolo Macchiarini, to enhance its standing. The Institute subsequently found it difficult to investigate Macchiarini, who was revealed to be transplanting unsanctioned and ultimately harmful artificial tracheas into his patients (see also Karabag and Berggren, 2016). Various actors tried unsuccessfully to expose the fraud before a Swedish TV documentary ultimately succeeded. The case illustrates the problem of fragmented control in the academic arena and how the boundaries between the academic world and wider society may shift during the actions related to MMG. It also shows that sometime the identification and exposure of misconduct does not come from academic organizations or governmental agencies but from elsewhere – in this case a documentary-maker.

5.2. Economic Interests and MMG

Some academic research has always had economic value, whether this be conducting clinical trials for the drug industry or creating software that can be commercialized. Increasing pressure from granting agencies and the emphasis many universities place on undertaking commercially valuable research have created incentives for re-orienting one's research agenda, and for 'spinning' research results in ways favorable to the research sponsor's interest (Sismondo, 2008) or to firms in which the professor has an economic interest (Slaughter et al., 2004). A meta-analysis of biomedical research studies found that industry-funded studies are more likely to find positive results (Bekelman et al., 2003). Industry-funded research is also more likely to engage in selective reporting (Salandra, 2018) and to make data unavailable for public review. None of this requires falsifying data, but merely taking advantage of the ‘researcher degrees of freedom’ to make design decisions likely to produce the desired result – what Biddle (2018) terms high ‘epistemic risk’.

Biomedical ghostwriting (Sismondo, 2009) could be put into this category. This involves academics who agree to be listed as authors on articles designed, paid for, and contracted to writing companies by pharmaceutical firms. This practice may appeal to some academics because of the generous honoraria offered to such ‘authors’, but may also be driven by a desire to appear more productive, since this can lead to more publications, citations and grants, a higher salary, and increased visibility (Sismondo, 2019). Hence, this behavior may reflect both the economic incentives of the study’s sponsors and the metrics-based incentives to academics. While such behavior does not necessarily change a study’s results, it presumably does enhance the impact of the study by associating it with more prestigious authors. Furthermore, even when there is no academic misconduct involved, even the perception of dubious conduct may be detrimental, particularly as the debate about the products of science becomes more politicized.

6. Detection, Enforcement, and Remedies

6.1. Detection and enforcement

While there have been studies suggesting that the academic community is relatively quick to recover once academic fraud has been exposed (e.g. Furman et al., 2012), the costs in terms of editorial and administrative time are often overlooked. In one case involving fraudulent biomedical publications, the total cost of the investigation was over half a million dollars (Michalek et al., 2010). The traditional forms of detection and enforcement for misconduct depend mostly upon the efforts of the scholarly community in the form, primarily, of editors and reviewers. Unfortunately, the rapid growth of the academic community, both in numbers and geographical scope, has almost certainly decreased the influence of community pressures mitigating wrongdoing and diluted agreement overt behavioral norms (Honig et al., 2014). At the same time, increasing specialization within fields suggests that editors may no longer be as intimately acquainted with their burgeoning field and hence less well equipped to ensure that appropriate referees are chosen.

The damage caused by misconduct is difficult to judge, partly because of the variety of the constituencies affected, and the network effects of the damage. In a study of retractions, Furman et al. (2012) finds that retracted papers fairly quickly lose their citations.[14] Salandra et al. (2018) show that papers labelled as being suspected of biased reporting also lose some of their expected citations. Hence, from an internalist perspective the damage to ongoing science may not be that severe, unless the retracted publications involved medical and therapeutic claims that have been adopted prior to the retraction. This can often happen as readers who downloaded an article prior to retraction may never hear that it was subsequently retracted. Yet while the damage to science as a body of knowledge may be limited, it is virtually impossible to assess the damage that fraudulent publications inflict on those honest scientists who failed to secure the jobs, promotions, or grants that instead went to dishonest scientists who looked good because of their fraudulently inflated vitae.

For the scientist forced to make the retraction, there is also a significant citation penalty with regards to their non-retracted articles, as these are subsequently less cited (Lu et al., 2013). The scale of the impact upon researchers whose misconduct or misrepresentation is discovered obviously varies as does the severity of sanctions, which can include a multi-year ban from government funding that may slow a scientist's career, especially in the case of junior researchers, and in some cases, result in the loss of employment. However, for much system-gaming behavior such as excessive self-citation or even more dubious actions such as agreeing to be a guest author, there are few if any sanctions unless there are also questions regarding the research itself (Teixeira da Silva and Dobranski, 2015). Of course, if it does prove necessary to retract an article, then this has a negative impact on the author's prior and future work (Lu et al.,

[14] Although there is a significant drop in citations to retracted papers, generally the drop is not to zero (Pfeifer and Snodgrass, 1990).

2013, Azoulay et al., 2015; 2017). In this Special Issue, Katrin Hussinger and Maikel Pellens show that the punishment extends beyond the authors, with co-authors on non-retracted articles also suffering an 8 to 9% citation penalty after the retraction. They explain this as "stigmatization by association", which may then evolve into a type of "shunning" behavior. They also suggest that this creates an indirect cost in the form of mistrust, perhaps discouraging whistleblowers from raising concerns in public – a result that resonates with the literature on the costs that whistleblowers face (Alford, 2002). The broader impacts of research malfeasance have begun to receive some attention (Hall and Martin, this issue), but more research is needed.

