Questions about questionable research practices in the field of management:

A guest commentary

George C. Banks
University of North Carolina at Charlotte

Ernest H. O’Boyle Jr.  Jeff H. Pollack
University of Iowa  North Carolina State University
Charles D. White  John H. Batchelor
Longwood University  University of West Florida
Christopher E. Whelpley  Kristie Abston
Virginia Commonwealth University  University of West Florida
Andrew A. Bennett  Cheryl L. Adkins
University of Alabama  Longwood University

Editor’s Note: This paper was originally submitted as a regular submission, where the JOM editorial team thought the overall theme and set of issues were better addressed Commentary piece to spur dialog concerning ethical research practices in our field. We are pleased that the authors have included the full write-up of the five studies that inform this work (downloadable as Supplementary Material; please have a look). As with all journals' editorial policies, JOM’s own editorial policy on data transparency, reporting, and other practices discussed in this piece continues to evolve. We have signed on to the Editor’s Code of Ethics (https://editorethics.uncc.edu/), are members of the Committee on Publication Ethics (COPE, http://publicationethics.org/), have recently made changes to our review policy in ways that increase reviewer accountability to professional standards, and are considering adopting additional practices and/or partnering with other groups that focus on developing high standards for science and ethics. The current commentary is not an officially reflection of JOM’s official policies; instead, it is intended to get us all thinking broadly about how we can increase the quality and integrity of our research, together. We will continue to update our readership on ways that JOM is addressing ethical issues in the conduct of research.

Author Note: Supplemental material for this article is available at http://xxx.sagepub.com/supplemental. Suggestions on earlier versions of this work by Sarai Blincoe, Ken Brown, Sven Kepes, Michael McDaniel, In-Sue Oh, and Steven Rogelberg were valuable in the improvement of this manuscript. We wish to thank Fred Oswald and five anonymous reviewers for their developmental comments as well as the participants in this research for their candor and constructive thoughts. The correspondence regarding this manuscript should be sent to George C. Banks, University of North Carolina at Charlotte, Belk College of Business, 9201 University City Blvd, Charlotte, NC, 28223. Electronic mail may be sent to gcbanks@gmail.com
Abstract

The discussion regarding questionable research practices (QRPs) in management as well as the broader natural and social sciences has increased substantially in recent years. Despite the attention, questions remain regarding research norms and the implications for both theoretical and practical advancements. The aim of the current article is to address these issues in a question-and-answer format while drawing upon both past research and the results of a series of new studies conducted using a mixed-methods design. Our goal is to encourage a systematic, collegial, and constructive dialogue regarding QRPs in management research.

Keywords: Questionable research practices (QRPs); Research methodology; Philosophy of science
Questions about questionable research practices in management: A guest commentary

A recent spate of high profile scandals in the social sciences, some of which occurred within management, has shaken the confidence of those within and outside of the profession leading to calls for greater transparency and oversight into the research process (Finkel, Eastwick, & Reis, 2015; Matlack, 2013; Schmidt & Hunter, 2015). Adding to concerns about the veracity of management research is evidence that our field and closely related fields may be susceptible to questionable research practices (QRPs; Bosco et al., in press; Francis, Tanzman, & Matthews, 2014; Franco, Malhotra, & Simonovits, 2014; O'Boyle, Banks, & Gonzalez-Mule, in press; Open Science Collaboration, 2015).

QRPs operate in the ambiguous space between what one might consider best practices and academic misconduct. Some examples of QRPs can include presenting post hoc findings as a priori, “cherry picking” fit indices, and selectively deleting outliers for the purpose of achieving statistical significance. The occurrence of these practices is not always questionable; in fact, some of these approaches are beneficial to management research under the right circumstances. For instance, exploratory data analysis has led to numerous discoveries in both the physical and social sciences (for a review see Locke, 2007). Further, certain fit indices are objectively better than others, and outliers should be examined and, at times, dropped from further analysis (Aguinis, Gottfredson, & Joo, 2013). However, those who see QRPs as a problem point to when the practices are either misreported or not reported rather than the practices themselves (Simmons, Nelson, & Simonsohn, 2011).

The purpose of this commentary is to highlight the wide-ranging perspectives about QRPs in terms of their possible causes, their prevalence, and journal policies that may prevent them. As we review the literature on QRPs, we provide the results from five new studies. Details
on the methods and results of these studies are presented in online Appendices A-F (see online supplementary materials). In short, these studies examined engagement in QRPs (Study 1; \( n = 405 \) active management researchers), motivations for engaging in and perceptions of the appropriateness of QRPs (Study 2; \( n = 344 \) additional active management researchers), current graduate training practices (Study 3; \( n = 126 \) doctoral students; \( n = 29 \) program coordinators), journal guidelines involving QRPs (Study 4; \( n = 160 \) management journals), and QRP engagement relative to other fields (Study 5; \( n = 201 \) researchers in supply chain management and sociology). In this paper we integrate the results of these studies and use them to make recommendations. We refer the reader to supplementary readings in several areas of this commentary as space constraints limit the comprehensiveness of discussions.

