In Defense of the Questionable: Defining the Basis of Research Scientists’ Engagement in Questionable Research Practices

Donald F. Sacco1, Samuel V. Bruton1, and Mitch Brown1

Abstract
National Institutes of Health principal investigators reported their perceptions of the ethical defensibility, prevalence in their field, and their personal willingness to engage in questionable research practices (QRPs). Using ethical defensibility ratings, an exploratory factor analysis yielded a two-factor solution: behaviors considered unambiguously ethically indefensible and behaviors whose ethical defensibility was more ambiguous. In addition, increasing perceptions that QRPs affect science predicted reduced acceptability of QRPs, whereas increasing beliefs that QRPs are normative or necessary for career success predicted increased acceptability of QRPs. Perceptions that QRPs are risky were unrelated to QRP acceptability but predicted reduced extramural funding (i.e., researchers’ lifetime extramural grants and total funding secured). These results identify risk (i.e., beliefs that QRPs are normative to stay competitive in one’s field) and protective factors (i.e., beliefs that QRPs have a significant negative impact on society) related to QRP endorsement that could inform educational interventions for training research scientists.

Keywords
questionable research practices, ethics, research funding, motives, behavioral social science research

Introduction
The past decade has seen growing interest in understanding research practices that undermine the quality of scientific research across numerous disciplines, including psychology and the medical science fields. The scope of behaviors deemed threatening to quality scientific research has expanded beyond research misconduct, standardly defined in terms of fabrication, falsification, and plagiarism (FFP), to include other questionable research practices (QRPs; John, Loewenstein, & Prelec, 2012). Unlike FFP, QRPs often resist simple definition or are indefensible in special contexts. Because of their subtlety, and because standard definitions of research misconduct are comparatively straightforward, QRPs may be perceived as more ethically defensible than FFP. Such perceptions could lead individuals to easily rationalize or defend QRPs and overlook the extent to which QRPs undermine scientific integrity. For example, when deciding whether to report the results of all measures or procedures from an experiment, one might infer that because failing to do so is not research misconduct, it is more defensible than FFP. However, studies demonstrate that such “researcher degrees of freedom” in data analysis and scientific reporting (e.g., inclusion of theoretically irrelevant covariates in statistical models) increase Type I error rates, thereby generating research findings that may be neither valid nor replicable (Ioannidis, 2005a, 2005b; Simmons, Nelson, & Simonsohn, 2011). Further doubts have been fueled by empirical demonstrations of several high-profile failures to replicate (e.g., Lynott et al., 2014; Open Science Collaboration, 2015; Ritchie, Wiseman, & French, 2012; Steele, Bass, & Crook, 1999).

The frequency with which researchers apparently engage in QRPs warrants concern. For example, whereas only about 2% of researchers admit to severe forms of research misconduct (FFP), more than 70% of sampled researchers admit to engaging in behaviors that would likely constitute QRPs (Fanelli, Costas, & Larivière, 2015). Thus, it seems that many researchers find QRPs sufficiently defensible to use in their own research, which may suggest that many reported scientific findings have been compromised. Recent research further highlights another potential problem: a 22% increase in reported positive results from between

1The University of Southern Mississippi, Hattiesburg, USA

Corresponding Author:
Donald F. Sacco, Department of Psychology, The University of Southern Mississippi, Owings-McQuagge Hall, 118 College Drive 5025, Hattiesburg, MS 39406, USA.
Email: Donald.Sacco@usm.edu
1990 and 2007 (Fanelli, 2010, 2011). Such an increase could indicate scientists’ reluctance to publish null findings, which would preclude the scientific community and society from having the most complete picture of current scientific knowledge (Fanelli, 2010, 2011; Ferguson & Heene, 2012). These trends have spurred discussions about a “reproducibility crisis,” the implied degradation of research quality, and a concomitant erosion of public trust (Begley & Ioannidis, 2015; cf. Maxwell, Lau, & Howard, 2015).

Given their apparent prevalence, many observers have called for the development of educational and intervention strategies to reduce researchers’ QRPs. Such strategies may involve enacting enhanced research integrity policies, improving mentoring on research integrity, and encouraging transparent communication between researchers, such as readily sharing research data (Fanelli et al., 2015). For example, some have suggested that journal outlets require researchers to make their data publicly available in a data repository (e.g., Open Science Framework) and encourage researchers to consult statisticians when analyzing their data to reduce both errors and secrecy (Sijtsma, 2016). Others’ practical recommendations include encouraging larger sample sizes, increasing reliability of study measures, avoiding multiple underpowered studies, and preregistering studies (Asendorpf et al., 2013).

