

Hoover, Gary A.; Hopp, Christian

Working Paper

What Crisis? Taking Stock of Management Researchers' Experiences with and Views of Scholarly Misconduct

CESifo Working Paper, No. 6611

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Hoover, Gary A.; Hopp, Christian (2017) : What Crisis? Taking Stock of Management Researchers' Experiences with and Views of Scholarly Misconduct, CESifo Working Paper, No. 6611, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/171075>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

What Crisis? Taking Stock of Management Researchers' Experiences with and Views of Scholarly Misconduct

Gary A. Hoover, Christian Hopp

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl

www.cesifo-group.org/wp

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

What Crisis? Taking Stock of Management Researchers' Experiences with and Views of Scholarly Misconduct

Abstract

This research presents the results of a survey regarding scientific misconduct elicited from a sample of 1,215 management researchers. We find that misconduct (research that was either fabricated or falsified) is not encountered often by reviewers nor editors. Yet, there is a strong prevalence of misrepresentations (method inadequacy, omission or withholding of contradictory results, dropping of unsupported hypotheses). Despite these findings, respondents put a fair deal of trust in the replicability and robustness of findings being published. A sizeable majority of editors and authors eschew open data policies but sees value in replication studies to ensure credibility in empirical research.

JEL-Codes: K300, A110.

Keywords: scientific misconduct, data fabrication, data misrepresentation, ethics.

Gary A. Hoover
Department of Economics
University of Oklahoma
308 Cate Center Drive
USA - 73019 Norman OK
ghoover@ou.edu

Christian Hopp
Time Research Area
RWTH Aachen University
Kackertstr. 7
Germany - 52072 Aachen
Hopp@time.rwth-aachen.de

INTRODUCTION

Harley, Faems, and Corbett (2014: 1361) reflect on the “general unease in the scholarly community about academic misconduct”. As editors, they indicate the growing concerns at their journal (*Journal of Management Studies*) but also with misconduct discussed in (private) conversations with editors at other leading management journals. Similarly, the number of retractions related to fraudulent behavior is troubling with some disciplines reporting a fourfold increase (Steen, 2010; Azoulay, Bonatti, and Krieger, 2015).

The field of management has come under increasing scrutiny recently, in part due to the incidences of peer-reviewed journal retractions (Honig, Lampel, Siegel, & Drnevich, 2014; Karabag, & Berggren, 2012; Karabag, & Berggren, 2016). One prominent management scholar has accumulated sixteen retractions from major management journals (Retraction Watch, 2016), and another has racked up seven (Retraction Watch, 2016b). While some errors can be attributed to accident or incompetent research, a growing body of work asserts that retractions are the result of academic dishonesty by scholars facing what could be considered an increasingly competitive environment (Honig et al., 2014).

In Psychology, Pashler and Wagenmakers (2012) discuss the transition from embarrassment (individual instances of fraud, dubious publications, and public mockery) to crisis (researchers not willing to share data and engaging in questionable research practices). John, Loewenstein, and Prelec (2012) even argue that minor questionable research practices may, in fact, be more common, and damaging, to the profession. Relatedly, researchers cast doubt on the general replicability of research in psychology, with Hartshorne & Schachner (2012) having only 50 percent replication success. The efforts have now led to large-scale scientific research projects investigating the reproducibility of psychology research (Open

Science Collaboration, 2015). Effect sizes were reported to be 50 percent lower in replications; only two thirds of the statistically significant effects were corroborated by the replications.

Yet, while there are noticeable developments in other (related) disciplines, there is little evidence of the amount of misconduct in management research and the general views on the disciplines replicability and general veracity. Bedeian, Taylor, and Miller (2010) contacted every tenured and tenure-track faculty (1,940 observations) of AACSB accredited PhD granting institutions concerning the prevalence of misconduct along two dimensions: 1) Fabrication and Falsification, and 2) Questionable Research Practices. The authors report that 80 percent of their respondents had knowledge of colleagues withholding methodological details and/or dropping unsupported hypotheses. Hence, researchers were exerting wide discretion in their data preparation and analyses. The dilemma is that individuals put personal gains over scientific peers and society, which may endanger the veracity and legitimacy of scientific research.

In the following, we shed light on the practiced ethics in empirical management research and the impact this might have on the way members view the status of their profession. We believe that integrity and veracity is the most important building block of empirical management research and especially in determining, how stakeholders perceive the implications its research may hold. It is our firm belief that individual malpractices that have been reported are anything but individual frailties. They reflect on the incentives in our publishing system, the way we organize our scrutiny through peer review, and the openness with which we lead our debates about scholarly work. Anything that hampers this process damages everyone in the system. We agree with the notion sketched out in Harley, Faems, and Corbett (2014) that, as disturbing as it might appear, it might help to expose these forms of misconduct and to discuss ways forward.

In this work, we elicit the views of researchers involved in all levels of the publishing process in management research. As a result, we targeted individuals (in their roles as editors and reviewers) attending past Academy of Management meetings and asked about their experience with scientific malpractices. In addition, we wanted to take stock of the general views individual respondents held about the current state of their discipline and whether or not the time is ripe for change. As Honig et al. (2014) state: “We are clearly in transition, but from what to where?”

Overall, we find that although several instances of misconduct have occurred, especially misrepresentation of results, respondents see empirical management research as being in fairly good shape and put a great deal of trust in the replicability and robustness of findings being published. Similarly, many of them regard replication studies as a way forward to ensure credibility, yet are skeptical as to whether or not data should be made publicly available.

THEORETICAL BACKGROUND

The nature of scientific progress

In his landmark work on scientific revolutions Kuhn (1962) challenged the notion of science developing through the accumulation of (accepted) facts and theories. Rather, he argues, scientific progress is driven by discontinuity, periods of revolution, and disruptions to existing paradigms. Paradigms ("universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners" (p. x)) only reign for some time and are replaced by new ways of problem solving, analyzing data, change to the rules, or new ways of doing research time and again.

Individual academics receive credit from and influence the thinking of their peers and the wider academic community (Azoulay, Graff Zivin, & Wang, 2010; Azoulay, Bonatti, Krieger,

and Murray, 2015). Research topics discussed and analyzed within academic peer groups only reach broader consent and acceptance in the academic field if a critical mass of publications, citations, and general outreach can be achieved to convince editors that topics have merit. Similarly, an individual's chance for publication is strongly affected by a researcher's reputation (Laband and Piette, 1994).

Legitimacy and reputation are crucial ways in which the field affects peers and individual academics; it lends support to academics confirming with performance standards (publications). Yet, it may also attest to appropriate behavior or sanction inappropriate behavior, thus, providing guidance on what the field views as acceptable (Azoulay, Bonatti, and Krieger, 2015). Yet, when scientists compete over the interpretation of new paradigms, they are on difficult terrain.

Quoting the transition to Lavoisier's and Newton's new paradigm, Kuhn notes not only a scientific gain but also even more so, a loss of a permissible question and an accepted solution. "Though each may hope to convert the other to his way of seeing science and its problems, neither may hope to prove his case. The competition between paradigms is not the sort of battle that can be resolved by proofs." (Kuhn, 1962: 148). This rejection by the scientific community makes it extremely difficult for scientists to publish work in the first place, often facing rejection through peer review, and more likely, makes them hesitant to accept later scholars revisiting (critically) their prior contributions.

To this end, peer reviewing, serves multiple purposes. It ensures that (prevailing) standards of scientific inquiry are followed and may expose weaknesses and flaws before studies see the day of light, but it also serves a gate-keeper function where works are scanned to make a contribution to the field (Park, Peacey, and Munafò, 2014). As the latter function may induce

reviewers to reject papers that are at the contrary to the prevailing scientific norm, it may actually stifle revolutionary science.

In Tournament theory, an agent's optimal level of effort is achieved when agents compete for a prize that is given out based on relative rank (Lazear and Rosen, 1981; Connelly, Tihanyi, Crook, and Gangloff, 2014). Yet, when the prize spread gets too large, incentives are distorted and riskier behavior than the principal wished for is induced. In scientific research, this may lead to researchers taking short-cuts, engaging in malpractices, misreporting, and even fraud. Honig et al. (2014) speculate that these incentives induce scholars to take short-cuts in the fast paced academic environment, succumb to intense competition, and falter because of high rejection rates along with individualized incentives for publishing.

Park et al. (2014) refer to Mendel famously dropping observations to confirm hypotheses. Though, his theory has stood the test of time, this is now largely overlooked. While Merton (1942: 276) attested to "[...] the virtual absence of fraud in the annals of science" there is currently an ongoing and lively debate about the replicability of published results in psychology, economics, and management research.

Replication – Purpose and Perception

It appears, as if academic integrity and veracity hinges on a nested system of individual researchers, their academic peers, and the academic community (DiMaggio and Powell, 1983; Shadham and Lawrence, 2011). In his 1942 work titled the "The normative structure of science" Robert Merton ascribes four principles to comprise "the ethos of modern science": Universalism (truth claims being subject to impersonal criteria), communism (common ownership of goods), disinterestedness (striving for societal not personal gain), and organized skepticism (critical

scrutiny). The dilemma that pertains is that without organized skepticism it is possible that individuals put personal gains over scientific peers and society. Exercising scrutiny, yet, hinges on the availability of data for other researchers, an opportunity to publish critical results, to put the findings under the microscope and test whether the findings really hold up.

One of the most important avenues likely to reduce incidents of erroneous scholarship, is through systematic replication. Replication serves the purpose of verifying published results, and thus, bringing quantitative management research more in line with the natural sciences. Replication is considered best practice to protect against the uncritical assimilation of erroneous empirical results (Hubbard, Vetter, and Little, 1998).

Only through careful and unbiased re-examination of published data through rigorous replication procedures can scholars be certain that critical findings are trustworthy, and should thus be heeded by the general practitioner community. Evidence based management research demands systematic replication, in support of good and accurate science (Rousseau, 2006).