**Background on QRPs**

The first mention of “questionable research practices” was in the 1958 Code of Professional Ethics and Practices of Public Opinion Researchers. They called upon members to “subscribe to standards imposed everywhere by men of science…never to knowingly indulge in questionable research practices in order to arrive at some predetermined result or to ‘prove’ a case” (Riley, 1958, p. 218). Recent writing characterizes QRPs as design, analytic, or reporting practices that have been *questioned* because of the potential for the practice to be employed with the purpose of presenting biased evidence in favor of an assertion (Bakker, van Dijk, & Wicherts, 2012; Finkel et al., 2015; John, Loewenstein, & Prelec, 2012; Schmidt & Hunter, 2015; Simmons et al., 2011). Some have also suggested that QRPs be redefined as “Questionable Reporting Practices” (Fanelli, 2013; Sijtsma, Veldkamp, & Wicherts, 2015; Wigboldus & Dotsch, 2015) to emphasize that the issue is one of transparency. Table 1 presents QRPs that have received the most attention. Despite past efforts to categorize QRPs by level of severity
QUESTIONABLE RESEARCH PRACTICES

(e.g., Bedeian, Taylor, & Miller, 2010), the rates at which researchers engage in certain QRPs may ultimately make one type of QRP more detrimental than another. For instance, if falsifying data only occurs in very rare circumstances, the problem of the biased reporting of outcomes may be much more detrimental. Ultimately, the extent to which QRPs cause harm is dependent on how the QRPs are used and how the field responds to certain findings.

(Insert Table 1 about here)

To examine agreement in management regarding these QRPs, we asked 344 active management researchers about the appropriateness of specific practices (Study 2). For example, when asked about “rounding down” a significance test (e.g., a p-value of .054 is reported as p <.05), 13% (95% CI [9%, 17%]) of management researchers disagreed or strongly disagreed that this was inappropriate. Approximately 23% (95% CI [19%, 27%]) of management researchers disagreed or strongly disagreed that selectively reporting hypotheses based on whether or not they were statistically significant was an inappropriate practice. Additionally, 21% (95% CI [17%, 25%]) of management researchers disagreed or strongly disagreed that excluding data after looking at the impact of doing so on the results was inappropriate. Similar findings were reported for hypothesizing after results are known (HARKing; Kerr, 1998; 25% disagreed or strongly disagreed the practice was inappropriate; 95% CI [20%, 30%]) and selective inclusion or exclusion of control variables to achieve statistical significance (21% disagreed or strongly disagreed; 95% CI [17%, 25%]). In sum, and consistent with past discussions in related literatures on responsible conduct of research, social norms, mentoring, and organizational justice in science (Anderson, Martinson, & De Vries, 2007; Anderson, Ronning, De Vries, & Martinson, 2007; De Vries, Anderson, & Martinson, 2006; Martinson, Anderson, Crain, & De Vries, 2006), we found a lack of consensus among researchers regarding the appropriateness of
QRPs. In the following section, we propose a series of questions about QRPs, and present evidence from the existing literature as well as the findings from our series of empirical studies on QRPs.

**Reviewing the evidence on QRPs**

**Question 1**: What causes QRPs?

Researchers have suggested several root causes for QRPs (for reviews see Bakker et al., 2012; Finkel et al., 2015; John et al., 2012; Simmons et al., 2011). First, it has been suggested that an overemphasis on theory may result in QRPs. For example, Locke (2007) writes:

> Everyone who publishes in professional journals in the social sciences knows that you are supposed to start your article with a theory, then make deductions from it, then test it, and then revise the theory...In practice, however, I believe that this policy encourages—in fact demands, premature theorizing and often leads to making up hypotheses after the fact—which is contrary to the intent of hypothetico–deductive method (p. 867).

Locke’s (2007) assertion is that a rigid adherence to deductive research may lead to engagement in QRPs as it artificially narrows the routes to scientific discovery to not only a priori theory, but a priori theory that is supported. In this way, interesting findings that lack supporting theory (Hambrick, 2007) or are in contrast to “strong theory” may be either suppressed or manipulated to achieve something more consistent with reviewer and reader expectations (Harrison et al., in press).