Despite recommendations for improving the quality of scientific research, there is a relative dearth of literature exploring the underlying motives of researchers that predict the extent to which they find QRPs acceptable. One of the most widely studied factors contributing to unethical research practices is publication pressure. Specifically, researchers in the medical sciences report publication pressure as too high, and publication pressure is strongly associated with various forms of research misconduct (Tijdink, Verbeke, & Smulders, 2014). However, possibly many factors and beliefs influence the extent to which researchers view QRPs as ethically defensible and therefore predict researchers’ likelihood of engaging in them. Being able to identify the most robust underlying motives for their use thus could help deterrence efforts. The current study looks at how three such factors influence perceptions of QRPs: the extent to which researchers believe such behaviors significantly affect science or society, the extent to which researchers can rationalize their use of QRPs, and the extent to which researchers believe that QRPs are risky. These factors have been linked to other unethical, yet prevalent, types of behavior, including software piracy (Siponen, Vance, & Willison, 2012).

Impact, Rationalization, and Risk

One potentially useful theory for understanding motives for QRPs comes from the criminal justice literature. According to neutralization theory, offenders try to rationalize criminal behavior to justify it (Siponen et al., 2012). These techniques, called neutralization techniques, help to reduce the feelings of guilt and shame that would normally serve to inhibit objectionable behavior. Using neutralization theory as a framework, the current research focuses on three factors that may be implicated in attitudes toward QRPs. First, individuals may view QRPs as relatively inconsequential in that they are unlikely to cause harm to science or society. This factor is often referred to as denial of injury, whereby perpetrators regard their misdeeds as “victimless.” Endorsement of this denial of injury claim predicts individuals’ willingness to commit corporate crimes (Piquero, Tibbetts, & Blankenship, 2005).

In addition, individuals may convince themselves that others are using QRPs to their advantage, and therefore that by not engaging in QRPs, they put themselves at a competitive disadvantage. This is often referred to as the “defense of necessity” argument. Using observations of workplace health and safety in a factory in Canada, research documents how this rationalization process works (Gray, 2006). Specifically, a supervisor was willing to violate health and safety requirements on an assembly line, and justified these actions by claiming the company was under pressure and had to ship an important order. Indeed, research on white-collar crime suggests that organizational deviance is often justified due to it being in the best interest of the business and, therefore, necessary (Croall, 2001). Thus, a research scientist may feel that it is necessary, and even normative, to engage in questionable behaviors to stay competitive with respect to research publications and extramural funding.

Finally, whereas the denial of injury and defense of necessity arguments rationalize QRPs in virtue of their benefits, individuals must also weigh the personal costs associated with their use. According to deterrence theory, the threat of significant moral or ethical sanctions reduces willingness to engage in illicit activities. For example, the severity of moral sanctions has a significant negative effect on users’ willingness to commit computer abuse (D’Arcy, Hovav, & Galletta, 2009). Although unlike research misconduct, there are often no formal sanctions against using QRPs, engaging in them may have negative reputational effects (Lee, Lee, & Yoo, 2004). Thus, researchers may see a potential beneficial competitive edge to be gained from using QRPs, but the possibility of informal sanctions may inhibit their endorsement.

The Current Research

In the current research, we sought to identify potential categories of QRPs across a broad range of behaviors, including data analysis QRPs, scientific reporting QRPs, and financial or sponsor-related QRPs, among others, which could be characterized in terms of their ethical defensibility.
(or lack thereof). We created a scale tapping into these various QRP behaviors and sought to determine whether there was an underlying factor structure to these behaviors using an exploratory factor analysis. We also generated a scale involving three motives potentially related to QRP endorsement: QRP impact, QRP normalization, and QRP risk. We predicted that to the extent that participants could convince themselves that QRP behaviors are not harmful (denial of injury) or that they are necessary to remain competitive (defense of necessity), participants would indicate that QRP behaviors are more ethically defensible and would report greater willingness to engage in them. Conversely, we predicted that to the extent that individuals are concerned with moral sanctions for engaging in QRP behaviors, they would believe that QRP behaviors are less ethically defensible and would report less willingness to use them. Thus, this research sought to understand the potential decisional and motivational precursors that influence individuals’ attitudes toward QRP behaviors.

Importantly, we tested our hypotheses using a unique participant sample. We compiled a database of researchers across numerous disciplines who had at least one active National Institutes of Health (NIH) grant during 2016. This contact list contained approximately 3,500 individuals across 36 different research institutions across the United States. Whereas much research relies on convenience samples to understand QRP behaviors, we sought to recruit individuals who were active in their fields and successful at securing extramural funding. The QRP perceptions of such scientists are especially significant.

Method

IRB Statement

This project was approved by the University of Southern Mississippi’s Institutional Review Board (Protocol No. 16110904).

Participants

Using the NIH database of funded principal investigators (PIs; https://report.nih.gov/award/index.cfm), we recruited participants from various fields who had at least one active NIH grant during 2016. To obtain a diverse sample, we sent emails to researchers at 36 universities across the contiguous United States. Specifically, we divided the United States into six geographical regions (e.g., Northeastern US, Southwestern US) and identified six universities per region, with the only requirement being that each institution contained at least 100 NIH PIs. Following initial emails containing the invitation to participate in the study, participants were emailed a second invitation to participate 2 days later. We obtained 136 participants who rated the ethical defensibility of each of the 40 behaviors that provided the basis of our factor analysis. However, 17 participants failed to complete at least one of the questionnaires, thus truncating the amount of usable data in subsequent analyses; differences in degrees of freedom reflect these differences. Of those providing demographic information, the sample was comprised of 60 men and 59 women (M = 52.23 years, SD = 10.56); 21 participants did not report age, and 17 did not report gender. Participants came from more than 60 different self-reported disciplines (with psychology, cell biology, medicine, pediatrics, cardiology, and cancer research being some of the representative research disciplines). Participants were compensated with Amazon gift cards (US$5.00) for their participation; this compensation amount was selected to incentivize participation without being coercive.