While some consider exact replication the “scientific gold standard” (Jasny et al., 2011, p. 1225) others eschew the benefits since “a replicated phenomenon may not serve as a rigorous test” (Stroebe and Strack, 2014: 59). In fact, exact replications still have a hard time finding acceptance in academic practice.¹ Replications are still rare in management research, with the noticeable exception of a special issue published in *Strategic Management Journal*.² Also of note, many of the replication studies published in this special issue found, for the most part, the same effect as in the original articles.³

Pashler and Harris (2012) discuss what would happen in the imaginative scenario that investigators attempt a direct replication and fail to obtain effects. They conclude that in an ideal world, the results would become public knowledge through journals or other channels (say

academic meetings for example). They also note that a conceptual non-replication of an established finding is currently nothing more than interesting academic gossip. Especially if in conceptual replication the investigator is potentially to blame for the non-finding.

Notwithstanding the merits of conceptual replication, fatal flaws with the original data collection and preparation will go undetected. In fact, in neither instant was the work by Stapel (Levelt, Noort, & Drenth, 2012), Lichtenthaler (2010) nor Walumbwa et al. (2011) discovered through exact replication. Potentially an exact replication would have worked to uncover the fraudulent asterisks that appeared for insignificant coefficients (Lichtenthaler), yet it might have failed for other instances in which data was not available/existent (Stapel). Though, in some instances, replication helped to identify fraudulent behavior as evidenced in Broockman, Kalla, and Aronow (2015).

In the current system of “publish or perish” in management with an increased focus for theoretical contributions, it is worthwhile to introduce new ideas, yet not necessarily beneficial to validate older ideas (Nosek, Spies, & Motyl, 2012). The problem is, once results are reported, the theoretical implications drive the impact of the articles, thus, the analyses are mere accessories. In essence, there is the potential for a “graveyard of undead theories” (Ferguson & Heene, 2012, p. 555), a situation where theories are “unkillable”, due to an absence of rigorous replication and falsification. Similarly, due to reduced incentives to exactly replicate, retractions in economics and other social sciences are rare (Azoulay, Bonatti, Krieger, and Murray, 2015).

The Open Science Collaboration project denote a studies’ reproducibility as the demarcation between the science and the non-science (Open Science Collaboration, 2012). Using studies published in leading economics journals Chang and Li (2015) were only able to replicate some 50 percent (29 of 59 papers) with help from the authors and conclude that, by and large,

economic research is not reproducible. These effects are similar to other areas of inquiry such as physiology (25 percent replication success; Prinz, Schlange, & Asadullah, 2011) and psychology (about 50 percent replication success; Hartshorne & Schachner, 2012). Camerer et al. (2016) replicated a large number of published experimental analyses (being able to replicate between 60 and 80 percent) while also testing the beliefs of knowledgeable experts in the field using prediction markets (there are no material differences between beliefs about replicability and actual replication efforts).

Failing to investigate and ignoring a low reproducibility rate is problematic. “Self-critique, and the promise of self-correction, is what makes science such an important part of humanity’s effort to understand nature and ourselves (Open Science Collaboration, 2012: 659).” Which bears the question of how scientific (and replicable) is management research?

DATA AND METHODOLOGY

We sent out emails to participants at the Academy of Management (AoM) annual meetings. We identified 38,426 unique email addresses of individuals that have participated in any AoM between 2005 and 2015. Emails were sent separately in August and September of 2016. After the initial round of emails were sent, we let 4 weeks pass to before sending out a reminder email. In sum, we received 1,215 useable replies which correlates to a direct response rate of 4.47 percent (8,682 emails bounced back with an additional 2,575 with out-of-office replies and institution changes). The survey was not linked to any identifiable author information to preserve anonymity.⁴

[Insert table 1 about here]

Table 1 reports our sample composition. Out of the 1,215 we eliminated non-academics, and created non-overlapping groups to distinguish the view of those in an editorial role, and those involved as reviewers or authors. In total, we have 208 editors responding to our questionnaire, 767 individuals that indicated some reviewing activity and 831 that indicate a peer-reviewed publication. Both groups, reviewers and authors are used to contrast the view of editors and to reflect on the broader perspective of the field.

Because we allowed individuals to omit certain answers, the actual number of respondents for subsequent tables may vary. Among the editors, the respondents indicated they had been involved in an editorial role (editor, associate, or department editor) for at least 19 out of the 45 journals represented in the Financial Times (FT) -45 journal ranking. When it came to reviewing, we found that 39 of the FT-45 journals are represented and some 60 percent of the reviewers indicated reviewing activities for these journals.

We asked editors, reviewers, and authors about their experiences in publishing and their views on scholarly management research in general. In line with previous research, we used an indirect approach and asked whether or not respondents have encountered a case of scientific misconduct. This does not necessarily involve a case in which they engaged in the behavior themselves. In fact, work by Fanelli (2009) shows that self-report surveys generally return lower percentages of misconduct. It is important to note, that (similar to other work in this area) our results may not necessarily reflect the true extent of scientific wrong-doings in the profession given that researchers may have different perspectives on acts witnessed as opposed to personally encountered.

Misconduct

In what follows, we explore incidences of scientific malpractice highlighted in other scientific fields. However, we focus explicitly on misconduct (fabrication, data deletion, deceptive representation, or data tampering) and misrepresentations (method inadequacy, omission or withholding of contradictory results, dropping of unsupported hypotheses). Results are tabulated in table 2. The questionable research practice questions (Misconduct) 1-3 are adapted from Gardner, Lidz and Hardwig (2005) and Kattenbraker (2007, as cited in Fanelli, 2009). Question 4 is a variant of a question asked in Kattenbraker (2007, as cited in Fanelli, 2009) and Eastwood, Derish, Leash, and Ordway (1996). Questions related to misconduct originate from prior research in medicine and physiology.

Misrepresentation

We also investigate the prevalence of misrepresentations (method inadequacy, omission or withholding of contradictory results, dropping of unsupported hypotheses). Questions 1 and 2 in Table 3 concerning misrepresentation are adapted from Bebeau and Davis (1996) and Kattenbraker (2007), as cited in Fanelli, 2009). Questions 3 and 4 are adapted from Martinson, Anderson, and D Vries (2005) and Bedeian et al. (2010). For both, misconduct and misrepresentation we not only asked respondents if they had encountered an incident, but how often. Also, because the respondents are asked anonymously, it may be that respondents are referring to the same incident. Questions related to misrepresentation have been previously used in medicine and physiology (Bebeau and Davis (1996); Martinson, Anderson, and D Vries (2005); Kattenbraker (2007), as cited in Fanelli, 2009) and management and economics (Bedeian et al., 2010).

Views on Empirical Management Research

To gain insights into the professions' views regarding the current state of empirical management research, we asked respondents to provide information as to how they agreed or disagreed with a series of statements that aims to capture malpractices reported in earlier research. Answers were elicited on a 5-point Likert scale (Strongly Disagree, Disagree, Neither, Agree, Strongly Agree). We focused on incentives to publish statistically significant results (similar to the questionnaire in Devaney (2001), and as reported in Brodeur et al., 2016); overreliance on supporting theory (based on the theoretical reasoning in Hambrick, 2007, and the discussion in Leung, 2011); the extent to which management results reflect true effect sizes (especially in light of dropping unsupported hypotheses: see O'Boyle, Banks, and Gonzalez-Mulé, 2014); an over-representation of p-values just below 0.1 (Schooler, 2011, O'Boyle, Banks, and Gonzalez-Mulé, 2014); general replicability in the field (Devaney (2001), List et al. (2001), Open Science Collaboration, 2015); and the prevalence of false-positives (Schooler, 2011; Simmons, J. P., Nelson, L. D., & Simonsohn, U., 2011). Topics are based on prior in psychology (Devaney, 2001; Open Science Collaboration, 2015; Simmons, J. P., Nelson, L. D., & Simonsohn, U., 2011; Schooler, 2011) and management and economics (Brodeur et al., 2016; List et al., 2001; O'Boyle, Banks, and Gonzalez-Mulé, 2014; Hambrick, 2007; Leung, 2011).

General Views on Replication Studies

We also asked respondents about their general views of replication that are now widely discussed. Questions originate from prior research in psychology. We elicit (again on 5-point Likert scale) whether respondents see the need to repeat research that others have already

conducted (Open Science Collaboration, 2015) and whether replications contribute in new ways to the field (Pashler & Harris, 2012). Also, we examine whether hesitations exist with respect to replication studies, which might eventually hamper replications from being conducted, published, and whether they will find acceptance in the field. We asked about potential disincentives, such that replication studies may not receive as much attention as the work to be replicated (Pashler and Harris, 2012; Stroebe and Strack, 2014), whether they may lack creativity (Pashler and Harris, 2012; Stroebe and Strack, 2014), and whether or not one would recommend replication studies to be carried out by doctoral students (Pashler and Harris, 2012; Pashler and Wagenmakers, 2012).

Ramifications to the Current State-of-the-Art in Empirical Research

As a result of their analyses of the prevalence of false-positive findings in psychology research, Simmons, Nelson, and Simonsohn (2011) suggested several concrete requirements for authors to restore confidence in the publication process. We build upon their work, and asked respondents whether they see these suggested guidelines and requirements as beneficial for restoring confidence in findings from empirical management studies. We elicit (on a 5-point Likert scale) whether or not authors find agreement concerning rules for data collection termination before such a collection begins; whether authors should list all variables collected for a study; whether reviewers should be tolerant of study imperfections; whether data should be made publicly available; whether all conditions of analyses should be reported; and whether authors should provide analyses of robustness to potential outlier omission.

RESULTS

In this section, we discuss the results of respondents to our survey. We contrast the responses of journal editors (or those involved in the editorial process) to those who have a reviewing and author role in the profession.