A second cause for QRPs may be the relative emphasis placed on confirmatory analysis over exploratory analysis (for a more in-depth review please see Wagenmakers et al., 2012; Wagenmakers, Wetzels, Borsboom, & Van Der Maas, 2011; Wigboldus & Dotsch, 2015). For example, confirmatory research often relies on strongly established cutoff points (e.g., $p < .05$; RMSEA < .08 and CFI > .95) that could encourage some QRPs. Yet, exploratory research may also be susceptible to QRPs as researchers may conduct exploratory analyses of their data, but
report post hoc findings as a priori because of perceptions that confirmatory analyses are superior to exploratory (Kerr, 1998). This sentiment was echoed by an active management researcher we surveyed:

When you consider the style of writing that is expected by JAP [Journal of Applied Psychology], AMJ [Academy of Management Journal], JPSP [Journal of Personality and Social Psychology], and the like then you basically MUST do some of these things. When [have you] last seen a study published that was quantitative and exploratory?

A third and related potential cause of QRPs could be the enormous pressure to publish (Banks & O'Boyle, 2013; Schmidt & Hunter, 2015). With increasing emphasis on journal lists by research-intensive universities, scholars may feel the need to publish to satisfy tenure and accreditation demands as well as to pursue promotion and raises (Bedeian et al., 2010; O'Boyle et al., in press). For instance, one PhD student surveyed wrote,

It's currently impossible to publish anything in a top journal …with too many unsupported hypotheses. I don't think that is a good thing, but I do think jobs and tenure are good things, as I currently have neither.

Another participant wrote “It would be impossible for most researchers to make tenure at any research intensive institution if they were to strictly follow these high standards of research.”

If QRPs are simply a pragmatic method of achieving early success in publishing, one would expect that doctoral students and assistant professors are the primary source of QRPs. Similarly, QRPs should primarily be nested in universities with greater research intensity. Yet, in the current research, when untenured professors were compared to tenured professors, there were no significant differences in QRP engagement ($t (447) = 1.44, p > .05; 95\% \text{ CI } [-.06, .36];$ Cohen’s $d = .12$), nor was there a notable difference in QRP engagement between research-intensive universities and non-research-intensive schools ($t (639) = -.09, p > .05; 95\% \text{ CI } [-.22, .20];$ Cohen’s $d = .01$). If engagement in QRPs is learned during graduate training and pressure to use such practices is encouraged in assistant professors, perhaps the use of such practices
could become a habit. Alternatively, as Schmidt and Hunter (2015) suggested, other incentives, such as promotions, raises, and legacy may motivate QRPs throughout one’s career.

**Question 2:** How prevalent are QRPs?

Any attempt to uncover the true base rates of QRPs is obfuscated by the very practices that make them questionable (i.e., lack of reporting or misreporting). However, recent research in management has started to investigate QRP frequency. For instance, O’Boyle et al. (in press) argued that QRPs were the cause for the ratio of supported to unsupported hypotheses more than doubling from defended dissertation to journal publication. Other studies in management have found evidence for QRPs by examining changes from dissertation to published articles (Mazzola & Deuling, 2013), inspecting correlation matrices in published articles (Bosco et al., in press), and probing the accuracy of $p$-value reporting in multiple moderated regression analyses (O’Boyle et al., 2015). Using objective measures and observer ratings of engagement in QRPs, these studies provided evidence of potential engagement in practices, such as HARKing, suppressing null results, and misreporting $p$-values.

In our surveys of active management researchers (Studies 1 and 2), less than 1% (95% CI [0%, 1%]) reported that they falsified data and 11% (95% CI [9%, 13%]) said that they had “rounded off” $p$-values. These are relatively low base rates—possibly encouraging to those who believe QRPs cause harm. However, about one-third of the researchers stated that for the express purpose of supporting hypotheses with statistical significance, they engaged in post hoc exclusion of data (29%; 95% CI [26%, 32%]) or selectively included control variables (33%; 95% CI [30%, 36%]). Similarly, 50% of management researchers responded that they selectively reported hypotheses based on statistical significance (95% CI [46%, 54%]) and presented a post hoc hypothesis as if it were developed a priori (95% CI [46%, 54%]). Many of the surveyed
researchers provided additional insight regarding observations of their own and others’
engagement in QRPs. For instance, one wrote:

> I’ve seen top researchers (i.e., SIOP [Society for Industrial and Organizational Psychology] and AOM [Academy of Management] fellows) at top universities (e.g., [redacted]) conduct clear fishing and report only the significant results (about 5% of the runs) at top journals in our field.