With misconduct and misrepresentation apparently on the increase, much attention has focused on how these might be curtailed. Governments have established offices such as the U.S. National Institutes of Health Office of Research Integrity tasked with investigating research malpractice. However, a 2008 survey of research malpractice concluded it was dramatically under-reported (Titus et al., 2008) – perhaps not surprising given that those reporting research malpractice may suffer not only condemnation from fellow researchers, but are also likely to have to invest significant time and resources in any investigation. Furthermore, in the U.S., they run the risk of being sued for libel or defamation by the individuals whose wrong-doing they expose (Edwards and Roy, 2017). Programs to reward those exposing scientific fraud, such as the US False Claims Act (which allows whistleblowers to recover a significant share of the funds recovered by the government), may create incentives to take on these burdens, although we do not yet have reliable evidence whether such incentives work with academic misconduct. However, they could also lead to unfounded attacks, especially on research critical of particular interests such as pharmaceutical or fossil-fuel firms.

More recently, concern about academic misconduct has attracted not only scholarly research but also the attention of self-appointed watchdogs. The widely followed Retraction Watch, often used as a source by academic misconduct researchers, was established by Ivan Oransky and Adam Marcus, two science writers (Guaspare and Didier, 2018). Retraction Watch is not affiliated with any government agency or academic organization, but is supported by philanthropic organizations like the Arnold Foundation, and by the donations from users. Similarly, Macchiarini's fraud was publicly exposed by a documentary filmmaker not a government agency (Berggren and Karabag, this volume), while the most influential critic of pseudo journals was Jeffrey Beall, a university-employed librarian. External non-governmental organizations and individuals have thus become part of an academic fraud detection 'ecosystem'. Prominent within this new environment is PubPeer, a public website that started out as an 'online journal club' but has quickly become the premier fraud-detection website based on a crowd-sourcing model (Guaspare and Didier, 2018). Participants post their anonymous concerns about a given publication, inviting both the authors and other readers to further clarify and analyze the questions. Unlike the Office of Research Integrity where just a few experts investigate misconduct claims, the PubPeer model can potentially draw upon the expertise of hundreds of volunteer specialists. (The analogies with the open-source software model are obvious). And while PubPeer has no legal authority over the scientists whose work it analyzes, its discussions and findings have the ability to go viral, creating a level of global publicity that universities and government agencies cannot ignore.

The PubPeer model helps to overcome key hurdles. The difficulty in responding to research malpractice is that investigation is, almost invariably, far more expensive than committing the fraud. The fraud must be proved following a quasi-legal process, and that proof generally requires establishing intent – a hurdle that involves reconstructing the sequence by which the fraud was committed. Unfortunately, those most familiar with the data, decisions, and techniques may have little interest in cooperating and committing time, given this will do little to further their career. That does not apply, however, to the many part-time volunteer analysts contributing their expertise to PubPeer. Conversely, recommendations such as those by Gall and Maniadis (this issue) to address research misconduct through greater scrutiny to raise the cost of questionable research practices may prove difficult to implement, since yet more resources and time would need to be devoted to investigating and punishing those involved.

In this Special Issue, Thomas Gall and Zacharias Maniadis create a stylized game-theoretic model of the research and publication process to explain scientific misconduct and how it might be detected and prevented (see also Lacetera and Zirulia, 2011). Gall and Maniadis begin by observing that there are differing levels of severity of research misconduct. Moreover, the types of fraud detected are not necessarily the most prevalent ones and may not even provide a reliable picture of the range of fraud being perpetrated. Their model suggests that the research enterprise will secure the greatest benefit by mandating research transparency to prevent the more widespread and milder forms of questionable research practices. They argue further that the pressure to publish might even reduce misconduct by encouraging closer scrutiny, in line with what Merton and Zuckerman (1971) previously suggested.

One striking feature of the literature on scientific misconduct and attempts to combat it is the continuing dominance of an individualistic perspective on the problem, which is surprising given the team nature of much scientific research (see, for example, Funk, 2017). Walsh et al. (this issue) and Hall and Martin (also this issue) review the literature on misconduct in organizations, arguing that a more organizational-based understanding of pathologies in academic research may help in identifying potential solutions. This is reinforced by the findings of a recent study by Davis (2018), who found that systemic competitive pressures lay behind many of the most common forms of detrimental research practices. Consequently, "Discussion of research integrity that locates it solely in the behaviours of individuals, and makes no effort to incorporate or reflect on wider injustices in the system of science, runs the risk of being ignored by the very researchers it is directed at." (ibid., p.17)

6.2. Remedies

While there are various voices raising concerns about MMG, there is little agreement on possible solutions. Much of the current focus is on specifying guidelines, improving the culture and providing more extensive training in the hope that all this will lead to researchers behaving more ethically (Funk, 2017). Similarly, the problems associated with p-hacking and HARKing and other misuses of researcher degrees of freedom may lessen as new norms and practices regarding the uses of statistical testing and the need for independent testing of hypotheses derived from existing datasets begin becoming standard in graduate school training. Although these might be useful in reducing the variation in the norms noted above, one doubts whether it will be effective either for serious offenders such as Gottinger, Schön or Stapel, nor in addressing the 'researcher degrees of freedom' problem. Furthermore, even if norms of proper behavior and judgment may be stabilized within a discipline, interdisciplinary research will continue to bring together very different fields, thus creating a source of potentially ongoing confusion. Lastly, as many examples cited here show, misconduct is very much a moving target, spawning new practices in response to new detection processes[15] as well as changing incentive systems such as metrics (Biagioli and Lippman 2019) and the availability of new tools like Photoshop which facilitate fraudulent manipulations of visual evidence. Finding remedies for misconduct thus requires constant attention to, and research on, the new strains that keep evolving.

Some observers such as Nosek et al. (2015) argue that the current incentive system in science may over-emphasize innovation to the detriment of verification. In the case of collaborative research, Chubin (1985) and Walsh et al. (this issue) argue for applying lessons from high-reliability systems (such as those for airplanes or nuclear plants) by incorporating more overlapping duties and quality checks into the system, even in the lab, before publication. Regarding the responsibilities of co-authors, the committee investigating the Schön case argued that co-authors have a clear responsibility for ensuring that papers with their names attached are based on sound results (Beasley et al., 2002), in line with the ICMJE definition of authorship.