Materials

QRP behaviors. We generated a series of 40 different research behaviors that were potentially QRP behaviors based on previous studies of common behaviors in psychology and medical science; these behaviors varied with respect to their potential severity (e.g., John et al., 2012; Tijdink et al., 2014). Participants responded to each item three times with a different question that assessed the extent to which participants found each QRP behavior to be ethically defensible (1 = completely indefensible, 4 = neither defensible nor indefensible, 7 = completely defensible), prevalent in the participant’s respective field (see Fanelli, 2009; 1 = very uncommon, 4 = neither common nor uncommon, 7 = very common), and a practice they would be willing to engage in (1 = completely unwilling to engage in this behavior, 4 = neither willing nor unwilling to engage in this behavior, 7 = completely willing to engage in this behavior). Higher scores indicated greater perceptions of defensibility, commonality, and willingness to use, respectively. These 40 items represented various categories of QRP behaviors, including data analytic strategies, scientific reporting, and interactions with financial sponsors. For each behavior, participants indicated its defensibility, its commonality, and their willingness to engage in it. Within each series of questions, QRP behaviors appeared in random order. (See Table 1 for the finalized items for the two-factor structure with accompanying descriptive statistics for each individual item.)

QRP motivations. Based on previous research exploring the mechanisms that predict consumer digital piracy (Koklic, Kukar-Kinney, & Vida, 2016), we adapted a three-item scale that assessed participants’ denial of injury, in terms of the impact to science or society, of engaging in QRP behaviors (e.g., “What would be the magnitude of impact on society?”), a three-item scale assessing defense of necessity through how scientists rationalized QRP behaviors (e.g., “It would not be a big deal because no one would be directly hurt or affected.”),
Journal of Empirical Research on Human Research Ethics 00(0)

Table 1. Descriptive Statistics for Individual Items Included in the Final Factor Structure for Ethical Defensibility.