A final point related to QRP prevalence is that it has been suggested that those fields with highly developed theoretical paradigms and a lengthy history may have higher rates of QRPs than those fields with a lower level of theory development (Harrison et al., in press). However, we did not observe differences between various sub-fields, such as OBHR and strategy ($t(422) = 1.77, p > .05$; 95% CI [-.04, .67]; Cohen’s $d = .25$). In other words, despite the use of different methods and analytic approaches, we did not see a difference between these management sub-groups.

**Question 3:** When are QRPs learned?

Some researchers (e.g., Bedeian et al., 2010) have suggested that QRPs are learned behaviors that begin in graduate school. Consistent with this sentiment, one study participant in the current work wrote:

> The moment I started working with PhD students … these practices started to surface. They magnified during my training as a PhD, where assistant professors would provide me with a dataset to mine and find ‘something significant’ and, based on it, propose ‘interesting hypotheses’.

If QRPs are simply learned behaviors in doctoral training, then interventions into QRP reduction can target doctoral programs. We surveyed management doctoral students as well as doctoral program coordinators (see Study 3). In the survey, we were interested in both the formal training and the observed behaviors in the program. Overall, 12% (95% CI [6%, 17%]) of PhD students stated that they had observed the practice of “rounding off” $p$-values and 47% (95% CI
[38%, 55%]) said that they were instructed to avoid the practice. Concerning the selective reporting of hypotheses, 55% (95% CI [46%, 63%]) of doctoral students stated they had observed the selective reporting of results and 61% (95% CI [52%, 69%]) said they had been instructed to avoid the practice. Regarding excluding data post hoc, 33% (95% CI [25%, 41%]) said that they had observed the practice and 60% (95% CI [52%, 68%]) were instructed to avoid the practice. Further, 58% (95% CI [49%, 67%]) of students indicated that they had observed HARKing and 71% (95% CI [63%, 79%]) had been instructed to avoid it. For the selective inclusion of control variables to turn a null result into a significant one, 45% (95% CI [36%, 54%]) of PhD students observed the practice and 52% (95% CI [43%, 61%]) were instructed to avoid it. Finally, 7% (95% CI [3%, 11%]) stated that they had observed data falsification while 84% (95% CI [78%, 90%]) were instructed to avoid it. The students in this study provided useful insight into their experiences. For instance, one PhD student wrote, “They can teach the ethical practices, but the underlying message is to write something that will get published, as they like to say in my department, it depends how well you ‘tell the story’.”

In terms of the PhD program coordinators surveyed, roughly half of the responses noted that QRPs represent a “serious” issue, yet some responses claimed that too much is being made of the issue:

As long as there is replication, I don't see this as being an issue.

I think we may have gone too far now in the way we approach these concerns, ‘witch hunting’ is what I see a lot of as an AE [Associate Editor]. I don't think the concerns are invalid, it's the approach that concerns me.

Thus, there was little consensus regarding the degree to which the field of management should be concerned about QRPs in doctoral training. In terms of preparation, with respect to the design and execution of research, responses illustrated that two mechanisms are currently
QUESTIONABLE RESEARCH PRACTICES

integrated within the PhD training process. First, the majority of the responses stressed that mentorship (and the advisor-student relationship) is the key mode of training in ethics. In addition, roughly half of the respondents noted that ethics (in research) are covered in classes, brown bags, and special seminars. With regard to how the reporting of results and the publication process are taught, coordinators reported that “mentoring with faculty on research projects” as well as informal “feedback sessions with colleagues and senior staff” were the primary mechanisms for education.

**Question 4:** Do researchers perceive that the publication process encourages QRPs?

Bedeian (2003) surveyed researchers who had published in *Academy of Management Journal* and 25% reported that they were encouraged by editors/reviewers to make changes they knew to be incorrect. Further, experimental research has suggested that reviewers tend to perceive studies with results that achieve statistical significance to have greater relevance and methodological rigor (Emerson et al., 2010). In the literature on publication bias (for in-depth reviews see Banks, Kepes, & McDaniel, 2015; Schmidt & Hunter, 2015), there are some data to suggest that authors are the primary cause of this bias as null results are likely never written up and submitted (see "Myth #2: The Editorial Review Process Is the Primary Cause of Publication Bias" in Banks et al., 2015). Yet, there is little evidence that highlights the extent to which editors and reviewers demand the use of QRPs or the biased reporting of outcomes. To assess this, we asked active management researchers if they had ever been instructed to engage in QRPs by a reviewer or editor. No researchers reported being asked to falsify data and only 1% (95% CI [0%, 11%]) reported being asked to “round off” a p-value. However, 40% (95% CI [32%, 48%]) had been asked to selectively report hypotheses, 10% (95% CI [0%, 20%]) had been asked to exclude data post hoc, 33% (95% CI [25%, 41%]) had been asked to engage in HARKing, and
14% (95% CI [5%, 23%]) had been asked to selectively include control variables. Open-ended responses provided additional clarification. One study participant stated, “In moving through the review process (at a very prestigious journal), the reviewers recommend[ed] that we re-write the hypothesis and include our post hoc analyses as though they were intended as part of our a priori analysis.” Another researcher wrote:

> If the first story does not appeal, you rewrite it to appease the editors/reviewers. This is a key to getting one's work accepted. Anyone who pretends otherwise is a fool and probably unpublished. Sorry to burst your bubble. This is not science, it is the art of persuasion.