[15] A 'Red Queen' evolutionary effect (Van Velen, 1973) is at work here, with improved efforts to detect and police misconduct then encouraging devious researchers and organizations to find new ways of beating the system.

This has led Nosek et al. (2015) to propose what they call the Transparency and Openness Promotion (TOP) guidelines. These include: (1) standards for citations to data, code, and materials; (2) standards for data sharing; (3) code sharing; (4) research materials sharing; (5) design and analysis transparency; (6) preregistration of studies; (7) preregistration of analysis plans; and (8) encouraging replications. These standards can be adopted by journals in order to encourage reporting of studies in ways that increase the transparency and reliability of findings. Similarly, Bergh et al. (2017) emphasize the importance of reproducibility as a first check on the validity of scientific claims. In addition, a recent editorial in the *Strategic Management Journal* suggests that, in order to improve the evaluation of scientific findings and the cumulativeness of science, we should move away from authors, reviewers and editors fetishizing alpha cutoffs – such that .049 is 'valuable' and .051 is 'worthless'– emphasizing instead the difference in importance attached to findings (Bettis et al., 2015). The expectation is that by reporting exact p values rather than asterisks or p-value cutoffs, misrepresentation will be discouraged and research findings will be more nuanced.

Similarly, Simmons et al. (2011) suggest a set of reporting standards designed not so much to eliminate researcher degrees of freedom as to increase the transparency of the research report. Their suggestions include: authors reporting their predetermined rules for terminating data collection; collecting sufficient cases to ensure minimum power to the tests; listing all variables collected in the study; reporting all experimental conditions, including failed manipulations; reporting statistical results if eliminated cases (e.g., outliers) are included; and reporting results with and without a covariate. They also suggest that reviewers should: ensure that authors follow these requirements, while at the same time being more tolerant of imperfections in the results; insist that authors show that results do not depend on specific design and analysis decisions; and if the justifications are not sufficient, require authors to conduct an exact replication. Again, the goal is to have journals adopt these practices in order to increase the transparency, and perhaps the robustness, of published research findings. At the same time, it is critical that tougher reporting standards still allow space for serendipity in the research.

Some suggest addressing these problems by de-emphasizing broad performance metrics and focusing just on major contributions (Bedeian et al., 2010). As Simmons et al. (2011, p.1365) argue, “Our goal as scientists is not to publish as many articles as we can, but to discover and disseminate truth”. Brown’s emphasis on integrity in the face of methodological contingencies highlights the need to be thoughtful in exercising one’s degrees of freedom, and to incorporate shared values into the choices, as well as showing humility and contextual awareness in the research and reporting process (Brown, forthcoming). Bedeian et al. (2010) suggest basing hiring, promotion and other rewards on just a few key publications (a practice introduced decades ago at Harvard Medical School) to discourage the production of numerous trivial papers in lower-tier journals. This would also have the advantage of helping to unclog journals and reduce the burden on reviewers, and give scientists more time to thoroughly develop their key ideas, rather than racing to publish yet another paper before the next review. In addition, it would make it easier for evaluators to read the smaller number of papers submitted, rather than pushing them to rely on summary metrics because of the overwhelming volume of materials produced.[16]

China has recently proposed reforms aimed at increasing the penalties for misconduct and publicly shaming those who engage in such behavior. This includes blacklisting certain very low quality journals to prevent them from being included in researchers’ productivity statistics, and taking responsibility for enforcement away from universities and journals and centralizing it in the ministries (Cyranoski 2018). While some might view these measures as too draconian and unlikely to prove effective, such external policing of science may become more common if key actors conclude that self-monitoring has failed.

[16] For example, at some U.S. universities, only the five best papers (as selected by the candidate) are included in a tenure and promotion evaluation packet, although the packet still generally includes the full vita.

7. Conclusions and Areas for Future Research

Concern about academic misconduct has been increasing within the academic community and elsewhere. In preparing the Special Issue and this Introduction, we were struck by the breadth of worries about MMG, with review articles, anecdotes, and expressions of concern in a wide variety of disciplines. Yet there are far fewer research studies exploring MMG. Moreover, many of the most highly cited articles were written not by social scientists but by scientists in different fields motivated by unease regarding the state of the scientific community.

The stakes involved in MMG may well vary appreciably by field. For example, in literature studies the harm beyond that to the academic enterprise itself may be minimal. However, in fields such as biomedical sciences, real physical harm has been done either directly, as in the Macchiarini case discussed by Berggren and Karabag (this issue), or indirectly as when pharmaceutical firms ghost-write articles advocating approval for a particular drug. Other less immediately identifiable harms have come from extensive oil-firm funding of climate denial 'research' (Lewandowski et al., 2015) or suspect research on pesticides and herbicides.

This introduction and the article by Martin and Hall in this issue show that MMG now extends far beyond that by an individual bad actor or research team to include journal editors, publishers, and even academic administrators. Journal editors have been accused of pressuring authors to cite their journal even when the citations may not be relevant. While not discussed in this overview, recently some have attacked journal publishers for excessive charges and even the creation of their own 'low-barrier' journals in an effort to compete with other 'mega journals'.[17] Finally, universities worldwide increasingly see themselves locked into competition with each other, with the 'currency' being global rankings and impact factors (Kehm, 2019). This is strengthening the pressure on researchers to publish in top-tier journals by any means necessary – or, if that proves impossible, at least to publish more papers, even if that involves more

[17] Many of these journals cover the entire range of science rather than just one field.

questionable journals – simply to demonstrate to administrators, funding authorities and others that they are ‘productive’. In the U.S., such pressures are having a pernicious impact on the entire academic system (Breznitz and Kenney, 2018), while in other countries researchers lacking the resources to compete at higher levels may have little option but to resort to pseudo journals and conferences.