In Tables 2 and 3 we report the findings concerning the occurrence of misconduct and misrepresentation disaggregated by reviewers and editors. Tables 4, 5, and 6 report our results concerning views on empirical management research in general, individual's views on replication studies, and the ramifications to the system. Tables 4, 5, and 6 are disaggregated into the views of authors and editors. We condition the groups as non-overlapping, such that editors are not also counted within the reviewing/authoring group. Authors are considered to have at least one peer-reviewed publication. On average authors reported having between 6-10 papers that appeared in scientific peer-reviewed journals.

Misconduct and Misrepresentation

We first asked respondents whether they had encountered a case where the data presented in research was either fabricated or falsified. Both groups reported that the majority of the time they did not encounter such behaviors. However, editors reported that this never happened around 64 percent of the time while general faculty members (that are only involved in reviewing but not editing roles) reported that it never happened approximately 85 percent of the time. The differences here were due to instances when respondents reported that they rarely, sometimes, or often encountered such behaviors. Here, editors reported much higher instances. This is not necessarily surprising given that editors are involved with a greater number of

manuscripts than reviewing individuals would be although, presumably it is the reviewers, who uncover and report such activities.

[Insert table 2 about here]

We also asked whether respondents encountered instances where data in a manuscript had been deleted in a questionable fashion. Once again, the trend holds that both groups reported that the majority of the time, they never encounter such behaviors, however, the margin was only slightly so for editors. Editors reported at a much higher level that they rarely, sometimes, often, or very often encounter this type of behavior than reviewing respondents. In most categories, except for “rarely”, the response rate was double that for editors than reviewers.

Another interesting and revealing question we asked respondents was whether they had encountered manuscripts with deceptive or misleading reported study designs. Here, there was clear separation between results reported by editors and reviewers. While, a slim majority reported that they never encountered such actions, only 37 percent of editors reported the same. As a consequence, editors reported with much greater frequency that they rarely, sometimes, often, or very often encountered this behavior. In fact, the largest separation between the two groups occurred in the “often” or “very often” category where editors were approximately three times more likely to report that they found this activity taking place. Unlike some of the other questions asked, this difference is stark and troubling.

Finally, we asked respondents if they had encountered cases where authors had changed or omitted data points from a study. These results are more in line with previously reported responses. While both had majorities reporting that they never encountered such actions, the

responses were weaker for editors. In addition, for all of the other categories (rarely, sometimes, often, very often) editors reported encountering this form of questionable/unethical behavior at a much higher rate.

We then turned our attentions to issues revolving around misrepresenting findings in manuscripts, which are found in Table 3. Unlike the results presented above, the responses show that these problems are much more pervasive in the profession, despite the fact that editors are more likely to report having rarely, sometimes, often, or very often encountering such actions. Unlike the previous section, not one question in this series of questions had respondents reporting a majority “never” encountering such actions.

[Insert table 3 about here]

We asked respondents if they had encountered cases where authors had used inadequate or inappropriate research designs. Over 40 percent of editors and 25 percent of reviewers reported that they often or very often encounter colleagues engaging in such activities.

In addition, we asked respondents if they had encountered cases where manuscripts failed to present data that would have contradicted previous research. Presenting findings that do not conform with previously published works could lead to rejection or increased scrutiny. As a way of avoiding this, authors might be tempted to leave out such findings for fear that the reviewer might have published different results. This seems to be the case in the profession with nearly half of editors and over 30 percent of reviewers reporting that they sometimes, often, or very often find manuscripts where this has taken place.

In the vein of the previous question, we asked respondents if they had encountered cases where the details of the methodology or results were withheld. As stated, if results contradicted previously published work or if results did not support the a priori assumptions of the model, they were simply withheld from the manuscript. Only 18 percent of editors and 25 percent of reviewers reported that they had never encountered this activity. On the other end, 11 percent of editors and 5 percent of reviewers reported that they find this happening very often.

Finally, we asked respondents if they had encountered manuscripts where unsupported hypothesis had been dropped from a manuscript. A little over a quarter of editors reported that they had never encountered this behavior while slightly over a quarter of editors reported that this behavior was encountered often or very often. The responses from reviewers were similar with more reporting that they had never encountered such behavior and less reporting that they had often or very often found this in manuscripts.

Views on Empirical Management

The results of Table 4 reveal that we began by asking prospective groups whether they felt that there were strong incentives in place to publish results that were statistically significant. For both editors and reviewers nearly 90 percent agreed or strongly agreed.⁵ As mentioned before, this could be a result of authors not wanting to report findings contradictory to published works. However, it does go against the scientific notion of “being led by the data.” It is troubling and could cast doubt on the reliability of management research. Leung (2011), echoing Hambricks (2007) discussion of management researcher’s obsession with theory, explains that there might be temptations to engage in post-hoc theorizing, after results are known. Noted by Kerr (1998), this is often referred to as HARKing (Hypothesizing After the Results are Known).

In a similar fashion, O'Boyle, Banks, and Gonzalez-Mulé (2014) document how authors drop statistically nonsignificant hypotheses when papers come closer to publication.

Hence, another way of asking the same question is to ask if the power of established theories would be so great that only papers supporting said theories would find their way into journals. In this regard, only about 40 percent of journal editors and reviewers agreed or strongly agreed with this statement. For journal editors, a bit over a quarter, and for reviewers, about a third could neither agree nor disagree with this this statement. In some sense, these results differ from those reported above.

[Insert table 4 about here]

While significant results may not be replicable, effect sizes might even be declining in replication attempts. A strong and widespread engagement in questionable research practices may invalidate a large part of empirical results. Omitting non-significant variables, outliers, or dropping unsupported hypotheses may falsely inflate the effect size in the study reported (Schooler, 2011). Authors are keenly aware of the distribution of the effect of sizes. When asked whether published works in management research were truly reflective of these effects less than 10 percent for each group disagreed or strongly disagreed. In other words, the overwhelming majority of respondents were neutral, agreed, or strongly agreed that the results that appeared in print did, indeed, reflect the true distribution of effect sizes.

We asked respondents if there was an overrepresentation of p-values in the tail of the distribution just below $\text{Alpha} = 0.1$. The results were mixed. Almost half (for both groups) could neither agree nor disagree with this statement. For both editors and reviewers over 30 percent

either agreed or strongly agreed with the statement. Taken as a whole, it would appear that respondents were concerned that there was indeed an overrepresentation in published works of results with high level of statistical significance.

The key to empirical work is that it be replicable and then expanded upon. There was some skepticism about whether published work in management research was generally replicable, however it was not wide spread. For both editors and reviewers, over 40 percent of respondents agreed or strongly agreed that management research was generally replicable. Troubling was that for both groups, nearly a third responded that they neither agreed nor disagreed with the statement. Of course, this could reflect that respondents did not have enough information to make an informed opinion on the matter but it is more likely that they were not entirely sure regarding replicability.

Finally, we asked respondents how they felt about the notion of there being persistence in false negative results in management research. In other words, did results that had been disproved or unsupported by the data still appear in management manuscripts. For general reviewers, 50 percent could neither agree nor disagree with this notion. For editors the response rate was still a very high 39 percent. Of those abler to state definitively, over 30 percent of journal editors and nearly 25 percent of reviewers either agreed or strongly agreed with this notion.

Replication

As alluded to in the previous section, management scholars were not uniform in their opinions of the amount of work that could be replicated in journals. In Table 5 we explore the idea of replication even further.

We firstly asked both groups if they felt it was important to repeat work that had already be conducted or published. For both editors and reviewers, over 70 percent agreed or strongly agreed that such activities should take place. ⁶ Conversely, both groups had less than 10 percent of respondents reporting that they disagreed or strongly disagreed with this activity.

We asked respondents how they viewed the idea that when management research is replicated, early positive studies often receive more attention than later negative ones. For editors, over 50 percent either agreed or strongly agreed with the statement while for reviewers more than 60 percent of respondents either agreed or strongly agreed. However, the issue is not settled because for both groups a sizeable third neither agreed nor disagreed with the statement.

Unlike the natural sciences where replication of previous studies has value in and of itself, our respondents were less in agreement when asked whether replication studies lacked creativity and originality. Over 40 percent of editors and reviewers agreed or strongly agreed with the notion that replication studies were not creative nor original. However, for both groups over a quarter of respondents disagreed or strongly disagreed with this sentiment.

[Insert table 5 about here]

To follow up on the previous question we asked respondents how they felt about the notion that replication studies added to the field in new ways. Over 60 percent for editors and 70 percent for reviewers agreed or strongly agreed with the statement. However, this seems to stand in contrast to the previous question since so many felt that replication studies were not original or creative. The two notions do not have to be mutually exclusive of each other but do highlight the lack of uniformity in regards to this important area among management scholars.

Frank and Saxe (2012) point out that students may benefit from engaging in replication studies by witnessing the process of scientific discovery firsthand, as replicators, instead of only as passive readers. Students achieve this by moving beyond nicely written papers, paying careful attention to methods employed, and understanding and scrutinizing researchers' discretionary choices.

Finally, we therefore asked respondents if they would not advise doctoral students to start their careers by doing careful replication studies of previous research. Having students recreate the works of others is a common learning technique in graduate programs given that it teaches skills and gets students acquainted with the data. Interestingly, over half of the editors and reviewers either disagreed or strongly disagreed with this statement. Thus, they felt it was a worthy endeavor for beginning scholars. Going through the painful work and frustration of not being able to replicate another researcher's paper can help students become better researchers themselves. In addition, over 20 percent of both groups were neither in agreement or disagreement with the statement. Thus, a majority of management scholars see value in this type of exercise and felt that young scholars were engaged in worthy activities by doing so.

Ensuring Credibility

In Table 6 we turned our attention to what should and could be done to ensure that management research be perceived as credible. In the end, if the output of scholarly efforts is doubted, the profession and those engaged in it could be viewed with skepticism by end users.