<table>
<thead>
<tr>
<th>Items</th>
<th>M (SD)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Failing to attain appropriate informed consent.</td>
<td>1.13 (0.53)</td>
<td>6.52</td>
<td>52.43</td>
</tr>
<tr>
<td>Concealing data or results that contradict one's own previous research.</td>
<td>1.33 (0.65)</td>
<td>2.60</td>
<td>8.70</td>
</tr>
<tr>
<td>Use of others’ ideas, words, images, or other materials without citation.</td>
<td>1.18 (0.57)</td>
<td>4.16</td>
<td>20.19</td>
</tr>
<tr>
<td>Withholding relevant details of methodology or results in publication.</td>
<td>1.62 (1.07)</td>
<td>2.18</td>
<td>4.76</td>
</tr>
<tr>
<td>Reuse of one’s own previously published data, results, or analysis without citation.</td>
<td>1.72 (1.19)</td>
<td>1.90</td>
<td>2.94</td>
</tr>
<tr>
<td>Failing to disclose all potentially relevant conflicts of interest.</td>
<td>1.28 (0.53)</td>
<td>1.72</td>
<td>2.14</td>
</tr>
<tr>
<td>Withholding publication specifically to please a sponsor.</td>
<td>1.95 (1.45)</td>
<td>1.65</td>
<td>2.02</td>
</tr>
<tr>
<td>Submitting a manuscript in which the sponsor was involved in data analysis and/or manuscript preparation without disclosing the sponsor’s role.</td>
<td>1.28 (0.68)</td>
<td>2.94</td>
<td>9.53</td>
</tr>
<tr>
<td>Failing to acknowledge significant sources of financial assistance.</td>
<td>1.51 (0.99)</td>
<td>2.47</td>
<td>7.66</td>
</tr>
<tr>
<td>Adding or deleting coauthors to increase chances of acceptance.</td>
<td>1.92 (1.29)</td>
<td>1.45</td>
<td>1.34</td>
</tr>
<tr>
<td>Overlooking or ignoring others’ research misconduct.</td>
<td>1.49 (0.88)</td>
<td>2.24</td>
<td>5.12</td>
</tr>
<tr>
<td>Refusing to share data or materials with other researchers to prevent questions about the quality of your work from being raised.</td>
<td>1.27 (0.65)</td>
<td>3.45</td>
<td>15.22</td>
</tr>
<tr>
<td>Overlooking or ignoring others’ questionable research practices.</td>
<td>1.75 (1.01)</td>
<td>1.48</td>
<td>1.78</td>
</tr>
<tr>
<td>Providing a biased peer review of a manuscript submitted to a journal for publication to delay its publication so as to benefit one’s own research.</td>
<td>1.12 (0.36)</td>
<td>3.27</td>
<td>10.86</td>
</tr>
<tr>
<td>Refusing to provide experimental data to a coauthor.</td>
<td>1.36 (0.99)</td>
<td>3.75</td>
<td>15.29</td>
</tr>
<tr>
<td>Submitting research for publication without permission from all authors.</td>
<td>1.53 (0.95)</td>
<td>2.12</td>
<td>4.52</td>
</tr>
<tr>
<td>Factor 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>To enhance chances of publication, violating the ideal of “replace, reduce, refine” regarding the use of research animals.</td>
<td>2.06 (1.49)</td>
<td>1.68</td>
<td>2.21</td>
</tr>
<tr>
<td>Adding additional research participants because the results collected thus far are not yet statistically significant.</td>
<td>4.15 (2.05)</td>
<td>−0.08</td>
<td>−1.30</td>
</tr>
<tr>
<td>Stopping collecting data earlier than planned because the hypothesized result already had been attained.</td>
<td>4.04 (1.80)</td>
<td>−0.01</td>
<td>−0.98</td>
</tr>
<tr>
<td>Rounding off a ( p ) value simply to make results seem more significant, such as by reporting a ( p ) value of .044 to be ( p = .04 ).</td>
<td>3.26 (1.88)</td>
<td>0.44</td>
<td>−0.87</td>
</tr>
<tr>
<td>Deciding whether to include or exclude data after looking at the impact of doing so on the results.</td>
<td>1.96 (1.33)</td>
<td>1.54</td>
<td>1.83</td>
</tr>
<tr>
<td>Failing to report all of a study’s outcome measures.</td>
<td>2.96 (1.80)</td>
<td>0.60</td>
<td>−0.81</td>
</tr>
<tr>
<td>Selectively discussing only studies that supported the hypothesized result(s).</td>
<td>1.84 (1.00)</td>
<td>1.16</td>
<td>0.86</td>
</tr>
<tr>
<td>Reporting an unexpected result as having been hypothesized from the start.</td>
<td>3.13 (1.68)</td>
<td>0.60</td>
<td>−0.29</td>
</tr>
<tr>
<td>Drawing strong inferences from statistically significant but underpowered results.</td>
<td>2.73 (1.47)</td>
<td>0.66</td>
<td>−0.42</td>
</tr>
<tr>
<td>Selective reporting of subgroups, outcomes, and time points.</td>
<td>2.42 (1.64)</td>
<td>1.22</td>
<td>0.55</td>
</tr>
<tr>
<td>Deliberately delaying reporting results to publish findings in a higher impact journal.</td>
<td>4.85 (1.77)</td>
<td>−0.31</td>
<td>−0.92</td>
</tr>
<tr>
<td>Reuse of one’s own previously published ideas or words without citation, such as parts of a literature review section, introduction, or methodology, but without reusing data, results, or analysis.</td>
<td>3.60 (2.00)</td>
<td>0.33</td>
<td>−1.10</td>
</tr>
<tr>
<td>Publishing results of a single study as several articles simply to increase the number of publications derived from the research (the so-called “salami slicing” problem).</td>
<td>3.31 (1.69)</td>
<td>0.45</td>
<td>−0.77</td>
</tr>
<tr>
<td>Changing the design, methodology, or results of a study to please a sponsor.</td>
<td>1.96 (1.53)</td>
<td>1.64</td>
<td>1.74</td>
</tr>
<tr>
<td>Acknowledging another’s technical assistance in publication without that person’s permission.</td>
<td>3.21 (1.76)</td>
<td>0.37</td>
<td>−0.78</td>
</tr>
</tbody>
</table>

Notes. Higher scores connote greater defensibility for each practice.

and the personal risk associated with engaging in QRPs (e.g., “It might interfere with my ability to secure extramural funding.”). Each item was assessed using a 7-point scale. Impact subscale items were anchored between 1 (very small) and 7 (very large) whereas Rationalization and Risk subscale items were anchored between 1 (strongly disagree) and 7 (strongly agree). The three subscales were reliable (Impact, \( \alpha = .90 \); Rationalization, \( \alpha = .78 \); Risk, \( \alpha = .87 \)). As
such, we computed composite Impact, Rationalization, and Risk scores for each participant, where higher scores are associated with greater perceptions of impact for QRPs, rationalizations to justify QRPs, and greater perception of risk associated with engaging in QRPs, respectively.

Procedure
Participants received an email invitation describing the study. Participants were informed that the survey was anonymous insofar as their individual responses were not linked with any personally identifiable information. Interested parties clicked on the link and provided informed consent anonymously, which led them to respond first to the QRP questions, followed by the QRP motivation scales, and finally a demographics questionnaire, which included a single open-ended question allowing participants the opportunity to provide their general comments in the space provided. After being debriefed, participants had the option of clicking a link redirecting them to a separate questionnaire in which they could enter their name and email address to receive a redemption code for an Amazon e-gift card. The software automatically delinked participants’ names from their responses to the main questionnaires.