We would be remiss not to revisit the impact of technology, especially digitalization and computers, on MMG. Contemporary natural sciences rely heavily on visual evidence, but the advent of software that can alter images makes forgery easier and quicker (Biagioli, 2019). The impacts of digital technologies on scientific work are only beginning to receive attention, so not surprisingly far less is known about how these may increase the rate of fraud. In the social sciences, the presence of powerful, simple-to-use software has allowed social scientists to data-mine and to hypothesize after the fact (Kerr, 1998) in order to develop ‘theories’ that seemingly fit the data better. Pseudo journals and conferences would not be nearly as successful without the ability to create websites and publish online at a fraction of the cost of printing a journal issue or of mailing conference materials to potential participants.

For the *Research Policy* community, MMG and its impacts on the scientific enterprise would seem worthy of increased research, including the development of more sophisticated measurement and analysis. Pseudo journals and conferences deserve greater attention and even-handed assessments of whether they contribute to the overall expansion of scientific research in nations with poorer or emerging research systems. Have research assessment exercises and the increasing emphasis on metrics and ranking systems contributed to a deterioration in research? Can we study instances regarding organizational MMG, whether by universities, editors, or publishers, to understand and correct for their attempts to game the system? Better understanding of how outside interests and excessive hype affects behavior in research fields such as climate change and biotechnology would be of particular interest to *Research Policy* readers.

Across the natural, biological and social sciences, acknowledgement of the difficulty of reproducing research results in published articles is now widespread (Baker, 2016). Given the current incentive structure not only for potential miscreants, but also for the scientific community that would like to enforce standards, the ability to discipline deviants is constrained by the considerable costs imposed on those charged with maintaining standards, although assistance from volunteers like those contributing to PubPeer may help to spread the load. Discipline is further constrained by the fuzzy boundaries regarding what behaviors are within acceptable bounds and which are clearly over the line (Hall and Martin, this issue). Yet, while there may be a lack of clarity on the boundary regions, this cannot explain the considerable rise in egregious violations of research norms, some of which border on the criminal.

For the scientific community, there is at least some good news. According to a Pew survey, the U.S. public's trust in science, although quite "soft", is still relatively high (Funk, 2017), and that is likely to be true in most countries. This places the scientific community in a difficult position, since greater efforts to expose and discipline research malefactors could lead to publicity that erodes public trust. Conversely, a lack of action could result in embarrassing scandals exposed in the media (as in the Korean-German exposé of pseudo conferences discussed above). A public debate within the scientific community is now vital or the enterprise runs the risk of external supervision being imposed that could erode the relative independence of the community. Funding for science could also suffer. It is therefore incumbent on the scientific community to set about reducing the incidence of MMG, clearly demonstrating that we are capable of successfully policing ourselves.

Acknowledgements

The author order is alphabetical, with all the authors contributing equally to the Special Issue Introduction. This Special Issue was originally conceived by Mario Biagioli and Martin Kenney. However, due to unforeseen circumstances, Ben Martin and John Walsh later joined the editorial team after all the SI papers had been accepted. The authors then completed the introduction collectively. Walsh was funded by the US National Science Foundation (Grant#164645).

References

Abritis, A. & McCook, A., 2017. Cash bonuses for peer-reviewed papers go global. Science. (Downloaded on 1 October 2018 from http://www.sciencemag.org/news/2017/08/cash-bonuses-peer-reviewed-papers-go-global)

Alford, C.F., 2002. Whistleblowers: Broken Lives and Organizational Power. Cornell University Press, Ithaca, N.Y.

Azoulay, P., Bonatti, A. & Krieger, J.L., 2017. The career effects of scandal: evidence from scientific retractions. Research Policy, 46(9), 1552-1569. https://doi.org/10.1016/j.respol.2017.07.003

Azoulay, P., Furman, J.L., Krieger, J.L. & Murray, F., 2015. Retractions. Review of Economics and Statistics, 97(5), 1118-1136. https://doi.org/10.1162/REST_a_00469

Baker, M., 2016. Is there a reproducibility crisis? Nature, 533, 452-454. doi:10.1038/533452a

Banks, G.C., O'Boyle E.H., Pollack, J.M., White, C.D., Batchelor, J.H., Whelpley, C.E., Abston, K.A., Bennett, A.A. & Adkins, C.L., 2015. Questions about questionable research practices in the field of management: a guest commentary. Journal of Management, 42(1), 5-20. doi:10.1177/0149206315619011.

Bartlett, Tom. 2018. "'I Want to Burn Things to the Ground': Are the foot soldiers behind psychology's replication crisis saving science — or destroying it?" The Chronicle of Higher Education (September 11, 2018)

Basken, P., 2017. Why Beall's list died – and what it left unresolved about open access. Chronicle of Higher Education. (Downloaded on 1 October 2018 from https://www.chronicle.com/article/Why-Beall-s-List-Died-/241171)

Baudrillard, J., 1988. Simulacra and Simulations, in: Poster, M. (ed.), *Jean Baudrillard, Selected Writings*. Stanford University Press, Stanford, CA, pp. 166-184.

Beall, J., 2012. Criteria for determining predatory open-access publishers. 2nd edition. Scholarly Open Access, Denver, CO. http://scholarlyoa.files.wordpress.com/2012/11/criteria-2012-2.pdf.

Beasley, M. R., Datta, S., Kogelnik, H., Kroemer, H. & Monroe, D., 2002. Report of the investigation committee on the possibility of scientific misconduct in the work of Hendrik Schön and coauthors. Lucent Technologies. (Downloaded on 1 October 2018 from https://media-bell-labs-com.s3.amazonaws.com/pages/20170403_1709/misconduct-revew-report-lucent.pdf)

Bedeian, A.G., Taylor S.G. & Miller. A.N., 2010. Management science on the credibility bubble: cardinal sins and various misdemeanors. Academy of Management Learning & Education, 9(4), 715-25. doi: 10.5465/amle.9.4.zqr715.