We firstly asked respondents if authors must decide the rule for terminating data collection before data collection begins and report this rule in the manuscript. By a wide margin both editors and reviewers disagreed or strongly disagreed with this notion. In fact, for both

groups around 20 percent agreed or strongly agreed. Thus, this is not a solution likely to be seen in future manuscripts.

We asked if authors must list all variables collected for a study. Responses were mixed. For editors over 40 percent disagreed or strongly disagreed with the statement while approximately 30 percent agreed or strongly agreed. For reviewers the responses were about the same. Thus, this is another area where no consensus of opinion was found on how to increase the credibility of research findings by management scholars.

We also asked respondents if reviewers should be more tolerant of imperfections in reported results. The answer seems to be a resounding no. For both editors and reviewers over 60 percent disagreed or strongly disagreed with this sentiment. Only 20 percent of editors and 16 percent of reviewers agreed or strongly agreed.

[Insert table 6 about here]

Many journals are now requiring that when the manuscript is submitted for review, the original data and materials must be supplied also. Not all will make the data publically available but some will. We asked respondents if journals should require that the data and materials be provided and made publically available. A sizeable majority of editors disagreed or strongly disagreed with that statement. Nearly half of reviewers disagreed or strongly disagreed. Conversely, less than 20 percent of editors agreed or strongly agreed with the statement. This is important given that for such a policy to be implemented it would come from editors and editorial boards. Thus, it is not likely that wider adoption of this policy will be implemented immediately.

We also asked whether authors must report all experimental conditions including failed manipulations. For both editors and reviewers, over 70 percent disagreed or strongly disagreed with this statement. There clearly was little support among respondents for such a requirement for management scholars.

Finally, we asked respondents if observations were eliminated, should authors report what the statistical significance of findings would have been had the eliminated variables been kept. For editors, over 70 percent disagreed or strongly disagreed with this notion. For reviewers over 60 percent disagreed or strongly disagreed. Once again, this is not a suggestion that is likely to catch on with management scholars given the strong and overwhelming negative responses.

Post-hoc Analyses

To provide further evidence on how often individuals have encountered cases of misconduct and misrepresentation we provide further post-hoc tests. In doing so, Table 7 uses the variables misconduct and misrepresentation as the dependent variable. Here, a factor analysis reveals two unique factors on which the questions in Tables 2 and 3 load, respectively. Cronbach's alpha is 0.82 for misconduct and 0.80 for misrepresentation. Due to the nature (average of four individual items) of the variables, we employ an OLS regression for each.

Concerning the explanatory variables, we include the role individuals serve: Individuals can act as *editors* (Managing Editor, Editor-In-Chief =1, 0 otherwise), *department or associate editors* (Department/Associate Editor=1, 0 otherwise), and we further delineate their *reviewing* activities for journals (Reviewed for FT-45 journal=1, 0 otherwise). Editors can also be reviewers. Categorizations are not mutually exclusive.⁷ Similarly, we control for the productivity of the scholars by including the *number of scientific publications* (1=0 publications,

2=1-5 publications, 3=6-10 publications, 4=11-20 publications, and 5=>20 (co)-publications during 2006 and 2010). In addition, we also break down this variable into the *number of FT45 journal list publications* with the equivalent coding scheme. Regarding personal characteristics, we also include the *academic position* (1= Doctoral/PhD student, 2=PostDoc, 3=Assistant Professor, 4=Associate Professor, 5=Full Professor and 0=Other), *age* (real number), *gender* (female=1, male=0), whether they hold a *PhD* (PhD completed=1, 0 otherwise), and location of university where they are currently employed (US=1, Europe, South America, Asia Pacific, Africa = 0).

[Insert table 7 about here]

Overall, we had some 44 percent of respondents from US based institutions and 45 percent from European institutions. One third were women; 208 respondents indicated editor in-chief, department editor or associate editor responsibilities; about 50 percent review for FT 45 journals; on average scholars report 6-10 publications between 2006-2010; 40 percent report zero FT 45 publications; almost 50 percent report between 1 and 5 FT 45 publications; the average age is 47 with 23 percent being assistant professors; 29 percent associate professors and 38 percent full professors.⁸

Results are depicted in Table 7. Firstly, the results paint a fairly consistent picture that the more experienced scholars have experienced more instances of academic misconduct and misrepresentation. Yet, these effects are not embedded in life experience (such as age and academic position) but are more related to experience with the academic publishing process, be it as authors or reviewers. The strongest effects can be found for those individuals that serve as editor or department/associate editors. Interestingly, academics that review for FT-45 journals have experienced more instances of misrepresentation than those who serve as reviewers at other

journals, which attests to the role of misguided incentives in the publishing system laid out earlier.

Based on these findings, we also investigated how experience with misconduct and misrepresentation in turn, effects the views scholars hold on the state of the discipline. We report the results in Table 8. In here, we include the two variables misconduct and misrepresentation as explanatory variables and employ the same control variables as reported in Table 7.

[Insert table 8 about here]

Noteworthy, those academics that report higher levels of misrepresentation encountered, generally attest to stronger misguided incentives to publish significant results and to stronger incentives to support theories, which may result in the reporting of only statistically significant results. Among the control variables, we find mixed results with the only exception that US based scholars attest more to the notion of incentives for reporting significant results. In addition, reviewers and department editors of FT-45 journals are less likely to concur with an overrepresentation of results slightly below the 10 percent cut-off value. Interestingly though, we find mixed evidence as it relates to the trustworthiness of results. Those that have experienced misconduct and misrepresentation reflect positively on the true effect sizes reported and the general level of replicability, though they similarly attest to a prevalence of false positives. Noteworthy, especially reviewers at FT-45 journals hold more positive views on replicability, while females and US based scholars are far more skeptical.

[Insert table 9 about here]

In Table 9, we reveal the views that individuals hold on replication studies. Importantly, those that have experienced more instances of misreporting see greater value in replication studies. Hence, replication might safeguard against overstating results and other minor research

offenses. Yet, those who have experienced more misconduct do not see much value in replication studies, which potentially shows that replications can only be a part of the answer against individual frailties and that other (likely complementary) measures are called for. It is quite assuring that editors and department editors do not marginalize the value of replication studies and regard replications as beneficial. Yet, those who have experienced misrepresentation paint a different picture. There still seems to be a mismatch between misrepresentations observed and the number of replication studies published.

It is also interesting to note that scholars who published more in FT 45 journals and those who are holding higher academic positions see replication studies as somewhat original which may help to further their introduction into academic curricula. Noteworthy is that more productive scholars publishing in more prestigious outlets, would not discourage their PhD students from doing replication studies (while the productive scholars publishing in less prestigious outlets would). In addition, US based scholars and those who have experienced misrepresentations would be more encouraging to PhD students when engaging in replication studies. Hence, while we do not see many replication studies in management research yet, the incentives are set for engaging more intensely in this area and editors would not generally eschew on them being publishing.

[Insert table 10 about here]

Lastly, as it relates to potential ramifications, Table 10 corroborates the previous findings that open data policies are unlikely to catch up. Editors are generally opposed. Similarly, those who have experienced more instances of misrepresentation do not see open data and more transparency as a beneficial measure to ensure credibility. In addition, those who have experienced misconduct do not regard more information on data collection and variables as a

potential ramification. More transparency may potentially help against misrepresentation but not outright fraud. Among the control variables, we find that US based scholars are substantially more in favor of open data policies and more transparency, exhibiting an interesting divide.

Limitations

As with all research, our study is not without limitations. Firstly, our data resembles the Academy of Management composition fairly well in terms of US and European based scholars. However, we have received emails from Asian based scholars that they were not able to fill out the google-based questionnaire. Hence, this might limit the generalizability of our findings as it relates to non-US and non-European scholars. Similarly, our research touches upon misconduct and misrepresentation as they relate to quantitative empirical studies. These types of studies clearly represent only a part of the empirical work carried out in management research. By no means do we claim this study to be exhaustive of other similar and potentially equally malicious wrongdoings as they relate to qualitative research, case study based work, mathematical and simulation based modelling, structural equation modelling, among others.

Moreover, the list of journals for which scholars review gives some indication for which sub-disciplines our results are generalizable and for which sub-disciplines they are not. We have individuals reviewing for nearly all journals in the FT-45 list. Yet, we have a very small number of individuals reviewing for Accounting, Finance, and Economics journals. Hence, our results may have limited value for these sub-disciplines. Similarly, while being incorporated more frequently, Marketing and Operations Management journals are in the minority. Given particularities in terms of methods employed, these journals might warrant a closer look into certain types of misbehavior that our questionnaire cannot capture.

By and large, however, our results are reflective of journals like *Academy of Management Journal/Review/Perspectives*, *Journal of Business Ethics*, *Journal of Management Studies*, *Organization Studies*, *Management Science*, *Organization Science*, *Administrative Science Quarterly*, *Entrepreneurship Theory and Practice*, *Journal of Business Venturing* and *Strategic Management Journal*. For most of these journals we have more than 10 percent of respondents indicating reviewing activities.

Lastly, we do not claim our list of misrepresentation and misconduct to be exhaustive either, nor do we claim that we have found evidence at the individual level. It may well be that scholars report instances they have encountered, and the extent of misreporting and misconduct would better be investigated by asking scholars individually about their intentions to commit or their prior own wrongdoings directly.

CONCLUSIONS AND IMPLICATIONS

Our findings indicate that scientific misconduct is quite common in empirical management research. Editors and reviewers involved in management research report evidence especially on the prevalence of misrepresentation in empirical management research. However, we find it even more astonishing that not more replication studies are published to exert scrutiny. Especially now that it is possible to use online appendices, repositories and the like, to document how original work was carried out. Special sections of journals are devoted to book reviews and teaching cases, but not replication studies.