Results
Ethical Defensibility Factor Analysis
We used a factor analysis to identify covarying groups of items that would ultimately emerge as individual factors. Specifically, we used principal axis factoring with a Direct Oblimin rotation to analyze potential correlations between the proposed factors. Our initial analysis suggested that a two-factor solution was the most appropriate and parsimonious model for our QRPs. The Kaiser–Meyer–Olkin measure of sampling adequacy in this solution indicated that items shared sufficient common variance for the current analysis (0.749). Bartlett’s test of sphericity further indicated no violations of sphericity, $\chi^2(780, N = 139) = 2,343.27, p < .001$.

Although we had a priori reasons for imposing a two-factor model on these data, we found it prudent to determine whether the data justified this solution. The first 10 eigenvalues in our analysis were 9.44, 3.56, 2.13, 1.86, 1.72, 1.62, 1.48, 1.31, 1.28, and 1.12. We used a visual scree test from the generated plot of our eigenvalues, which occurs by determining how a line drawn between eigenvalues on the plot changes slopes between eigenvalues (Cattell, 1966; Horn, 1965). Two factors clearly emerged in our data. However, for the sake of being exhaustive, we conducted both a two- and a three-factor solution.

In determining how individual items loaded for the initial two-factor solution, we considered (a) whether the item’s factor loading coefficient met or exceeded 0.30 and (b) whether items double-loaded to one of the factors; our criterion for double-loading was whether both coefficients must have exceeded 0.30. The pattern matrix indicated that five items (e.g., “Reuse of a sample collected for one purpose without the participants’ consent to use it for an additional purpose.”) did not load onto either of the factors. Three items double-loaded (e.g., “Failing to report all of a study’s conditions.”). Finally, despite a coefficient of 0.714 on Factor 1, another item (“Fabricating data by adding data for participants who in fact did not participate.”) elicited no variability between participants, thus justifying its removal. We removed these items from further consideration in our construction of factors and considered only the remaining 31 items in our two-factor model.

Our two factors were readily interpretable given the theoretical predictions directing this analysis. Factor 1 consisted of 16 items that reflected QRPs scientists seemed to perceive as categorically indefensible, that is, items for which participants perceived no viable ethical justification for engaging in the behavior described. We thus named this factor Unambiguously Unethical QRPs (UU-QRPs). Among the items ultimately considered for future analyses, the three highest UU-QRPs’ factor loadings were for “Failing to attain appropriate informed consent,” “Concealing data or results that contradict one’s own previous research,” and “Refusing to share data or materials with other researchers to prevent questions about the quality of your work from being raised.” Factor 2 consisted of 15 QRPs that scientists appeared more ready to defend, that is, behaviors that seemed to be perceived as “researcher degrees of freedom” (e.g., Simmons et al., 2011). We named this factor Ambiguously Unethical QRP (AU-QRPs). The three highest AU-QRP loadings were for “Failing to report all of a study’s outcome measures,” “Drawing strong inferences from statistically significant but underpowered results,” and “Publishing results of a single study as several articles simply to increase the number of publications derived from the research (the so-called “salami slicing” problem).” A positive correlation emerged between these factors ($r = .331$). The factor analysis for ethical defensibility provided the basis for what items to include in subsequent correlational analyses with willingness and commonality scores.

When conducting an exhaustive three-factor solution for our data to determine that the two-factor model was indeed the best possible solution for our results, no clear distinction emerged that would warrant naming of factors in that model. Specifically, a majority of items double-loaded in this analysis, indicating a two-factor model was mathematically simpler and more theoretically sensible. Thus, we continued to consider our results in the two-factor model.

These two factors for ethical defensibility, commonality, and willingness were reliable: UU-QRP Ethically Defensible
(α = .85), UU-QRP Willingness (α = .78), UU-QRP Common (α = .92), AU-QRP Ethically Defensible (α = .81), AU-QRP Willingness (α = .86), and AU-QRP Common (α = .92). A paired-samples t test revealed that participants found UU-QRPs (M = 1.46, SD = 0.50) less ethically defensible than AU-QRPs (M = 3.03, SD = 0.90), t(135) = −23.92, p < .001, d = 4.12. Another paired-samples t test indicated participants reported less willingness to engage in UU-QRPs (M = 1.36, SD = 0.41) than AU-QRPs (M = 2.60, SD = 1.07), t(118) = −15.31, p < .001, d = 2.82. A final paired-samples t test revealed that participants indicated that UU-QRPs were less common in their field (M = 2.79, SD = 1.07) than AU-QRPs (M = 3.97, SD = 1.18), t(124) = −18.32, p < .001, d = 3.29. Participant sex did not qualify any of the findings reported above.