In general, the active management researchers in this study reported that at least some editors and reviewers encourage engagement in QRPs. The notion that reviewers are intolerant of imperfect results (Simmons et al., 2011) and that journals place great emphasis on novel, positive results in pursuit of higher impact factor ratings at the expense of promoting truth is a perspective held by some (Nosek, Spies, & Motyl, 2012). However, this view may create a false dichotomy between those involved in the research process and those involved in the publication process. With rare exception, reviewers and editors are active researchers who are submitting manuscripts to peer-review journals and are likely subjected to the same pressures conducive to QRPs. Nevertheless, the role of gatekeeper may lead reviewers and editors to rationalize QRPs in the hopes of improving the visibility and impact of their journals. For example, an associate editor surveyed wrote:

> Reviewers regularly instruct authors to change post hoc inductive analyses into hypotheses, to avoid unsupported hypotheses, etc. and editors rarely instruct authors against these practices in decision letters. From a pragmatic standpoint, since the pressure to publish is so high, if the keepers of the publishing/research system are asking for science that is not conducted via the scientific method, then I understand why authors would engage in these practices.
While many of the respondents we surveyed reflected this perspective, we offer an optimistic view. Many editors in management have signed the Editor Ethics 2.0 Code (https://editorethics.uncc.edu/). Efforts such as this are meant to prompt editors to work with authors to reduce the QRPs highlighted above. Further, editors are encouraged to reduce other behaviors, such as promoting self-citations and creating citation cartels with other journals. We see these and other related efforts to be encouraging. For instance, in addition to providing submitting authors with editorial policies, journals may often require authors to affirm that they have acted consistent with certain ethical standards. Here are examples of useful affirmations by some of the leading journals:

*Journal of Management* provides the following affirmation:

Confirm that the manuscript has been submitted solely to this journal and is not published, in press, or submitted elsewhere.

Confirm that all aspects of this research has been carried out in accordance to the ethical standards of the applicable discipline (e.g., management, psychology), as well as the legal/ethical standards of the country where both the data were collected and the research carried out.

*Academy of Management Journal* provides the following affirmation:

Confirm that all the research meets the ethical guidelines, including adherence to the legal requirements of the study country and obtaining human subjects permission where appropriate. Please see the Academy’s [Code of Ethics](#) for more information.

*Journal of Applied Psychology* provides the following affirmation:

I agree to comply with APA Ethics Code Standard 8.14a, Sharing Research Data for Verification, allowing other qualified professionals to confirm the analyses and results should my manuscript be accepted for publication. As suggested by APA guidelines, I will retain the raw data for a minimum of 5 years after the publication of this research.

**Question 5:** What are journal policies pertaining to QRPs?

The evidence reviewed to this point suggests there are systematic issues leading to engagement in QRPs. Hence, an additional study was designed to explore the extent to which
explicitly stated journal policies, in the form of guidance to submitting authors or questions and affirmations during the submission process, addressed the issue of QRPs and encouraged ideal research practices (see Study 4).

Of the 160 management journals we considered, 32% explicitly stated some form of research ethics policy. These policies ranged from the treatment of human subjects to disclosures of potential conflicts of interest. Specifically, 5% discouraged some form of coercive citation practices (e.g., spurious citations meant to improve a journal’s impact factor). Further, 17% included policies on data overlap (e.g., the use of the same data in multiple journal submissions), and 23% mentioned plagiarism checks. Roughly 15% of journals included policies about the inclusion of supplemental materials (e.g., additional tables or figures), and 13% mentioned policies about data sharing or data retention (e.g., the APA requirement for sharing data with other professionals for the purpose of confirming analyses). When considering journal policies more specific to engagement in QRPs, journals did not explicitly advise against presenting post hoc hypotheses as a priori. Further, journals did not mention a policy about publishing null results or a policy about reporting all data excluded from an analysis (e.g., outliers). It is important to note that 64% of journals held a Committee on Publication Ethics (COPE) membership and that 27% of management editors signed the Editors Ethics 2.0 Code policies (https://editorethics.uncc.edu/). The Editor Ethics 2.0 Code explicitly states that editors will take action to discourage HARKing and suppressing null results and, in general, will strive to promote transparent reporting. The results illustrate the degree to which management journals have various forms of ethics policy.