Bekelman, J.E., Li, Y. & Gross, P.C., 2003. Scope and impact of financial conflicts of interest in biomedical research. Journal of the American Medical Association, 289, 454-65. doi:10.1001/jama.289.4.454

Bergh, D.D., Sharp, B.M., Aguinis, H. & Li, M., 2017. Is there a credibility crisis in strategic management research? Evidence on the reproducibility of study findings. Strategic Organization, 15(3), 423-36. doi: 10.1177/1476127017701076.

Bettis, R.A., Ethiraj, S., Gambardella, A., Helfat, C. & Mitchell, W., 2015. Creating repeatable cumulative knowledge in strategic management. Strategic Management Journal, 37(2), 257-61. doi: 10.1002/smj.2477.

Bhattacharjee, Y., 2011. Saudi universities offer cash in exchange for academic prestige. Science, 334(6061), 1344-1345. doi: 10.1126/science.334.6061.1344.

Biagioli, M., 2003. Rights or rewards? Changing frameworks of scientific authorship, in: Biagioli, M. & Galison, P. (eds), Scientific Authorship: Credit and Intellectual Property in Science. Routledge, New York, pp.253-279.

Biagioli, M., 2012. Recycling texts or stealing time? Plagiarism, authorship, and credit in science. International Journal of Cultural Property, 19, 453-476. https://doi.org/10.1017/S0940739112000276

Biagioli, M., 2016. Watch out for cheats in citation game. Nature, 535, 201. doi: 10.1038/535201a.

Biagioli, M., 2019. From Artisanal to Industrial Fraud. Angewandte Chemie, 58 (in press).

Biagioli, M. & Lippman, A. (eds.), 2019. Metrics and Misconduct: New Ecologies of Academic Research. MIT Press. (In press)

Biddle, J.B., 2018. Antiscience zealotry? Values, epistemic risk, and the GMO debate. Philosophy of Science 85, 360-79. https://doi.org/10.1086/697749

Bohannon, J., 2013. Who's afraid of peer review? Science 342, 60-65. doi: 10.1126/science.342.6154.60

Boutron, I. & Ravaud, P., 2018. Misrepresentation and distortion of research in biomedical literature. Proceedings of the National Academy of Sciences, 115(11), 2613-2619. https://doi.org/10.1073/pnas.1710755115

Breznitz, S. & Kenney, M., 2018. Slouching toward the Downton Abbey university system. Issues in Science and Technology, 34(3), 74-82.

Broad, W. & Wade, N., 1983. Betrayers of Truth. Simon & Schuster, New York.

Brown, M.J., forthcoming. Is science really value free and objective? From objectivity to scientific integrity, in: McCain, K and Kampourakis, K. (eds),What Is Scientific Knowledge? An Introduction to Contemporary Epistemology of Science. Routledge, New York.

Buranyi, S., 2017. Is the staggeringly profitable business of scientific publishing bad for science? The Guardian. (Downloaded on 27 June 2018 from https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science)

Butler, N., Delaney, H. & Spoelstra, S., 2017. The gray zone: questionable research practices in the business school. Academy of Management Learning & Education, 16(1), 94-109. doi: 10.5465/amle.2015.0201.

Cambridge Economic Policy Associates., 2008. Activities, costs and funding flows in the scholarly communications system in the UK. (Downloaded on 27 June 2018 from www.rin.ac.uk/system/files/attachments/Activites-costs-flows-report.pdf)

Chubin, D., 1985. Beyond invisible colleges: inspirations and aspirations of post-1972 social studies of science. Scientometrics, 7(3-6), 221-254. https://doi.org/10.1007/BF02017148

Colussi, T., 2018. Social ties in academia: a friend is a treasure. Review of Economics and Statistics, 100(1), 45-50. https://doi.org/10.1162/REST_a_00666

Conner-Simons, A., 2015. How three MIT students fooled the world of scientific journals. MIT News. (Downloaded on 27 June 2018 from http://news.mit.edu/2015/how-three-mit-students-fooled-scientific-journals-0414)

Cress, P.E., 2017. Are predatory conferences the dark side of the open access movement? Aesthetic Surgery Journal, 37(6), 734–738. https://doi.org/10.1093/asj/sjw247

Csiszar, A., 2019. Gaming metrics before the game: citation and the bureaucratic virtuoso, in: Biagioli, M. & Lippman, A. (eds.), Metrics and Misconduct: New Ecologies of Academic Research. MIT Press. (In press)

Cyranoski, D., 2018. China introduces sweeping reforms to crack down on academic misconduct. Nature, 558,171. doi: 10.1038/d41586-018-05359-8.

Davies, S.R., 2018. An ethics of the system: talking to scientists about research integrity. Science and Engineering Ethics (forthcoming). https://doi.org/10.1007/s11948-018-0064-y

Davis, P., 2014. When a journal sinks, should the editors go down with the ship? The Scholarly Kitchen (6 October) (downloaded on 5 October 2018 from https://scholarlykitchen.sspnet.org/2014/10/06/when-a-journal-sinks-should-the-editors-go-down-with-the-ship/)

De Silva, P.U. & Vance, C.K., 2017. Preserving the quality of scientific research: peer review of research articles, in Scientific Scholarly Communication: The Changing Landscape. Fascinating Life Sciences. Springer, Cham, pp. 73-99. https://doi.org/10.1007/978-3-319-50627-2_6

Edwards, M.A. & Roy, S., 2017. Academic research in the 21st century: maintaining scientific integrity in a climate of perverse incentives and hyper competition. Environmental Engineering Science, 34(8), 616-616. https://doi.org/10.1089/ees.2016.0223