Relations Between QRP Motives, QRP Perceptions, and Self-Reported Behaviors

Correlational analyses revealed that the more impact participants perceived QRPs to have, the less ethically defensible they perceived both UU-QRPs, r(118) = −.289, p = .001, and AU-QRPs, r(118) = −.207, p = .023, to be. While perceptions of QRP Impact were not related to willingness to engage in UU-QRPs, r(116) = −.137, p = .139, higher perceptions of QRP Impact were associated with reduced willingness to engage in AU-QRPs, r(116) = −.196, p = .034. Perceptions of QRP impact were unrelated to the extent to which participants believed UU-QRPs were less common (M = 3.48, SD = 1.01), r(117) = −.135, p = .143, were common.

Participants’ greater ability to rationalize QRPs predicted higher ethical defensibility ratings for both UU-QRPs, r(118) = .274, p = .002, and AU-QRPs, r(118) = .377, p < .001. Participants’ greater ability to rationalize QRPs predicted greater willingness to engage in both UU-QRPs, r(116) = .338, p < .001, and AU-QRPs, r(115) = .428, p < .001. Participants’ greater ability to rationalize QRP behaviors predicted greater beliefs that both UU-QRPs, r(116) = .448, p < .001, and AU-QRPs, r(117) = .451, p < .001, are common in their field.

Perceptions of QRP risk were unrelated to perceptions of ethical defensibility for both UU-QRPs, r(118) = −.076, p = .407, and AU-QRPs, r(118) = −.042, p = .645. Perceptions of QRP risk were unrelated to participants’ self-reported willingness to engage in both UU-QRPs, r(117) = −.156, p = .090, and AU-QRPs, r(115) = −.100, p = .282. Perceptions of QRP risk were unrelated to the extent to which participants believed both UU QRPs, r(115) = −.126, p = .174, and AU-QRPs, r(116) = −.024, p = .800, are common in their field.

Relations Between Funding Success, QRP Perceptions, and Motives

Because we recruited only NIH-funded PIs to participate in this research, we asked participants to provide an estimate of (a) how much extramural funding (in U.S. dollars) they had secured over the course of their research careers and (b) how many extramural grants they had funded over the course of their careers. One hundred three participants provided an interpretable response for amount of funding received, and 115 participants provided an interpretable response for how many extramural grants they had funded. We excluded participants who provided responses such as “always been funded” or “32 years of continuous funding,” as there was no way to quantify these kinds of responses. In addition, if a person responded “at least $10,000,000” or “>20 grants,” we conservatively coded these as “$10,000,000” and “20 grants,” respectively. Participants reported securing an average of $10,368,398.17 over the course of their careers (SD = US$14,084,196.37) and 9.82 funded extramural grants (SD = 9.49).

Perhaps not surprisingly, gender differences emerged with respect to total extramural grants received and total extramural funding secured. Men reported securing significantly more extramural grants over their careers (M = 12.75, SD = 10.44) than did women (M = 7.12, SD = 7.75), t(99.35) = 3.25, p = .002, d = 0.61. Men also reported securing more extramural funding (M = US$14,543,396.23, SD = US$16,264,573.46) than did women (M = US$6,002,959.43, SD = US$9,752,549.29), t(86.19) = 3.24, p = .002, d = .70. For both analyses, adjusted degrees of freedom and t values are reported due to violations of homogeneity of variance. It is possible that these differences reflect documented gender disparities in representation and opportunities in STEM fields (e.g., van der Lee & Ellemers, 2015).

Interestingly, neither total grants secured nor total funding dollars received were significantly related to perceptions of QRP ethical defensibility, perceptions of QRP commonness, or willingness to engage in QRPs (all ps > .210). Thus, perceptions of QRP acceptability seem unrelated to extramural funding, at least in this sample.

Perceptions of QRP impact and rationalization were unrelated to total number of grants secured and total grant dollars received (all ps > .300). However, perceptions that engaging in QRPs is risky predicted both fewer grants, r(110) = −.278, p = .003, and fewer grant dollars secured over one’s research career, r(100) = −.260, p = .008.

Discussion

Given the prevalence of researchers who report engaging in QRPs across scientific disciplines, and given the fact that QRPs have been linked to increased Type I error rates and reductions in scientific reproducibility, it is incumbent on the scientific community to (a) identify factors that contribute to QRP utilization and to (b) develop interventions capable of decreasing QRP behavior (Begley & Ioannidis, 2015; Fanelli, 2009; John et al., 2012; Simmons et al., 2011). Until now, however, researchers have focused mainly
on what kinds of behaviors constitute QRPs, the extent to which researchers demonstrate willingness to engage in these behaviors, and how these behaviors may negatively affect the validity and reproducibility of scientific findings. Much less research has focused on the underlying psychological and social factors that predict relative endorsement of QRPs.

The current study sought to expand the scope of research on QRPs in several notable ways. First, we used a representative sample of NIH-funded PIs to validate a series of potential QRPs. That is, participants were any NIH PI holding at least one extramural NIH grant; no additional criteria for selection were included. These QRPs were both consolidated from previous research and created specifically for the current study. Indeed, our questionnaire contained more items than previously published work on QRPs (John et al., 2012; Tijdink et al., 2014). Given the notable research experience of our participants, their perceptions are especially useful in assessing the various behaviors described in our questionnaire, even for practices far removed from their own research (e.g., animal researchers responding to questions about informed consent for human subjects). Indeed, participants in our sample had secured an average of approximately 10 external grants totaling nearly US$10 million, suggesting these were highly active researchers. Thus, their views provide a particularly important perspective.