Noteworthy, our results also show that researchers at the same time paint a quite different picture when it comes to the widely acclaimed reproducibility of results. By and large, respondents believe empirical research to reflect the true distribution of effect sizes and they

consider results as being largely replicable. This is surprising given the high prevalence of reported cases to the contrary, especially when it comes to misrepresentation of data. In our view, the current practice of dealing with empirical papers leaves many things in the dark, on both sides of the process. As a reviewer, it is very difficult to assess the validity of a paper. No author wants to be scooped or to be the reviewer of the next Stapel, Lichtenthaler, or Walumbwa paper. In his seminal paper, Edward Leamer (1983) urged economists to “take the con out of econometrics”. Now, more than 30 years later, has there ever been a timelier and pointed call for action for empirical research?

When it comes to future ramifications and extensions of current processes, both, reviewers and editors take a strong stance against making data publicly available to limit degrees of freedom in empirical research. This is in line with concurrent work that reports the sharing of data openly as a hotly debated topic. A recent survey by Berghmans et al. (2017) found that while 73 percent of researchers agree that they would benefit from having access to other researcher’s data, still a third of the respondents would not share their data. Respondents cite reasons related to privacy, data ownership, and control/trust considerations as the primary reasons. Data ownership appears to be a major source of disagreement (Borgman, 2012, 2015). Also, editors fear that researchers might gain unearned glory from the work that other’s put into data collection efforts and that open data might be used to unfairly discredit authors or undermine original publications (Longo and Drazen, 2016; Deveraux et al., 2016). While it appears attractive at first glance, data sharing might bring about hidden costs that have hitherto not been studied in a broader fashion. Not surprisingly, we find management researchers are skeptical and hesitant to accept open data policies.

Yet, they seem to be much more open to replication based solutions that might expose mistakes in original findings or may extend previous work using newer or better suited methods that become available. Review articles and meta analyses have found wide acceptance in the field, and are regarded as a respected way to assess a cumulative body of knowledge. Yet, they strongly hinge on complete and accurate information. It is noteworthy to see, that meta analyses have gained substantial traction in the management field (and claim to be a step towards evidence based management) but the underlying issues of appropriate peer reviewing and ensuring appropriate empirical analyses have not been solved sufficiently (Jefferson, Wager, Davidoff, 2002; Honig et al., 2014).

Also, clearly not every paper needs to be replicated, but for the most influential ones (those that are cited often and carry implications for a great body of subsequent work), there clearly is a need for replication (Ioannidis and Khoury, 2014). We believe that management research would benefit from more replication as a tool to improve the quality of analysis and interpretation and to reduce the likelihood of incidences of malpractice. In physiology, molecules foreign to the body may induce an immune reaction and help to develop an adaptive response. We see incidences of scientific malpractice as such an intrusive encounter for which we need to develop an immunization response. Yet, our results also show that scholars believe that replication studies are important when dealing with misrepresentation but not necessarily fraud.

In addition, there may be many defensible assumptions that can be made when analyzing data. Hence, replications might not be a panacea for empirical research. Along these lines, other fields have already started to pave the way forward. If there is dispute about assumptions being made, adversarial collaboration would be more fruitful (Mellers, Hertwig, & Kahneman, 2001). This approach requires the involved parties to agree to work together on empirical tests to

resolve the dispute (with the help of an arbiter). In a laudable effort, for example, Silberzahn, Simonsohn, and Uhlman (2014) have even teamed up to reconcile ambiguous findings in Silberzahn and Uhlmann (2013) in a very professional and scientifically productive fashion.

Important scientific contributions should not rest on the assumption that researchers have done everything in their power to ensure the robustness of the initial findings. Especially in light of pressure to search for significance, for tailoring theory after the finding, file-drawer problems, there is a need for scrutiny (Nosek et al., 2012; Open Science Collaboration, 2012, Jasny et al. 2011; Karabag, & Berggren, 2016). To establish candor in science, replication plays a crucial role in building and extending our cumulative knowledge. Management science needs to learn to protect itself; replication may lead to fortification. As the Open Science Collaboration (2015: 943) vividly states “Innovation points out paths that are possible; replication points out paths that are likely; progress relies on both.”

Informed consent: Informed consent was obtained from all individual participants included in the study

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards

References

- Azoulay, P., Bonatti, A., & Krieger, J. L. (2015). *The career effects of scandal: Evidence from scientific retractions* (No. w21146). National Bureau of Economic Research.
- Azoulay, P., Furman, J. L., Krieger, J. L., & Murray, F. (2015). Retractions. *Review of Economics and Statistics*, 97(5), 1118-1136.
- Azoulay, P., Graff Zivin, J., & Wang, J. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, 125(2), 549-589. doi: doi:10.1162/qjec.2010.125.2.549
- Barkema HG, Vermeulen F. 1998. International expansion through start-up or acquisition: a learning perspective. *Academy of Management Journal* 41(1): 7–26.
- Bebeau MJ, Davis EL (1996) Survey of ethical issues in dental research. *Journal of Dental Research* 75: 845–855.
- 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), 715-725.
- Berghmans, Stephane; Cousijn, Helena; Deakin, Gemma; Meijer, Ingeborg; Mulligan, Adrian; Plume, Andrew; de Rijcke, Sarah; Rushforth, Alex; Tatum, Clifford; van Leeuwen, Thed; Waltman, Ludo (2017), "Open Data: the researcher perspective - survey and case studies", Mendeley Data, v1 <http://dx.doi.org/10.17632/bwrnfb4bvh.1>

- Borgman, C.L. (2012) The Conundrum of Sharing Research Data. *JASIST* 63 (6): 1059–78.
doi:10.1002/asi.22634.
- Borgman, C.L. (2015) Big data, little data, no data: scholarship in the networked world.
Cambridge, MA: The MIT Press. www.Mitpress.mit.edu/big-data
- Brodeur, A., Lé, M., Sangnier, M., & Zylberberg, Y. (2016). Star wars: The empirics strike back.
American Economic Journal: Applied Economics, 8(1), 1-32.
- Broockman, D., Kalla, J., & Aronow, P. (2015). Irregularities in LaCour (2014). *Work. pap.*,
Stanford Univ. http://stanford.edu/~dbroock/broockman_kalla_aronow_lg_irregularities.pdf.
- Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., ... & Heikensten, E. (2016). Evaluating replicability of laboratory experiments in economics. *Science*,
351(6280), 1433-1436.
- Chang, A. C., & Li, P. (2015). Is Economics Research Replicable? Sixty Published Papers from
Thirteen Journals Say 'Usually Not'.
- Connelly, B. L., Tihanyi, L., Crook, T. R., & Gangloff, K. A. (2014). Tournament theory thirty
years of contests and competitions. *Journal of Management*, 40(1), 16-47.
- DeVaney, T. A. (2001). Statistical significance, effect size, and replication: What do the journals
say?. *The Journal of Experimental Education*, 69(3), 310-320.
- Devereaux, P. J., Guyatt, G., Gerstein, H., Connolly, S., & Yusuf, S. (2016). Toward fairness in
data sharing. *The New England journal of medicine*, 375(5), 405-407.

- DiMaggio, P., & Powell, W. W. (1983). The iron cage revisited: Collective rationality and institutional isomorphism in organizational fields. *American Sociological Review*, 48(2), 147-160.
- Eastwood S, P. Derish, E. Leash, S. Ordway (1996). Ethical issues in biomedical research: Perceptions and practices of postdoctoral research fellows responding to a survey. *Science and Engineering Ethics* 2: 89–114.
- Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS one*, 4(5), e5738.
- Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), 555-561.
- Frank, M. C., & Saxe, R. 2012. Teaching replication. *Perspectives on Psychological Science*, 7(6): 600-604.
- Gardner W, Lidz CW, Hartwig KC (2005) Authors' reports about research integrity problems in clinical trials. *Contemporary Clinical Trials* 26: 244-251.
- Hambrick, D. C. 2007. The field of management's devotion to theory: Too much of a good thing? *Academy of Management Journal*, 50(6): 1346-1352.
- Harley, B., Faems, D., & Corbett, A. (2014). A few bad apples or the tip of an iceberg? Academic misconduct in publishing. *Journal of Management Studies*, 51(8), 1361-1363.
- Hartshorne, J. K., & Schachner, A. (2012). Tracking replicability as a method of post-publication open evaluation. *Beyond open access: visions for open evaluation of scientific papers by post-publication peer review*, 70.

- 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):118-142.
- Hubbard, R., Vetter, D. E., & Little, E. L. (1998). Replication in strategic management: Scientific testing for validity, generalizability, and usefulness. *Strategic Management Journal*, 243-254.
- Ioannidis, J. P., & Khoury, M. J. (2014). Assessing value in biomedical research: the PQRST of appraisal and reward. *JAMA*, 312(5), 483-484.
- Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. 2011. Again, and again, and again.... *Science*, 334(6060): 1225-1225.
- Jefferson, T., Wager, E., & Davidoff, F. (2002). Measuring the quality of editorial peer review. *Jama*, 287(21), 2786-2790.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological science*, 0956797611430953.
- Karabag, S. F., & Berggren, C. (2016). Misconduct, marginality and editorial practices in management, business and economics journals. *PloS one*, 11(7), e0159492.
- Karabag, S. F., & Berggren, C. 2012. Retraction, dishonesty and plagiarism: Analysis of a crucial issue for academic publishing, and the inadequate responses from leading journals in economics and management disciplines. *Journal of Applied Economics and Business Research*, 2(3): 172-183
- Kattenbraker MS (2007) Health education research and publication: ethical considerations and the response of health educators. Carbondale: Southern Illinois University. Doctoral thesis.

- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2(3), 196-217.
- Kuhn, Th. (1962). *The Structure of Scientific Revolutions* (1st ed.). University of Chicago Press.
- Laband, D. N., & Piette, M. J. (1994). Favoritism versus search for good papers: Empirical evidence regarding the behavior of journal editors. *Journal of Political Economy*, 102(1), 194-203.
- Lazear, E. P., & Rosen, S. (1981). Rank-Order Tournaments as Optimum Labor Contracts. *The Journal of Political Economy*, 89(5), 841-864.
- Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review*, 73(1), 31-43.
- Leung, K. (2011). Presenting post hoc hypotheses as a priori: Ethical and theoretical issues. *Management and Organization Review*, 7(3), 471-479.
- Levelt, C., Noort, C., & Drenth, C. (2012). Falende wetenschap: De frauduleuze onderzoekspraktijken van sociaal-psycholoog Diederik Stapel. *Final report of the Committee Levelt*, 28.
- Lichtenthaler, U. (2010). RETRACTED: Determinants of proactive and reactive technology licensing: A contingency perspective. *Research Policy*, 39(1), 55-66.
- List, J. A., Bailey, C. D., Euzent, P. J., & Martin, T. L. (2001). Academic economists behaving badly? A survey on three areas of unethical behavior. *Economic Inquiry*, 39(1), 162-170.
- Longo, D. and J. Drazen (2016): Data Sharing, *New England Journal of Medicine*, 374:276-277
- Martinson, B. C., Anderson, M. S., & De Vries, R. (2005). Scientists behaving badly. *Nature*, 435(7043), 737-738.

- Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, 12(4), 269-275.
- Merton, R. K. (1942). Note on Science and Democracy, *A. J. Legal & Pol. Soc.*, 1, 115.
- Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, 7(6), 615-631.
- O'Boyle, E. H., Banks, G. C., & Gonzalez-Mulé, E. (2014). The Chrysalis Effect How Ugly Initial Results Metamorphosize Into Beautiful Articles. *Journal of Management*, 0149206314527133.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
- Open Science Collaboration. 2012. An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, 7(6): 657-660.
- Park, I. U., Peacey, M. W., & Munafò, M. R. (2014). Modelling the effects of subjective and objective decision making in scientific peer review. *Nature*, 506(7486), 93-96.
- Park, U. D., Borah, A., & Kotha, S. (2016). Signaling revisited: The use of signals in the market for IPOs. *Strategic Management Journal*, 37(11), 2362-2377.
- Pashler, H., & Harris, C. R. 2012. Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, 7(6): 531-536.

- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science a crisis of confidence?. *Perspectives on Psychological Science*, 7(6), 528-530.
- Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets?. *Nature reviews Drug discovery*, 10(9), 712-712.
- Rousseau, D. M. 2006. Is there such a thing as "evidence-based management"? *Academy of management review*, 31(2): 256-269.
- Schooler, J. (2011). Unpublished results hide the decline effect. *Nature*, 470(7335), 437.
- Shadnam, M., & Lawrence, T. B. (2011). Understanding widespread misconduct in organizations: An institutional theory of moral collapse. *Business Ethics Quarterly*, 21(03), 379-407.
- Silberzahn, R., & Uhlmann, E. L. 2013. It pays to be Herr Kaiser: Germans with noble-sounding surnames more often work as managers than as employees. *Psychological Science*, 0956797613494851.
- Silberzahn, R., Simonsohn, U., & Uhlmann, E. L. 2014. Matched-Names Analysis Reveals No Evidence of Name-Meaning Effects: A Collaborative Commentary on Silberzahn and Uhlmann (2013). *Psychological science*, 25(7): 1504-1505.
- 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*, 0956797611417632.
- Steen, R. G. (2010). Retractions in the scientific literature: is the incidence of research fraud increasing?. *Journal of medical ethics*, jme-2010.

- Stroebe, W., & Strack, F. 2014. The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, 9(1): 59-71.
- Tsang, E. W., & Yamanoi, J. (2016). International expansion through start-up or acquisition: A replication. *Strategic Management Journal*, 37(11), 2291-2306.
- Walumbwa, F. O., Luthans, F., Avey, J. B., & Oke, A. 2011. Retracted: Authentically leading groups: The mediating role of collective psychological capital and trust. *Journal of Organizational Behavior*, 32(1): 4-24.

Appendix

Table 1: *Number of respondents in differing roles*

	EDITORS	REVIEWERS	AUTHORS
TOTAL RESPONDENTS: 1,215			
ACADEMIC	208	767	831
FEMALE	56	280	361
ASSISTANT PROFESSOR	10	192	200
ASSOCIATE PROFESSOR	57	193	203
FULL PROFESSOR	121	240	248

TABLE 2						
QUESTIONABLE RESEARCH PRACTICES (MISCONDUCT)						
	NEVER	RARELY	SOMETIMES	OFTEN	VERY OFTEN	TOTAL
HOW OFTEN HAVE YOU ENCOUNTERED THE CASE THAT AN AUTHOR...						
... FABRICATED OR FALSIFIED DATA (EDITORS)	125	53	11	6	1	196
	63.78%	27.04%	5.61%	3.06%	0.51%	
... FABRICATED OR FALSIFIED DATA (REVIEWERS)	648	87	18	7	4	764
	84.82%	11.39%	2.36%	0.92%	0.52%	
... DELETED DATA IN AN UNJUSTIFIED WAY (EDITORS)	99	58	25	11	3	196
	50.51%	29.59%	12.76%	5.61%	1.53%	
... DELETED DATA IN AN UNJUSTIFIED WAY (REVIEWERS)	537	156	47	17	5	762
	70.47%	20.47%	6.17%	2.23%	0.66%	
... DECEPTIVELY OR MISLEADINGLY REPORTED STUDY DESIGN (EDITORS)	72	62	38	18	6	196
	36.73%	31.63%	19.39%	9.18%	3.06%	
... DECEPTIVELY OR MISLEADINGLY REPORTED STUDY DESIGN (REVIEWERS)	388	215	120	30	8	761
	50.99%	28.25%	15.77%	3.94%	1.05%	
... CHANGED OR OMITTED DATA POINTS IN A STUDY (EDITORS)	103	48	27	14	3	195
	52.82%	24.62%	13.85%	7.18%	1.54%	
... CHANGED OR OMITTED DATA POINTS IN A STUDY (REVIEWERS)	511	162	58	21	7	759
	67.33%	21.34%	7.64%	2.77%	0.92%	

TABLE 3						
QUESTIONABLE RESEARCH PRACTICES (MISREPRESENTATION)						
	NEVER	RARELY	SOMETIMES	OFTEN	VERY OFTEN	TOTAL
HOW OFTEN HAVE YOU ENCOUNTERED THE CASE THAT AN AUTHOR...						
... USED INADEQUATE OR INAPPROPRIATE RESEARCH DESIGNS (EDITORS)	21 10.66%	30 15.23%	62 31.47%	60 30.46%	24 12.18%	197
... USED INADEQUATE OR INAPPROPRIATE RESEARCH DESIGNS (REVIEWERS)	134 17.61%	185 24.31%	252 33.11%	134 17.61%	56 7.36%	761
.. FAILED TO PRESENT DATA THAT CONTRADICTS PREVIOUS RESEARCH (EDITORS)	51 26.29%	55 28.35%	46 23.71%	31 15.98%	11 5.67%	194
.. FAILED TO PRESENT DATA THAT CONTRADICTS PREVIOUS RESEARCH (REVIEWERS)	330 43.42%	192 25.26%	141 18.55%	75 9.87%	22 2.89%	760
... WITHHELD DETAILS OF METHODOLOGY OR RESULTS (EDITORS)	35 17.86%	51 26.02%	55 28.06%	33 16.84%	22 11.22%	196
... WITHHELD DETAILS OF METHODOLOGY OR RESULTS (REVIEWERS)	191 25.13%	247 32.50%	174 22.89%	108 14.21%	40 5.26%	760
... DROPPED UNSUPPORTED HYPOTHESES FROM AN ARTICLE (EDITORS)	51 26.15%	46 23.59%	41 21.03%	36 18.46%	21 10.77%	195
... DROPPED UNSUPPORTED HYPOTHESES FROM AN ARTICLE (REVIEWERS)	276 36.46%	203 26.82%	133 17.57%	82 10.83%	63 8.32%	757

