The University of Southern Mississippi

The Aquila Digital Community

12-21-2019

Ethical Consistency and Experience: An Attempt to Influence Researcher Attitudes Toward Questionable Research Practices Through Reading Prompts

Samuel V. Bruton
University of Southern Mississippi, samuel.bruton@usm.edu

Mitch Brown
Fairleigh Dickinson University

Donald F. Sacco
University of Southern Mississippi, Donald.Sacco@usm.edu

Follow this and additional works at: https://aquila.usm.edu/fac_pubs

Part of the Applied Ethics Commons

Recommended Citation

Ethical Consistency and Experience: An Attempt to Influence Researcher Attitudes Toward Questionable Research Practices Through Reading Prompts

Samuel V. Bruton[1] ☉, Mitch Brown[2], and Donald F. Sacco[1] ☉

Abstract

Over the past couple of decades, the apparent widespread occurrence of Questionable Research Practices (QRPs) in scientific research has been widely discussed in the research ethics literature as a source of concern. Various ways of reducing their use have been proposed and implemented, ranging from improved training and incentives for adopting best practices to systematic reforms. This article reports on the results of two studies that investigated the efficacy of simple, psychological interventions aimed at changing researcher attitudes toward QRPs. While the interventions did not significantly modify researchers' reactions to QRPs, they showed differential efficacy depending on scientists' experience, suggesting complexities in researcher psychology and the ethics of QRPs that merit further study.

Keywords

questionable research practices, research ethics, psychology, statistics, research ethics, RCR

Introduction

Over the past couple of decades, Questionable Research Practices (QRPs)—methods of research, analysis, and reporting techniques that raise questions about the integrity of the work or the way it is presented—have been widely discussed in the research ethics literature as a source of concern (see, for example, Begley & Ioannidis, 2015; Bouter et al., 2016; Martinson et al., 2005; Sacco et al., 2018). In contrast to the concept of research misconduct, which is commonly understood in terms of the U.S. federal government's narrow definition of it as fabrication, falsification, and plagiarism (Resnik et al., 2015), there is no definitive and exhaustive list of QRPs. Several practices are nonetheless frequently mentioned as examples. These include publication bias (Fanelli, 2012), significance chasing (Ware & Munafò, 2015), misleading or manipulated authorship designations and citation practices (Fong & Wilhite, 2017; Wislar et al., 2011), lack of statistical power (Button et al., 2013; Crutzen & Peters, 2017), turning a blind eye to others' use of flawed data or questionable citations (Tijdink et al., 2014), citation bias (Fanelli et al., 2017), presentational "spin" (Chiu et al., 2017), and several others.

It is difficult to determine the prevalence of QRPs precisely; some studies indicate that it may be quite high. Fanelli (2009) reports that almost 34% scientists surveyed admitted to using QRPs other than research misconduct, and in surveys about the behavior of colleagues up to 72%

acknowledged QRP use. Such findings have produced a general consensus that QRPs tend to negatively impact science and its progress. QPRs are frequently implicated in the fact that so many experimental findings apparently cannot be replicated or reproduced (Agnoli et al., 2017; Ioannidis, 2005). Their use inflates Type 1 error rates (Simmons et al., 2011), and they tend to reduce overall transparency in scientific practice and reporting (Munafò et al., 2017). These difficulties not only point toward the ways QRPs can blight the research literature, with attendant negative effects of skewing meta-analyses, misleading other scientists who mistakenly rely on flawed findings, and so forth. Their negative effects also include wasted research funding (Chalmers & Glasziou, 2009), inhibited scientific collaboration, imperiled human and animal health, and weakened public trust in science (Joynson & Leyser, 2015).