**Question 6:** Does journal prestige influence QRP engagement?
It has been suggested that the motivation for engaging in QRPs has been to increase one’s chance of publishing in prestigious outlets (O’Boyle et al., in press). Past research in a review of major journals has provided evidence that journals with higher impact factors tend to have a higher level of retractions owing in part to QRPs (Fang & Casadevall, 2011). Other investigations into the misreporting of p-values and corresponding journal impact factors provides additional inconclusive evidence (Bakker & Wicherts, 2011). In management, empirical evidence has illustrated that engagement in QRPs tends to be correlated with the impact factor of journals in which the research is published (O’Boyle et al., in press: \( r = .25 \)). Yet, one could argue that QRPs may be used more frequently in lower impact journals where the methodological rigor of studies is potentially weaker.

To further examine this question, we compared authors who had published in higher impact factor journals (see Appendix A, Study 1) with those that had published in lower impact factor journals (see Appendix B, Study 2). For Study 1, anonymous surveys were sent to authors of articles in 10 journals. For Study 2, similar surveys were sent to authors of articles in another 10 journals with lower impact factors. The findings showed that total engagement in QRPs was likely to be higher for those researchers who published in Study 1 compared to those in Study 2 \( (t(739) = 2.10, p = .04; 95\% \text{ CI} [0.13, 0.40]; \text{Cohen’s } d = .15) \). More specifically, the findings showed that participants in Study 1 were more likely to selectively report findings. However, there was no difference for the other specific QRPs. Anecdotal evidence provided by participants in the open-ended response section provided some support. One researcher wrote:

HARKING is a requirement of our journals, and the more prestigious the more it is emphasized. You MUST MUST MUST MUST invent some theory to explain your findings, and inductive research won't be published.
In sum, there does seem to be evidence that QRPs are used more frequently in articles published in more prestigious journals. While this finding is consistent with past research (e.g., Fan & Casadevall, 2011; O’Boyle et al., in press), we do acknowledge the limitation that, in Study 2, participants were also asked whether or not they believed their actions to be appropriate. It is not clear whether the introduction of this additional question may have influenced participant responses.

**Question 7**: How does engagement in QRPs in the field of management compare to other academic areas in business and the broader social sciences?

Past reviews have suggested that there may be a greater potential for engagement in QRPs in behavioral science where small sample laboratory experiments are used more frequently (John et al., 2012; Schmidt & Hunter, 2015). A meta-analytic study by Fanelli (2009) suggested that approximately 2% of scientists across fields have either falsified or modified data and that 34% have employed other QRPs. Comparably, less than 2% of psychologists stated that they had falsified data and approximately a third specified that they had presented post hoc hypotheses as if they had been developed a priori (John et al., 2012). John et al. also estimated that approximately 22% of psychologists reported that they had “rounded off” p-values, and 40% excluded data post hoc. When comparing these results to the findings from management in the current study, in general, there seems to be a reasonable degree of overlap. We also compared the field of management to two related social science fields: supply chain and sociology (see Study 5). In general, the evidence suggests that supply chain and sociology are similar to management research in some ways. However, management researchers did report engaging in QRPs to a greater extent. Further, more management researchers, than supply chain and
sociology researchers, reported that they were encouraged by reviewers to inappropriately round $p$-values, engage in HARKing, and selectively report hypotheses.

**Recommendations for promoting transparency, openness, and reproducibility**

QRPs are suboptimal practices that appear to occur at a non-ideal frequency in management research. Moving from a descriptive to prescriptive approach, we close with a discussion of recommendations meant to promote transparency, openness, and reproducibility. These recommendations are focused not only on discouraging QRPs but also on promoting greater sharing and collaboration within our field (for review papers on these issues see Fang & Casadevall, 2015; Nosek & Bar-Anan, 2012; Nosek et al., 2012).

**Recommendations for management journals**

As one study participant wrote, “authors should be encourage[d] to be honest and open about all the natural quirks and iterative steps taken as part of the complex endeavor of research.” We agree that much of what makes QRPs questionable is the misreporting or lack of reporting of certain behaviors and that journals can play a key role in any reform efforts. A PhD student suggested that “change must occur top-down. Get journals to publish articles containing both significant and insignificant findings, get tenure to be less about ‘5 articles in A journals,’ and the appropriate behaviors will follow.” As such, many past recommendations directly deal with standards of reporting in journals that promote transparency, openness, and reproducibility (Finkel et al., 2015; Kepes & McDaniel, 2013; Nosek & Bar-Anan, 2012; Nosek et al., 2012).