Ellis, T., 2013. How Tulane's B-school is moving past scandal. Fortune. (Downloaded on 27 June 2018 from http://fortune.com/2013/07/29/how-tulanes-b-school-is-moving-past-scandal/)

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, 109(42), 17028-17033. https://doi.org/10.1073/pnas.1212247109

Ferguson, C., Marcus, A. & Oransky, I., 2014. Publishing: the peer-review scam. Nature. (Downloaded on 27 June 2018 from https://www.nature.com/news/publishing-the-peer-review-scam-1.16400)

Flanagin, A., Carey, L.A., Fontanarosa, P.B., Phillips, S.G., Pace, B.P., Lundberg, G.D. & Rennie, D.,1998. Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. Journal of the American Medical Association, 280(3), 222-224. doi:10.1001/jama.280.3.222

Fong, E.A. & Wilhite, A.W., 2017. Authorship and citation manipulation in academic research. PloS One, 12(12), e0187394. https://doi.org/10.1371/journal.pone.0187394

Frandsen, T.F., 2017. Are predatory journals undermining the credibility of science? A bibliometric analysis of citers. Scientometrics, 113(3), 1513-1528. https://doi.org/10.1007/s11192-017-2520-x

Franzoni, C., Scellato, G. & Stephan, P., 2011. Changing incentives to publish. Science 333(6043), 702-03. doi: 10.1126/science.1197286

Frey, B.S., 2003. Publishing as prostitution? – Choosing between one's own ideas and academic success. Public Choice, 116(1), 205-23. doi: 10.1023/A:1024208701874.

Funk, G., 2017. Mixed messages about public trust in science. Pew Foundation. (Downloaded on 17 September 2018 from http://www.pewinternet.org/2017/12/08/mixed-messages-about-public-trust-in-science/)

Furman, J.L., Jensen, K. & Murray, F., 2012. Governing knowledge in the scientific community: exploring the role of retractions in biomedicine. Research Policy, 41(2), 276-290. https://doi.org/10.1016/j.respol.2011.11.001

Goodhart, C.A., 1984. Problems of monetary management: the UK experience, in Monetary Theory and Practice. Palgrave, London, pp.91-121.

Guaspare, C. & Didier, E., 2018. The new watchdogs' vision of science: A roundtable with Ivan Oransky (Retraction Watch) and Brandon Stell (PubPeer). Social Studies of Science, 48(1),165-167. doi: 10.1177/0306312718756202.

Hackett, E.J., 1994. A social control perspective on scientific misconduct. Journal of Higher Education, 65(3), 242-60. doi: 10.1080/00221546.1994.11778499.

Haug, C.J., 2015. Peer-review fraud—hacking the scientific publication process. New England Journal of Medicine, 373(25), 2393-2395. doi: 10.1056/NEJMp1512330

Honig, B. & Akanksha, B., 2012. The fox in the hen house: a critical examination of plagiarism among members of the academy of management. Academy of Management Learning & Education, 11(1), 101-23. doi: 10.5465/amle.2010.0084.

Honig, B., Lampel, J., Siegel, D. & Drnevich, P., 2014. Ethics in the production and dissemination of management research: institutional failure or individual fallibility? Journal of Management Studies, 51(1), 118-42. doi: 10.1111/joms.12056.

Ioannidis. J., 2012. Why science is not necessarily self-correcting. Perspectives on Psychological Science, 7(6), 645-654. https://doi.org/10.1177%2F1745691612464056

Jabbehdari, S. & Walsh, J.P., 2017. Authorship norms and project structures in science. Science, Technology, & Human Values, 42(5), 872-900. doi: 10.1177/0162243917697192.

Jalalian, M., 2015. The story of fake impact factor companies and how we detected them. Electronic Physician, 7(2), 1069–1072. doi: 10.14661/2015.1069-1072

Kakuk, P., 2009. The legacy of the Hwang case: research misconduct in biosciences. Science and Engineering Ethics, 15(4), 545-562. https://doi.org/10.1007/s11948-009-9121-x

Karabag, S. & Berggren, C., 2016. Misconduct, marginality and editorial practices in management, business and economics journals, PLoS ONE, 11, e0159492, doi.org/10.1371/journal.pone.0159492.

Kehm, B., 2019. Global university rankings: impacts and applications, in: Biagioli, M. & Lippman, A. (eds.), Metrics and Misconduct: New Ecologies of Academic Research. MIT Press. (In press)

Kenney, M. & Patton, D., 2018. Sub-national technology policy and commerce: evaluating the impacts of the California Institute for Regenerative Medicine. Journal of Technology Transfer, 43(1), 47-68. doi: 10.1007/s10961-017-9580-1

Kerr, N.L., 1998. HARKing: hypothesizing after the results are known. Personality and Social Psychology Review, 2(3), 196-217. https://doi.org/10.1207%2Fs15327957pspr0203_4

Kevles, D., 2000. The Baltimore Case. WW Norton, New York.

Kupferschmidt, K., 2018. Tide of lies. Science, 361(6403), 636-641. doi: 10.1126/science.361.6403.636

Kutner, M., 2014. How to game the college rankings. Boston Magazine. (Downloaded on 27 June 2018 from https://www.bostonmagazine.com/news/2014/08/26/how-northeastern-gamed-the-college-rankings/)

Lacetera, N. & Zirulia, L., 2011. The economics of scientific misconduct. Journal of Law, Economics, and Organization, 27(3), 568-603. https://doi.org/10.1093/jleo/ewr025

Lancaster, C., 2016. The acid test for biological science: STAP cells, trust, and replication. Science and Engineering Ethics, 22(1), 147-167. doi: 10.1007/s11948-015-9628-2

Larkin, I. & Pierce, L., 2015. Compensation and employee misconduct: the inseparability of productive and counterproductive behavior in firms, in: Palmer, D. Greenwood, R. & Smith-Crowe, K. (eds.), Organizational Wrongdoing, Cambridge University Press, Cambridge, pp.1-27.