While there is no clear consensus among researchers regarding the extent of the QRP problem or the degree to which it can be remedied (Motyl et al., 2017), most scientists and observers concur that improved research

[1]The University of Southern Mississippi, Hattiesburg, USA
[2]Fairleigh Dickinson University, Teaneck, NJ, USA

Corresponding Author:
Samuel V. Bruton, The University of Southern Mississippi, 118 College Drive, #5037, Hattiesburg, MS 39406, USA.
Email: Samuel.Bruton@usm.edu

methods and practices need to become more normative. This is true for example in psychology (John et al., 2012; Open Science Collaboration, 2015) and some medical fields (Button et al., 2013; Gerritts et al., 2019), where QRP use has been shown to be especially common. Various recommendations for improving the integrity and replicability of science have been made or are currently being implemented. These suggestions range from ways of changing researcher behavior, such as improved training (Casadevall & Fang, 2018; Munafò et al., 2017) and motivational badges (Kidwell et al., 2016), to macro-level reforms. These broader reforms include institutional rules for archiving data, lab journals, and scripts for data analysis (Bouter, 2015), enhanced policies (3rd World Conference on Research Integrity, 2013; World Medical Association, 2013), reporting guidelines (Fanelli, 2013), more rigorous journal practices (Vazire, 2016), and modifications to traditional modes of peer review (Herron, 2012).

While QRP use and disappointingly low rates of reproducibility in the sciences undoubtedly stem from a variety of causes, many commentators note that researchers' cognitive biases play an important role (Mazar & Ariely, 2015; Nosek et al., 2012). Avoiding flawed reasoning, however, is easier said than done. As Richard Feynman (1974) put the point memorably, the "first principle" of science is that "you must not fool yourself—and you are the easiest person to fool." Substantial evidence from psychology supports this concern. Confirmation bias is a major problem in scientific research as in other aspects of life. Humans naturally gravitate toward conclusions that support existing beliefs and hypotheses, downplaying or ignoring countervailing evidence. It is difficult for even the best scientists to avoid settling on results and interpretations that support their preconceptions (Mynatt et al., 1977; Nickerson, 1998). In academic contexts, the tendency toward confirmation bias is augmented by the fact that the reward structure of science often gives scientists a motive for favoring some scientific conclusions more than others. Given that the dominant "coin of the realm" is peer-reviewed research publications, scientists are quite naturally motivated to find publishable results. In light of well-known problems of publication bias in the scientific literature (Fanelli, 2012), this means that researchers have an incentive to produce positive findings that will be of interest to leading, "high impact" journals.

Recent work demonstrates the difficulty of adhering to objective standards when they do not wholly align with our goals. Even though we are always eager to preserve the sense of ourselves as ethically conscientious, most people, when given the opportunity, cheat at least a little (Ariely, 2013; Gino et al., 2016; Mazar et al., 2008). Rather than ameliorating this tendency, scientists' originality can exacerbate the problem. In part, this stems from the fact that creative people are more adept at justifying their behavior (Gini & Ariely, 2012). The relevance of these points in the present

context is that QRPs can increase the likelihood of publishable results without resorting to conscious deception or outright fraud. As the human tendency toward confirmation bias and motivated reasoning meets the incentives and career-oriented imperatives of academic science, it is reasonable to suppose that a significant aspect of science's QRP problem is the result of such psychological tendencies.

Such biases can manifest themselves in various ways. One problem is "hypothesis myopia," in which explanations for phenomena other than the research hypothesis under investigation are ignored (Nuzzo, 2015). A related temptation is the so-called "Texas sharpshooter fallacy" which consists of mistaking random patterns in the data for significant findings. This problem can be magnified by "hindsight bias," the tendency to think outcomes were more predictable than they were before experiments were conducted (Fischhoff, 1975). The "bias blind spot" makes such flaws in our thinking opaque to even ourselves (Scopelliti et al., 2015), suggesting that conscientiousness alone may not be an adequate corrective. Scientific reasoning is vulnerable to such motivated reasoning and confirmation bias in part because the issues are difficult, and the available information is ambiguous. Research findings are typically susceptible to multiple interpretations, and in the face of such flexibility, humans are prone to settle upon the interpretation that supports their goals (Bastardi et al., 2011). It is often tempting to conclude that the more publishable interpretation of the data is the most reasonable and best supported (Wagenmakers et al., 2011).