Furthermore, whereas other research has organized QRPs around mean scores such as severity or prevalence, we sought to identify classifications of QRPs based on factor analysis. While our questionnaire included a broad array of research practices, these behaviors seemed to load on two distinct factors. The first factor contained behaviors that most closely fell under the standard definition of research misconduct (e.g., “Concealing data or results that contradict one’s own previous research.”). Perhaps not surprisingly, participants in our sample rated these practices as highly ethically indefensible. However, the second factor contained a variety of behaviors that are consistent with most other definitions of QRPs: behaviors that may be detrimental to science, but in a more ambiguous or contextually dependent way (e.g., “Deciding whether to include or exclude data after looking at the impact of doing so on the results.”). For these latter behaviors, participants reported them as significantly more ethically defensible than the others. Indeed, in an open-ended comments section at the end of our study, several participants noted that many of these behaviors are not “black and white” but depend on the nature of study phase, study design, characteristics of the data, and other factors. These qualitative participant responses suggest researchers’ awareness of the context-dependent ethicality for some of these practices.

Perhaps the largest contribution of our work involves an assessment of motivational factors that may predict perceptions of QRP acceptability. Specifically, we explored how the extent to which participants believed the impact of QRPs to be inconsequential, the extent to which they believed QRPs to be normative and necessary, and the extent to which QRPs are risky influence QRP acceptability. Our research revealed that the more participants believed QRPs to be inconsequential, the more ethically defensible they found them to be and the more personal willingness they expressed in engaging in them. Thus, perhaps one means of reducing QRP use is to better communicate the impact they have on science.

Furthermore, to the extent that people believed QRP behavior was normative, the more they found QRPs to be ethically defensible and the more willing they were to engage in them. Thus, if researchers believe it is necessary to engage in these behaviors to keep up with peers, then they find them more acceptable. This line of thinking is not without basis. A sample of research scientists indicated that 72% of respondents communicated that they had engaged in some kind of QRP behavior; perceptions of high normativity seem to accurately reflect actual researcher behavior (Fanelli, 2009). As such, educational training should focus on eliminating the inference that because QRP behavior is normative, it is acceptable. Such logic would align with how group norms facilitate adoption of categorically appropriate behavior in a given context. Whereas competitive norms may elicit questionable behavior, adoption of certain ethical standards addressing QRPs may mitigate their use. Indeed, previous research has demonstrated that prosocial norms foster prosocial behaviors (Cialdini, 2003; Nook, Ong, Morelli, Mitchell, & Zaki, 2016). Thus, inducing similar prosocial norms in research domains might well encourage ethical research behaviors. If researchers are more likely to perceive other researchers as not engaging in behavior they believe undesirable, they may be less likely to behave undesirably themselves.

Interestingly, perceptions that QRPs are risky were largely unrelated to participants’ perceptions of the ethical defensibility of QRPs and their willingness to engage in QRPs. Thus, concern with moral sanctions did not induce reduced willingness to engage in QRPs or reduced perceptions of their defensibility. Perhaps researchers are aware that these behaviors are risky but see the benefits as outweighing the costs. More likely, however, individuals do not believe that there is a significant likelihood of sanctions for these behaviors, perhaps even if caught, thereby reducing the risks for engaging in them. Future research should attempt to frame QRPs in the same context as categorically unethical behavior, which may foster perspectives that any behavior undermining scientific integrity is indefensible (e.g., Kahneman & Tversky, 1984).

Although researchers may not concern themselves with sanctions associated with QRPs, reshaping their attitudes about ethical research practices may prevent ideas that
enable rationalization of QRPs. Such reframing could occur through interventions defining research ethics in terms of traditional conceptualizations of research integrity (e.g., avoiding FFP) and providing additional framing in terms of QRPs. Although FFP-related behaviors may be seen as universally indefensible, framing QRPs as unethical as well might meaningfully affect researchers’ views on these behaviors.

Despite the emergence of theoretically sensible data, the current study is not without limitations. First, our response rate was relatively low. Although we compiled a list of 3,600 NIH PIs, we only obtained response from 136 participants. It is possible that this low response rate was due to insufficient compensation for participation, aversion to responding to sensitive research behavior questions, or both. Primarily, several participants abstained from responding to certain measures (e.g., willingness), which precludes the current study of having a fully comprehensive, data-driven picture of the extent QRPs are prevalent or deemed acceptable by the scientific community. Several participants commented that they found some of the questions “inflammatory.” Future investigations may benefit from reminding participants that not all of these behaviors are wrong in every research context; thus, endorsing them may not reflect indifference to research ethics. Rather, follow-up studies may benefit from framing QRPs as part of an ongoing process of dialogue and refinement of research ethics standards. Nonetheless, a reluctance to respond to certain questions may indicate a certain measure of defensiveness.