This topic was the point of emphasis in a recent article published in *Science* (Nosek, Alter, Banks, Borsboom, Bowman et al., in press) that presents a series of recommendations for the policies and practices of journals. The eight standards highlighted pertained to (1) citation standards, (2) data transparency, (3) analytic methods (code) transparency, (4) research materials
transparency, (5) design and analysis transparency, (6) preregistration of studies, (7) preregistration of analysis plans, and (8) replication. It is important to note that the guidelines have the advantage that there are four levels for each standard to facilitate flexibility and customization by journals. Level 1 is intended to require little to no effort for adoption. Consider data transparency as an example. To date, many efforts by journals to encourage data sharing have been ineffective (Alsheikh-Ali, Qureshi, Al-Mallah, & Ioannidis, 2011; Eich, 2014; Wicherts, Borsboom, Kats, & Molenaar, 2006). At a basic level, journals might simply request that all authors state in their manuscripts whether the data from the study are available for dissemination. If the data cannot be shared, authors could provide a reason (e.g., proprietary rights; confidentially). Approximately 13% of management journals have an explicitly stated policy about data sharing or retention (see Study 4). Adoption of Level 1 would mean that 100% of journals would have a stated policy while not dictating the terms of that policy.

Level 2 is associated with higher standards. In the context of data sharing, a journal might switch to a default standard which requires data sharing where data are placed in an online repository. A one to three year embargo may be included as a temporary restriction which would allow the author(s) an additional opportunity to use the data before the data became publicly available. Exceptions would still be allowed through a formal appeal to the editor at the point of initial submission. Level 3 is the strongest standard, but is also the most difficult to implement and would not always be applicable. In the context of data sharing, a journal might require that data be deposited in a trusted repository and an independent party might reproduce the analyses with the data prior to publication. A fourth level, Level 0, was provided as a point of comparison for journal policies that do not address the transparency standards advocated in the guidelines. The link to the full standards and list of journal signatories committed to considering the
guidelines, including many signatories from the field of management, is provided for interested readers (http://centerforopenscience.org/top/).

**Recommendations for management graduate training**

One participant in our studies wrote, “I generally avoid using significance testing if I can at all help it (the associated perils of NHST [null hypothesis significance testing] were stressed throughout my graduate training). It was interesting how easy it made it to check ‘never’ to many of the questions you posed.” This comment illustrates the importance of graduate training. Scholars have consistently called for an increase in teaching professional ethics in management, often focused at the graduate level (e.g., Wiese & Fullick, 2014). At this level, we contend that professional ethics should extend beyond legal ethics and protection of human subjects to include discussions about research design, analysis, and reporting ethics. Explicitly incorporating a discussion of QRP s into a doctoral course could be accomplished when teaching topics like research design, data analysis, and IRB training. We know that training individuals in QRP awareness leads to having improved ethical reasoning skills and enhanced awareness of ethical issues after formal training, however, this improvement may not last over time. Consequently, QRP s should also be an explicit part of the conversation among experienced scholars. At the university level, colleagues can informally engage others, or departments can formally discuss this at seminars or meetings. At the professional level, seminars, themes at conferences or meetings, and presidential addresses (e.g., Cortina, 2015) can help enhance this topic’s visibility.

**Recommendations for the role of individual management researchers**

It is important to note that individual researchers do not have to wait for journal policies to change to start promoting open scientific practices. For example, when conducting the latter series of studies for the current work, the studies, materials, and hypotheses were preregistered.
The following steps were taken: (1) After approval from the IRB, but prior to data collection, the information for the studies was pre-registered via the Open Science Framework operated by the Center for Open Science (https://cos.io/); (2) An anonymized link was created that provides a time stamped “frozen picture” of the studies’ information (https://osf.io/5thpc/?view_only=41843b91eea145ab939097c0f5ae1409); (3) The pre-registration included a priori hypotheses, supplemental research questions, the exact surveys to be used, as well as the design of the study. Pre-registration by no means limits the exploration of one’s data, nor does pre-registration prevent researchers from deviating from the original plan. If changes are made, these changes are documented and can simply be explained.

Individual researchers should also consider the approaches they take to collaboration. There are growing concerns about the issue of “gift authorship” or “hyperauthorship” in the natural and social sciences in which one might wonder whether all authors really provided an intellectual contribution. For instance, there was a recent paper published in physics with 5,154 authors (Hotz, 2015). Whether or not “gift authorship” is an issue in management is unclear. There are other potential QRPs pertaining to author collaboration or lack thereof. For instance, some have argued that researchers work too much in isolation leading to a disconnected literature. Others have also argued that dividing research efforts into multiple publications may also present an ethical problem (Elsevier, 2015). In this paper, we believe that the important contribution of this research would not have been possible without a large scale collaboration. We encourage future research to continue to explore the potential of large scale collaboration while simultaneously respecting APA guidelines regarding intellectual contributions.