TABLE 4						
VIEWS ON EMPIRICAL MANAGEMENT RESEARCH (EDITORS)						
	STRONGLY DISAGREE	DISAGREE	NEITHER	AGREE	STRONGLY AGREE	TOTAL
TO WHICH EXTENT DO YOU AGREE WITH THE FOLLOWING STATEMENTS?						
THERE ARE STRONG INCENTIVES TO PUBLISH STATISTICALLY SIGNIFICANT RESULTS IN EMPIRICAL MANAGEMENT RESEARCH (<i>EDITORS</i>)	1 0.48%	1 0.48%	21 10.10%	60 28.85%	125 60.10%	208
THERE ARE STRONG INCENTIVES TO PUBLISH STATISTICALLY SIGNIFICANT RESULTS IN EMPIRICAL MANAGEMENT RESEARCH (<i>AUTHORS</i>)	5 0.60%	15 1.81%	60 7.22%	187 22.50%	564 67.87%	831
THE POWER OF ESTABLISHED THEORIES IS SO GREAT THAT ONLY RESULTS THAT SUPPORT THEM FIND THEIR WAY INTO JOURNALS (<i>EDITORS</i>)	17 8.17%	53 25.48%	56 26.92%	54 25.96%	28 13.46%	208
THE POWER OF ESTABLISHED THEORIES IS SO GREAT THAT ONLY RESULTS THAT SUPPORT THEM FIND THEIR WAY INTO JOURNALS (<i>AUTHORS</i>)	57 6.86%	142 17.09%	282 33.94%	248 29.84%	102 12.27%	831
RESULTS PUBLISHED IN MANAGEMENT RESEARCH REFLECT THE TRUE DISTRIBUTION OF EFFECT SIZES. (<i>EDITORS</i>)	3 1.44%	13 6.25%	84 40.38%	72 34.62%	36 17.31%	208
RESULTS PUBLISHED IN MANAGEMENT RESEARCH REFLECT THE TRUE DISTRIBUTION OF EFFECT SIZES. (<i>AUTHORS</i>)	2 0.24%	60 7.22%	312 37.55%	329 39.59%	128 15.40%	831
THERE IS AN OVERREPRESENTATION OF P-VALUES IN THE TAIL OF THE DISTRIBUTION JUST BELOW ALPHA = 0.1. (<i>EDITORS</i>)	12 5.77%	35 16.83%	90 43.27%	56 26.92%	15 7.21%	208
THERE IS AN OVERREPRESENTATION OF P-VALUES IN THE TAIL OF THE DISTRIBUTION JUST BELOW ALPHA = 0.1. (<i>AUTHORS</i>)	22 2.65%	123 14.80%	377 45.37%	214 25.75%	95 11.43%	831
RESULTS PUBLISHED IN MANAGEMENT RESEARCH ARE GENERALLY REPLICABLE. (<i>EDITORS</i>)	1 0.48%	43 20.67%	66 31.73%	80 38.46%	18 8.65%	208
RESULTS PUBLISHED IN MANAGEMENT RESEARCH ARE GENERALLY REPLICABLE. (<i>AUTHORS</i>)	10 1.20%	145 17.45%	319 38.39%	259 31.17%	98 11.79%	831
FALSE POSITIVE RESULTS IN MANAGEMENT RESEARCH ARE VERY PERSISTENT. (<i>EDITORS</i>)	15 7.21%	43 20.67%	81 38.94%	59 28.37%	10 4.81%	208
FALSE POSITIVE RESULTS IN MANAGEMENT RESEARCH ARE VERY PERSISTENT. (<i>AUTHORS</i>)	27 3.25%	174 20.94%	422 50.78%	163 19.61%	45 5.42%	831

TABLE 5 VIEWS ON REPLICATION STUDIES						
	STRONGLY DISAGREE	DISAGREE	NEITHER	AGREE	STRONGLY AGREE	TOTAL
TO WHICH EXTENT DO YOU AGREE WITH THE FOLLOWING STATEMENTS?						
IT IS IMPORTANT TO REPEAT RESEARCH OTHERS HAVE ALREADY CONDUCTED AND PUBLISHED. (<i>EDITORS</i>)	2	15	42	75	74	208
	0.96%	7.21%	20.19%	36.06%	35.58%	
IT IS IMPORTANT TO REPEAT RESEARCH OTHERS HAVE ALREADY CONDUCTED AND PUBLISHED. (<i>AUTHORS</i>)	14	49	164	302	302	831
	1.68%	5.90%	19.74%	36.34%	36.34%	
WHEN MANAGEMENT RESEARCH IS REPLICATED, EARLY POSITIVE STUDIES OFTEN RECEIVE MORE ATTENTION THAN LATER NEGATIVE ONES. (<i>EDITORS</i>)	2	26	71	71	38	208
	0.96%	12.50%	34.13%	34.13%	18.27%	
WHEN MANAGEMENT RESEARCH IS REPLICATED, EARLY POSITIVE STUDIES OFTEN RECEIVE MORE ATTENTION THAN LATER NEGATIVE ONES. (<i>AUTHORS</i>)	13	60	251	282	225	831
	1.56%	7.22%	30.20%	33.94%	27.08%	
REPLICATION STUDIES LACK ORIGINALITY AND CREATIVITY. (<i>EDITORS</i>)	9	43	63	58	35	208
	4.33%	20.67%	30.29%	27.88%	16.83%	
REPLICATION STUDIES LACK ORIGINALITY AND CREATIVITY. (<i>AUTHORS</i>)	57	169	222	252	131	831
	6.86%	20.34%	26.71%	30.32%	15.76%	
REPLICATION STUDIES CONTRIBUTE IN NEW WAYS TO THE FIELD. (<i>EDITORS</i>)	3	22	56	82	45	208
	1.44%	10.58%	26.92%	39.42%	21.63%	
REPLICATION STUDIES CONTRIBUTE IN NEW WAYS TO THE FIELD. (<i>AUTHORS</i>)	18	82	236	333	162	831
	2.17%	9.87%	28.40%	40.07%	19.49%	
I WOULD ADVISE PhD STUDENTS NOT TO LAUNCH THEIR CAREERS BY CONDUCTING CAREFUL REPLICATIONS OF PUBLISHED RESEARCH. (<i>EDITORS</i>)	56	56	42	42	12	208
	26.92%	26.92%	20.19%	20.19%	5.77%	
I WOULD ADVISE PhD STUDENTS NOT TO LAUNCH THEIR CAREERS BY CONDUCTING CAREFUL REPLICATIONS OF PUBLISHED RESEARCH. (<i>AUTHORS</i>)	245	237	193	101	55	831
	29.48%	28.52%	23.23%	12.15%	6.62%	

TABLE 6						
VIEWS ON ENSURING CREDIBILITY IN EMPIRICAL MANAGEMENT RESEARCH						
	STRONGLY DISAGREE	DISAGREE	NEITHER	AGREE	STRONGLY AGREE	TOTAL
TO ENSURE CREDIBILITY IN EMPIRICAL MANAGEMENT						
... AUTHORS MUST DECIDE THE RULE FOR TERMINATING DATA COLLECTION BEFORE DATA COLLECTION BEGINS AND REPORT THIS RULE IN THE ARTICLE. (<i>EDITORS</i>)	51	53	58	34	12	208
	24.52%	25.48%	27.88%	16.35%	5.77%	
... AUTHORS MUST DECIDE THE RULE FOR TERMINATING DATA COLLECTION BEFORE DATA COLLECTION BEGINS AND REPORT THIS RULE IN THE ARTICLE. (<i>AUTHORS</i>)	158	219	284	107	63	831
	19.01%	26.35%	34.18%	12.88%	7.58%	
... AUTHORS MUST LIST ALL VARIABLES COLLECTED IN A STUDY. (<i>EDITORS</i>)	40	52	52	34	30	208
	19.23%	25.00%	25.00%	16.35%	14.42%	
... AUTHORS MUST LIST ALL VARIABLES COLLECTED IN A STUDY. (<i>AUTHORS</i>)	173	223	173	183	79	831
	20.82%	26.84%	20.82%	22.02%	9.51%	
... REVIEWERS SHOULD BE MORE TOLERANT OF IMPERFECTIONS IN RESULTS. (<i>EDITORS</i>)	53	74	38	30	13	208
	25.48%	35.58%	18.27%	14.42%	6.25%	
... REVIEWERS SHOULD BE MORE TOLERANT OF IMPERFECTIONS IN RESULTS. (<i>AUTHORS</i>)	237	299	156	94	45	831
	28.52%	35.98%	18.77%	11.31%	5.42%	
... JOURNALS SHOULD REQUIRE AUTHORS TO MAKE THEIR ORIGINAL MATERIALS AND DATA PUBLICLY AVAILABLE. (<i>EDITORS</i>)	54	66	44	26	18	208
	25.96%	31.73%	21.15%	12.50%	8.65%	
... JOURNALS SHOULD REQUIRE AUTHORS TO MAKE THEIR ORIGINAL MATERIALS AND DATA PUBLICLY AVAILABLE. (<i>AUTHORS</i>)	177	229	198	147	80	831
	21.30%	27.56%	23.83%	17.69%	9.63%	
... AUTHORS MUST REPORT ALL EXPERIMENTAL CONDITIONS, INCLUDING FAILED MANIPULATION. (<i>EDITORS</i>)	75	80	40	12	1	208
	36.06%	38.46%	19.23%	5.77%	0.48%	
... AUTHORS MUST REPORT ALL EXPERIMENTAL CONDITIONS, INCLUDING FAILED MANIPULATION. (<i>AUTHORS</i>)	295	309	142	62	23	831
	35.50%	37.18%	17.09%	7.46%	2.77%	
... IF OBSERVATIONS ARE ELIMINATED, AUTHORS MUST ALSO REPORT WHAT THE STATISTICAL RESULTS ARE IF THOSE OBSERVATIONS ARE INCLUDED. (<i>EDITORS</i>)	75	81	38	13	1	208
	36.06%	38.94%	18.27%	6.25%	0.48%	
... IF OBSERVATIONS ARE ELIMINATED, AUTHORS MUST ALSO REPORT WHAT THE STATISTICAL RESULTS ARE IF THOSE OBSERVATIONS ARE INCLUDED. (<i>AUTHORS</i>)	272	303	175	66	15	831
	32.73%	36.46%	21.06%	7.94%	1.81%	

Table 7: Regression experience with misconduct and misrepresentation

	(1) Misconduct	(2) Misconduct	(3) Misconduct	(4) Misrepresentation	(5) Misrepresentation	(6) Misrepresentation
Managing Editor			0.213** (0.040)			0.228* (0.078)
Department/Associate Editor			0.196** (0.011)			0.254*** (0.007)
Reviewer for FT 45 journal			0.113** (0.021)			0.222*** (0.002)
No. of Scientific Publications		0.065*** (0.003)	0.047** (0.045)		0.112*** (0.000)	0.093*** (0.004)
No. of FT-45 Publications		0.103*** (0.001)	0.063* (0.068)		0.113** (0.011)	0.037 (0.459)
Academic Position		-0.009 (0.691)	-0.017 (0.455)		0.031 (0.304)	0.018 (0.556)
Age	0.003 (0.131)	0.000 (0.906)	-0.000 (0.901)	-0.003 (0.236)	-0.009*** (0.005)	-0.009*** (0.003)
Holds a PhD	-0.010 (0.917)	-0.122 (0.220)	-0.120 (0.226)	0.136 (0.261)	-0.075 (0.548)	-0.086 (0.486)
Female	-0.051 (0.258)	-0.018 (0.697)	-0.018 (0.686)	-0.061 (0.357)	-0.020 (0.762)	-0.024 (0.712)
US Scholar	-0.048 (0.298)	-0.048 (0.290)	-0.046 (0.302)	-0.025 (0.700)	0.003 (0.960)	0.008 (0.900)
F-Value	1.21	4.10	5.01	0.77	4.75	5.90
P > F-Valued	0.304	0.000	0.000	0.548	0.000	0.000
Observations	938	894	894	938	894	894