Lee, C. J., Sugimoto, C. R., Zhang, G. & Cronin, B., 2013. Bias in peer review. Journal of the American Society for Information Science and Technology, 64(1), 2-17. https://doi.org/10.1002/asi.22784

Lerner, A.C., 2013. Statement of Allison C. Lerner, Inspector General, National Science Foundation, Committee on Science, Space, and Technology Subcommittee on Investigations and Oversight, United States House of Representatives: Top challenges for science agencies: Reports from the Inspectors General. (Downloaded on 27 June 2018 from https://www.nsf.gov/oig/_pdf/testimony/Lerner_Testimony030917.pdf)

Lewandowsky, S., Oreskes, N., Risbey, J.S., Newell, B.R. & Smithson, M., 2015. Seepage: climate change denial and its effect on the scientific community. Global Environmental Change, 33, 1-13. https://doi.org/10.1016/j.gloenvcha.2015.02.013

Lewellyn, K.B., Judge, W.Q. & Smith, A., 2017. Exploring the questionable academic practice of conference paper double dipping. Academy of Management Learning & Education, 16(2), 217-36. https://doi.org/10.5465/amle.2015.0033

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.

Macfarlane B., Zhang J. & Pun A., 2014. Academic integrity: a review of the literature. Studies in Higher Education. 39(2), 339-58. https://doi.org/10.1080/03075079.2012.709495

Martin, B.R., 2012. Does peer review work as a self-policing mechanism in preventing misconduct: a case study of a serial plagiarist, in: Mayer¸ T. & Steneck, N. (eds.), Promoting Research Integrity in a Global Environment, World Scientific Publishing/Imperial College Press, London and Singapore, pp. 97-114

Martin, B.R., 2013. Whither research integrity? Plagiarism, self-plagiarism and coercive citation in an age of research assessment. Research Policy, 42(5), 1005-1014. https://doi.org/10.1016/j.respol.2013.03.011

Martin, B.R., 2016. Editors' JIF-boosting stratagems – which are legitimate and which not? Research Policy, 45 (1), 1-7. http://dx.doi.org/10.1016/j.respol.2015.09.001

Martin, B.R., 2016b. What's happening to our universities? Prometheus, 34(1), 7-24. https://doi.org/10.1080/08109028.2016.1222123

Martin, B.R. et al., 2007. Keeping plagiarism at bay – a salutary tale. Research Policy, 36(7), 905-911. https://doi.org/10.1016/j.respol.2007.08.001

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

McCook, A., 2017. One way to boost your uni's ranking: ask faculty to cite each other. Retraction Watch. (Downloaded on 27 September 2018 from https://retractionwatch.com/2017/08/22/one-way-boost-unis-ranking-ask-faculty-cite/)

McCook, A., 2018. A publisher just retracted ten papers whose peer review was 'engineered'. Retraction Watch. (Downloaded on 26 September 2018 from https://retractionwatch.com/2018/07/page/3/)

Merton, R.K., 1973. The Sociology of Science: Theoretical and Empirical Investigations. University of Chicago Press, Chicago.

Merton, R.K. & Zuckerman, H., 1971. Patterns of evaluation in science: institutionalization, structure and function of the referee system. Minerva, 9, 66-100. doi: 10.1007/BF01553188

Michalek, A.M., Hutson, A.D., Wicher, C.P. & Trump, D.L., 2010. The costs and underappreciated consequences of research misconduct: a case study. PLoS Medicine, 7(8), e1000318. https://doi.org/10.1371/journal.pmed.1000318

Mohliver, A., 2018. How misconduct spreads: auditors' role in the diffusion of stock-option backdating. Administrative Science Quarterly (forthcoming). https://doi.org/10.1177%2F0001839218763595

National Academies of Sciences, Engineering, and Medicine, 2017. Fostering Integrity in Research. National Academies Press, Washington, D.C.

Negro, G., Visentin, F. and Swaminathan, A., 2014. Resource partitioning and the organizational dynamics of 'fringe banking'. American Sociological Review, 79(4), 680-704. https://doi.org/10.1177%2F0003122414537644

Nosek, B.A. et al., 2015. Promoting an open research culture: author guidelines for journals to promote transparency, openness, and reproducibility. Science 348, 1422-25. doi: 10.1126/science.aab2374

Oakley, K.P. & Weiner, J.S., 1955. Piltdown man. American Scientist, 43(4), 573-583.

Oravec, J.A., 2017. The manipulation of scholarly rating and measurement systems: constructing excellence in an era of academic stardom. Teaching in Higher Education, 22(4), 423-436. https://doi.org/10.1080/13562517.2017.1301909

Pachter, L., 2014. To some a citation is worth $3 per year. (Downloaded on 27 June 2018 from https://liorpachter.wordpress.com/2014/10/31/to-some-a-citation-is-worth-3-per-year/)

Park, J., Howe, J.D. & Sholl, D.S., 2017. How reproducible are isotherm measurements in metal–organic frameworks?. Chemistry of Materials, 29(24), 10487-95. doi: 10.1021/acs.chemmater.7b04287.

Pfeifer, M.P. & Snodgrass, G.L., 1990. The continued use of retracted, invalid scientific literature. Journal of the American Medical Association, 263(10), 1420-23. doi:10.1001/jama.1990.03440100140020

Pisanski, K., Sorokowski, P. & Kulczycki, E., 2017. Predatory journals recruit fake editor. Nature, 543(7646), 481-483. https://www.nature.com/polopoly_fs/1.21662!/menu/main/topColumns/topLeftColumn/pdf/543481a.pdf

Priyadarshini, S., 2018. India targets universities in predatory-journal crackdown. Nature, 560, 537-538. https://www.nature.com/articles/d41586-018-06048-2.