Finally, an anonymous reviewer asked if QRPs would be less of a concern if researchers abandoned NHST for non-frequentist approaches, such as Bayesian analysis (Zyphur & Oswald,
Our answer is “it depends.” QRP$s can occur regardless of the analytic framework used. For example, a priori specifications must be made in hypothesis testing whether one uses frequentist or Bayesian approaches (Andraszewicz et al., 2014). If researchers explore their data and then test hypotheses using the same data, they deviate from both the frequentist and non-frequentist perspective. Currently, the heavy focus on $p$-values when considering QRP$s is “just the tip of the iceberg” (Leek & Peng, 2015, p. 612), and more work needs to be completed concerning how different analytical approaches may influence QRP$s.

**Conclusion**

We reviewed common questions regarding QRP$s in the field of management to help researchers better understand the means for improving general research practices. The collected evidence illustrates engagement in QRP$s. The results also shed light on the role of editors, reviewers, graduate training, and journal policies in addressing QRP$s. Further, we made recommendations for management journals, graduate training, and individual management researchers to reduce QRPs. Throughout the manuscript, and indirectly through our bibliography, we directed readers to additional resources on the topic. We wish to end with the following quote from one study participant:

> We could make the practice of research better if we aimed to define and educate our community about when each practice is or is not reasonable to use. In some cases, it is not questionable but clearly unacceptable -- whereas in other cases, it may be a reasonable interpretation because of the separation between the data generation process and the analysis. That is a valuable conversation to have.

> We, too, believe that a systematic, collegial, and constructive dialogue around the discussion of QRP$s in the management literature is valuable. As this conversation evolves, multiple directions for future work will emerge. We provide thoughts on the limitations of the present work and directions for future research in Appendix F.
References


Table 1. Definitions of questionable research practices (QRPs)

<table>
<thead>
<tr>
<th>QRPs</th>
<th>Description</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Selectively report hypotheses</td>
<td>Hypotheses with statistically non-significant results were less likely to be reported than hypotheses that achieved statistical significance.</td>
<td>Banks &amp; McDaniel (2011); Banks et al. (2012); Bedeian et al. (2010); Kepes et al. (2012); Leung (2011); O’Boyle et al. (in press); Pigott et al. (2013); Schmidt &amp; Hunter (2015); Simmons et al. (2011)</td>
</tr>
<tr>
<td>(2) Exclude data post hoc</td>
<td>A researcher conducts hypothesis testing. Some initial results are not statistically significant. After removing potential outliers, some of the initial results become statistically significant.</td>
<td>Bedeian et al. (2010); De Vries et al. (2006); Kepes &amp; McDaniel (2013); O’Boyle et al. (in press); Schmidt &amp; Hunter (2015); Simmons et al. (2011)</td>
</tr>
<tr>
<td>(3) HARKing</td>
<td>HARKing or “hypothesizing after results are known” occurs when a researcher analyzes data. After the data analysis, the researcher develops and reports post hoc hypotheses which suggest that findings were defined a priori, rather than identified post hoc.</td>
<td>Bedeian et al. (2010); Hitchcock &amp; Sober (2004); John et al. (2012); Kepes &amp; McDaniel (2013); Kerr (1998); Leung (2011); O’Boyle et al. (in press); Schmidt &amp; Hunter (2015)</td>
</tr>
<tr>
<td>(4) Selectively include control variables</td>
<td>Occurs when a researcher conducts multiple analyses to test the same hypothesis, each time adding or removing different control variables. The researcher only reports the use of control variables that allow for a statistically significant result.</td>
<td>John et al. (2012); Kepes &amp; McDaniel (2013); O’Boyle et al. (in press); Simmons et al. (2011)</td>
</tr>
<tr>
<td>(5) Falsify data</td>
<td>Fabricating a dataset rather than engaging in an actual data collection.</td>
<td>Bedeian et al. (2010); John et al. (2012); Schmidt &amp; Hunter (2015)</td>
</tr>
<tr>
<td>(6) “Round off” a p-value</td>
<td>A p-value: Reporting that a p-value of .054 is &lt; .05 rather than p = .05</td>
<td>Bakker and Wicherts (2011); Bakker and Wicherts (2014); John et al. (2012); Nuijten et al. (in press)</td>
</tr>
</tbody>
</table>

Note: The list of references for each QRP is an illustrative list and does not include all authors who have questioned the application of each of these practices.