* p<0.1, ** p<.05, *** p<.01

Table 8: Regression views on general trustworthiness of empirical research

	(1)	(2)	(3)	(4)	(5)	(6)
		<u>Publishing Incentives</u>			<u>Trustworthiness</u>	
	Fishing for Significance	Large Power of Theories	Overrepresentation of $p < 0.1$	True Effect Sizes	Generally Replicable	Prevalence of False Positives
Managing Editor	-0.061 (0.518)	-0.132 (0.356)	-0.016 (0.895)	-0.132 (0.246)	-0.150 (0.201)	-0.005 (0.972)
Department/Associate Editor	-0.108 (0.129)	-0.096 (0.394)	-0.263*** (0.005)	-0.150 (0.100)	-0.044 (0.629)	-0.174* (0.057)
Reviewer for FT 45 journal	-0.050 (0.355)	-0.154* (0.067)	-0.184** (0.013)	0.097 (0.146)	0.159** (0.030)	0.041 (0.553)
No. of Scientific Publications	-0.001 (0.977)	0.028 (0.445)	-0.038 (0.241)	0.031 (0.266)	0.004 (0.889)	0.020 (0.511)
No. of FT-45 Publications	0.077** (0.025)	-0.055 (0.343)	0.094* (0.069)	0.015 (0.734)	-0.028 (0.593)	0.008 (0.858)
Academic Position	-0.022 (0.396)	0.005 (0.879)	-0.006 (0.812)	-0.050** (0.041)	-0.042 (0.144)	-0.041 (0.143)
Misconduct	-0.008 (0.856)	0.081 (0.300)	0.059 (0.348)	0.204*** (0.000)	0.143** (0.024)	0.198*** (0.004)
Misrepresentation	0.122*** (0.000)	0.142*** (0.005)	0.150*** (0.000)	0.089** (0.024)	0.060 (0.159)	0.098** (0.014)
Age	-0.002 (0.407)	-0.002 (0.550)	-0.006** (0.038)	-0.004 (0.149)	0.002 (0.416)	0.000 (0.943)
Holds a PhD	-0.030 (0.755)	0.271* (0.082)	-0.078 (0.464)	0.138 (0.236)	0.064 (0.622)	0.030 (0.775)
Female	0.069 (0.183)	0.108 (0.168)	-0.070 (0.286)	-0.099* (0.092)	-0.198*** (0.002)	-0.113* (0.063)
US Scholar	0.240*** (0.000)	-0.064 (0.416)	0.011 (0.864)	0.083 (0.157)	-0.184*** (0.004)	0.022 (0.735)
F-Value	5.03	2.83	4.51	6.16	3.81	3.51
P > F-Valued	0.000	0.001	0.000	0.000	0.000	0.000
Observations	893	893	893	893	893	893

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Regression views on value of replication studies

	(1) Importance of Repetition	(2) Less Credit for Replications	(3) Lack of Originality	(4) Replications contribute in new ways	(5) Advise PhDs against Replications
Managing Editor	0.033 (0.767)	-0.272** (0.021)	-0.046 (0.736)	0.013 (0.907)	-0.022 (0.892)
Department/Associate Editor	-0.108 (0.273)	-0.162* (0.090)	-0.027 (0.816)	-0.034 (0.738)	0.075 (0.531)
Reviewer for FT 45 journal	-0.060 (0.426)	-0.098 (0.174)	0.005 (0.957)	-0.065 (0.397)	-0.002 (0.985)
No. of Scientific Publications	0.036 (0.256)	-0.011 (0.732)	0.070* (0.074)	0.051 (0.140)	0.091** (0.026)
No. of FT-45 Publications	0.077 (0.122)	0.015 (0.763)	-0.104* (0.094)	-0.056 (0.281)	-0.158** (0.015)
Academic Position	-0.014 (0.605)	0.002 (0.938)	-0.071* (0.051)	-0.034 (0.275)	0.001 (0.985)
Misconduct	-0.184*** (0.003)	-0.003 (0.963)	-0.060 (0.447)	0.029 (0.663)	0.017 (0.828)
Misrepresentation	0.140*** (0.002)	0.135*** (0.003)	-0.022 (0.688)	0.015 (0.754)	-0.132** (0.017)
Age	-0.001 (0.829)	0.001 (0.833)	0.001 (0.806)	0.002 (0.453)	0.010*** (0.008)
Holds a PhD	-0.108 (0.414)	0.054 (0.668)	0.092 (0.577)	-0.006 (0.963)	0.045 (0.760)
Female	-0.198*** (0.004)	0.078 (0.248)	-0.071 (0.371)	-0.085 (0.215)	-0.070 (0.417)
US Scholar	0.122* (0.075)	0.082 (0.238)	0.093 (0.253)	0.049 (0.466)	-0.411*** (0.000)
F-Value	2.43	2.21	1.26	0.82	5.01
P > F-Valued	0.004	0.010	0.239	0.632	0.000
Observations	893	893	893	893	893

* p<0.1, ** p<.05, *** p<.01

Table 10: Regression views on possible ramifications

	(1) Rules for Termination	(2) List all Variables Collected	(3) More Reviewer Tolerance	(4) Data Availability	(5) Condition Transparency	(6) Outlier Transparency
Managing Editor	-0.114 (0.425)	0.052 (0.733)	0.195 (0.184)	-0.273* (0.065)	-0.044 (0.678)	0.148 (0.211)
Department/Associate Editor	-0.038 (0.752)	0.140 (0.279)	-0.046 (0.674)	-0.039 (0.751)	0.125 (0.187)	-0.095 (0.298)
Reviewer for FT 45 journal	0.037 (0.684)	0.188* (0.054)	-0.299*** (0.000)	-0.019 (0.842)	0.040 (0.601)	-0.030 (0.688)
No. of Scientific Publications	-0.032 (0.411)	-0.010 (0.810)	-0.001 (0.980)	0.014 (0.741)	-0.021 (0.531)	0.030 (0.369)
No. of FT-45 Publications	0.198*** (0.002)	0.094 (0.158)	0.055 (0.346)	-0.014 (0.832)	0.046 (0.360)	0.061 (0.240)
Academic Position	0.034 (0.359)	0.016 (0.685)	-0.014 (0.713)	0.084** (0.020)	0.013 (0.633)	0.053* (0.058)
Misconduct	-0.169** (0.036)	-0.238*** (0.004)	-0.009 (0.900)	0.024 (0.772)	-0.081 (0.141)	-0.080 (0.139)
Misrepresentation	0.017 (0.744)	0.007 (0.914)	-0.083 (0.113)	-0.115** (0.038)	-0.114*** (0.008)	-0.067 (0.106)
Age	-0.003 (0.331)	-0.005 (0.246)	0.024*** (0.000)	-0.010*** (0.006)	-0.008*** (0.005)	-0.009*** (0.003)
Holds a PhD	0.251 (0.127)	-0.063 (0.718)	-0.159 (0.318)	0.004 (0.982)	0.139 (0.305)	0.113 (0.350)
Female	0.030 (0.708)	0.097 (0.284)	-0.039 (0.609)	0.446*** (0.000)	0.195*** (0.008)	0.161** (0.022)
US Scholar	-0.047 (0.564)	0.259*** (0.004)	-0.315*** (0.000)	0.229** (0.010)	0.243*** (0.001)	0.147** (0.031)
F-Value	2.40	2.80	6.77	4.93	4.32	3.06
P > F-Valued	0.005	0.001	0.000	0.000	0.000	0.000
Observations	893	893	893	893	893	893

* p<0.1, ** p<.05, *** p<.01

¹ We also searched the Web of Science for articles that involve the term “replication” in the title and found only 125 articles in journals included in the Financial Times 45 journal list.

² See <http://onlinelibrary.wiley.com/doi/10.1002/smj.2016.37.issue-11/issuetoc>

³ Not all replication studies in the SMJ special issue draw similar findings as the original work. Tsang and Yamanoi (2016), for example, point out inconsistencies in hypothesis development along with a lack of generalizability based on a sample from Barkema and Vermeulen’s (1998) study. Park, Borah, and Kotha (2016) fail to replicate the major findings of three studies they were intending to replicate.

⁴ The questionnaire design and analysis was carried out outside of the US, as such no university institutional review board has been involved in the oversight of this research. The American based co-author was not involved in consenting, data collection, or had access to identifiable data.

⁵ To reflect on the other side of the process, we also asked journal editors whether manuscripts *should* contain statistically significant results. This question is based on the work of Devaney (2001). Of the 191 respondents to this question, 131 responded that yes, manuscripts should, indeed contain significant results. Which in part, may relate to the perception of reviewers here.

⁶ We also had a question for those with editorial responsibilities only, that (following Devaney (2001)) asked whether replication studies were appropriate for publication. An overwhelming majority of management editors, 84 percent, responded that these types of studies were appropriate.

⁷ Respondents were given the link to the journal list to corroborate journal inclusion in the corresponding list.

⁸ When it comes to differences across the groups, three out of four editors report at least one FT-45 publication, with one third of editors having more than five FT-45 publications. For those in non-editor roles, more than 40 percent report zero FT-45 publications. Among those reviewing for FT-45 journals 82 percent report at least one FT-45 publication; 20% have more than five FT-45 publications, while two out of three of those not reviewing for FT-45 journals report zero FT-45 publications.