Quan, W., Chen, B. & Shu, F., 2017. Publish or impoverish: an investigation of the monetary reward system of science in China (1999-2016). Aslib Journal of Information Management, 69(5), 486-502. doi: 10.1108/AJIM-01-2017-0014

Reich, E.S., 2009. Plastic Fantastic. Palgrave Macmillan, New York.

Ryan, K.A., Sanders, M.A., Wang, D.D. & Levine, A.D., 2010. Tracking the rise of stem cell tourism. Regenerative Medicine, 5(1), 27-33. doi: 10.2217/rme.09.70.

Saisana, M., d'Hombres, B. and Saltelli, A., 2011. Rickety numbers: volatility of university rankings and policy implications. Research Policy, 40(1), 165-177. https://doi.org/10.1016/j.respol.2010.09.003

Salandra, R., 2018. Knowledge dissemination in clinical trials: exploring influences of institutional support and type of innovation on selective reporting. Research Policy, 47(7), 1215-1228. https://doi.org/10.1016/j.respol.2018.04.005.

Salandra, R., Criscuolo, P. and Salter, A., 2018. Impact of bias (detection) on follow-on research: evidence from the medical literature. Paper presented at BRICK workshop in Bath, UK.

Shen, C. & Björk, B.C., 2015. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. BMC Medicine, 13(1), 230. https://doi.org/10.1186/s12916-015-0469-2

Shibayama, S. & Yasunori, B., 2015. Dishonest conformity in peer review. Prometheus, 33(3), 215-33. doi: 10.1080/08109028.2015.1114745.

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), 1359-66. doi: 10.1177/0956797611417632.

Sismondo, S., 2008. Pharmaceutical company funding and its consequences: a qualitative systematic review. Contemporary Clinical Trials, 29(2), 109-113. doi: 10.1016/j.cct.2007.08.001

Sismondo, S., 2009. Ghosts in the machine: publication planning in the medical sciences. Social Studies of Science, 39(2), 171-198. https://doi.org/10.1177%2F0306312708101047

Sismondo, S., 2019. Ghost-managing and gaming pharmaceutical knowledge, in: Biagioli, M. & Lippman, A. (eds.), Metrics and Misconduct: New Ecologies of Academic Research. MIT Press. (In press)

Slaughter, S., Slaughter, S.A. & Rhoades, G., 2004. Academic Capitalism and the New Economy: Markets, State, and Higher Education. Johns Hopkins University Press, Baltimore.

Sorooshian, S., 2017. Conference wolves in sheep's clothing. Science and Engineering Ethics, 23(6), 1805-1806. https://doi.org/10.1007/s11948-016-9788-8

Stell, B. & Barbour, B., 2019. Scientific assessment without metrics, in: Biagioli, M. & Lippman, A. (eds.), Metrics and Misconduct: New Ecologies of Academic Research. MIT Press. (In press).

Steneck, N.H., 1994. Research universities and scientific misconduct: history, policies, and the future. Journal of Higher Education, 65(3), 310-330. doi: 10.2307/2943970

Teixeira da Silva, J.A. & Dobranszki J., 2015. Multiple authorship in scientific manuscripts: ethical challenges, ghost and guest/gift authorship, and the cultural/disciplinary perspective. Science and Engineering Ethics, 22(5), 1457-1472. doi: 10.1007/s11948-015-9716-3

Titus, S.L., Wells, J.A. & Rhoades, L.J., 2008. Repairing research integrity. Nature, 453(7198), 980. doi: 10.1038/453980a.

Tort, A.B., Targino, Z.H. & Amaral, O.B., 2012. Rising publication delays inflate journal impact factors. PLoS One, 7(12), e53374. https://doi.org/10.1371/journal.pone.0053374

Tourish, D. & Craig, R., 2018. Research misconduct in business and management studies: causes, consequences and possible remedies. Journal of Management Inquiry (forthcoming). https://doi.org/10.1177%2F1056492618792621

Van Noorden, R., 2011. The trouble with retractions, Nature, 478, 26-28. doi: 10.1038/478026a.

Van Noorden, R. 2013. Brazilian citation scheme outed. Nature News (27 August). (Downloaded on 27 June from https://www.nature.com/news/brazilian-citation-scheme-outed-1.13604#/citations)

Van Noorden, R., 2014. Publishers withdraw more than 120 gibberish papers. Nature News (24 February). (Downloaded on 3 June 2018 from http://www.nature.com/news/publishers-withdraw-morethan-120-gibberish-papers-1.14763)

Van Valen, L., 1973. A new evolutionary law, Evolutionary Theory, 1, 1–30.

Wager, E., 2009. Recognition, reward and responsibility: why the authorship of scientific papers matters. Maturitas, 62(2), 109-12. doi: 10.1016/j.maturitas.2008.12.001.

Wang, M.Q., Yan, A.F. & Katz, R.V., 2018. Researcher requests for inappropriate analysis and reporting: a U.S. survey of consulting biostatisticians. Annals of Internal Medicine (forthcoming). doi:10.7326/M18-1230

Whitley, R., 2007. Changing governance of the public sciences, in: Whitley, R. and Gläser J. (eds), The Changing Governance of the Sciences: The Advent of Research Evaluation Systems. Springer, Netherlands, pp.1-25.

Wilhite, A.W. & Fong, E A., 2012. Coercive citation in academic publishing. Science, 335(6068), 542-543. doi: 10.1126/science.1212540

Wislar, J.S., Flanagin, A., Fontanarosa, P.B. & DeAngelis, C.D., 2011. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. BMJ. https://doi.org/10.1136/bmj.d6128

Zuckerman, H., 1988. The sociology of science, in Smelser, N.J. (ed), Handbook of Sociology, Sage, Newbury Park, CA, pp.511-574