Finally, it is worth noting that in our sample, neither the total number of secured extramural grants nor total funding secured over participants’ careers was related to any aspect of QRP acceptability. Thus, it was not the case that those who have secured more grants and grant money are more accepting of, or more willing to engage in, QRP behaviors. It is possible that these highly experienced and successful researchers are acutely aware of acceptable research practices. Alternatively, given the self-reporting nature of the current study, they simply may be unwilling to share their actual perceptions of these behaviors. The only interesting relationship between funding success was related to QRP risk perception. Specifically, participants who expressed greater concern with the personal risks associated with engaging in QRPs obtained significantly fewer extramural grants and less extramural grant funding over the course of their research careers. While the reasons for this are unclear, it seems that those who are concerned about being sanctioned for QRP behavior are less successful at obtaining extramural funding.

Best Practices

The current results highlight a stratification between behaviors based on their degrees of unethicality. Whereas some behaviors are perceived as categorically indefensible, some behaviors are perceived as being more acceptable, depending on the extent to which they can be rationalized by the researcher. Nonetheless, all of these practices retain “researcher degrees of freedom” that may prevent replicability of findings, which may undermine public trust. This article highlights a systematic difference in how questionable research behavior is viewed, and provides a starting point for improving educational efforts directed toward best practices of research ethics.

Because the consequences associated with ambiguous QRPs may be less obvious to researchers, ethics training should make clear the adverse effects of these practices. In addition, early-career scientists likely would benefit from greater integration of research methods classes, such as those addressing techniques of statistical analysis, and research ethics training. Current research ethics training focuses heavily on research misconduct, narrowly defined, but more attention needs to be paid to the ethically “messier” contextually dependent behaviors that constitute QRPs. The earlier this training occurs in researchers’ careers, the better; it is important that good research ethics practices become habitual. Ideally, it would seem appropriate for institutions to implement ethics training modules for early-career researchers to regard ambiguous QRPs as more damaging and less defensible to discourage formation of habitually questionable behaviors. Such interventions ultimately may contribute to a more unified and collective conception of QRPs and research integrity. However, given mixed evidence for the efficacy of training as it relates to the promotion of research integrity and prevention of research misconduct (e.g., Marusic, Wager, Utrobic, Rothstein, & Sambunjak, 2016), future research should identify additional means of preventing QRPs and facilitating ethical research practices.

Research Agenda

Research scientists in this study perceived some questionable research behaviors as less damaging than others, which could facilitate their rationalization. Future studies should identify techniques for increasing researchers’ awareness of the ethics and adverse effects of QRPs. Such interventions could include real-world examples and case studies of context-dependent QRPs and their scientific impact. For example, Simmons and colleagues (2011) reported results from two experiments supporting an extraordinary hypothesis. However, in their initial report of results, they did not disclose certain flexibilities in analyses that ultimately undermined what they reported. Discussions of such nonhypothetical examples and their consequences may help convey the significance of even merely questionable practices.
Educational Implications

These findings afford institutions and mentors potential starting points for improving research integrity pedagogy. By fostering a mind-set that encourages openness and transparency, research scientists may feel more motivated to conduct their work ethically and be empowered to behave appropriately with their data. Training researchers to recognize the potential for a variety of QRPs to be damaging to science could prevent their occurrence.

Conclusion

Although scientists seem to be aware of the potential for damage caused by QRPs, perceptions of the acceptability of these practices may indicate considerable persisiveness of behaviors that significantly undermine the integrity of scientific research. This study identified a specific set of research behaviors that, while nonoptimal, seem to some degree tolerated in science. This implies the general vulnerability of science to researchers’ interests. Furthermore, we identified some specific motives related to the use of QRPs. These data provide the bases for further investigating researchers’ motives to commit QRPs to prevent them.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by a grant from the U.S. Department of Health and Human Services Office of Research Integrity (Grant 1 ORIR160021-01-00) awarded to the first and second authors.

References


Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. Scientometrics, 90, 891-904.

Fanelli, D., Costas, R., & Larivière, V. (2015). Misconduct policies, academic culture and career stage, not gender or pressures to publish, affect scientific integrity. PLoS ONE, 10, e0127556.


**Author Biographies**

**Donald F. Sacco** is an assistant professor of social psychology at The University of Southern Mississippi, USA. His research interests include the impact of financial conflict of interest disclosure on perceptions of researcher ethicality as well as how educational experience and academic background influence these perceptions. He was involved in the conceptualization of the research project, contributed to the creation of the study materials, served a primary role regarding data analysis, and assisted in the writing of the article.

**Samuel V. Bruton** is an associate professor of philosophy and director of the Office of Research Integrity at The University of Southern Mississippi, USA. His research interests include research ethics and discrimination. He was involved in conceptualization of the research project, contributed to the creation of the study materials, and assisted in the writing of the article.

**Mitch Brown** is a graduate student in social psychology at The University of Southern Mississippi, USA. His research primarily concerns the various bases of social perception, including those pertaining to questionable research practices. For the current article, he recruited participants via email, contributed to data analysis, and assisted in writing the article.