One way to dampen the effects of cognitive shortcomings and motivated reasoning on scientists is by changing research practices. Single- or double-blinded study designs are examples of this kind of approach (Munafò et al., 2017), though journals still do not demand blinding as much as is advocated by some commentators (Begley & Ioannidis, 2015). Preregistration of experiments and providing open access to data are other examples that have many recent proponents. Preregistration websites such as the Open Science Framework (OSF: https://osf.io/) and AsPredicted (https://aspredicted.org/) facilitate archiving the entire research process to make it more transparent and prevent the occurrence of QRPs (Nosek et al., 2018). The fact that some journals now essentially commit to publish studies based on preregistration plans alone, regardless of positive findings, supports these efforts. But not all study designs permit blinding, and neither is preregistration is a cure-all. Apart from being vulnerable to being hacked by researchers with devious intent, preregistration does not preclude the unintended use of QRPs (Ikeda et al., 2020). Also, absent systemic pressures, researchers can be surprisingly recalcitrant about making the extra efforts required by preregistration and open access (Washburn et al., 2018).

Rather than directly aiming to change research practices, the strategy behind the two studies reported below took a

different tack. Both attempted to change researcher attitudes about QRPs through a simple intervention. Specifically, both experiments aimed to mitigate researchers' approval of QRPs and their self-reported willingness to engage in them using short written statements about research integrity and ethical motivations. Thus, the strategy was to modify the way researchers think about certain ethical aspects of their work rather than advocate for specific behaviors.

There is theoretical support for interventions of this sort. In general, research ethicists and educators are recognizing the importance of better integrating the work of social scientists and psychologists to support research integrity (Redman & Caplan, 2016). Our studies were informed by research in behavioral ethics and psychology. One intervention consisted of emphasizing the overall negative impact of QRPs on science, an attempt to surpass the "tipping point" of motivated reasoning (Redlawsk et al., 2010). Similar interventions designed to change individuals' perceptions of risks have shown some effectiveness at increasing vaccination rates (Sheeran et al., 2014). The second intervention attempted to negate the natural tendency to rationalize possible misdeeds by modifying researchers' views of what behaviors were normative in their fields (Tsang, 2002). In effect, this intervention attempted to leverage perceived social norms to change attitudes, similar to the way experimentally manipulated social norms have been shown effective at changing hotel towel reuse (Goldstein et al., 2008) and energy conservation (Nolan et al., 2008). A third intervention prompted researchers to consider their conceptions of themselves as ethical scientists. It was informed by self-perception theory (Bem, 1972) and built on work showing that a desire to maintain consistency with oneself and past good deeds can be invoked by asking individuals to recall past moral behavior (Conway & Peetz, 2012). The fourth intervention involved activating a commitment to good science, an approach grounded in findings on reducing cognitive dissonance (Shu et al., 2012). Precommitment strategies have been shown successful in a variety of contexts, from helping students avoid procrastination (Ariely & Wertenbroch, 2002) to relationship satisfaction (Rusbult & Buunk, 1993).

Study 1 Method

Participants

Prior to contacting potential participants, institutional review board (IRB) approval was secured (Protocols 17102605 and CH17102605). In keeping with a method utilized in previous studies (Sacco et al., 2018), we created a list of scientists from various academic disciplines (e.g., biology, medicine, neuroscience, psychology) from

National Institutes of Health (NIH) and National Science Foundation (NSF) public websites that listed names and contact information for approximately 5,000 currently federally funded U. S. researchers. We focused on principal investigators (PIs) funded by these two agencies to represent the disciplines that are most commonly discussed in the literature on QRPs.

After creating a listserv to solicit participation, PIs were contacted by email and encouraged to participate through a hyperlink to the research materials on Qualtrics. Waves of emails to approximately 200 prospective respondents were dispersed every other day over the course of approximately 1.5 months. More specifically, we drew from our list of emails and created waves of approximately 200 participants by randomly selecting prospective respondents from the list while trying to equalize the number of men and women per wave. An a priori power analysis indicated that 200 participants would suffice to detect medium-sized effects (Cohen's $f = 0.25$, $\beta = 0.80$), although we deliberately oversampled to account for possible attrition. Approximately 4,200 researchers were solicited via email, and 287 participants completed the study (7% completion rate; $M_{age} = 45.84$, $SD = 9.72$; 139 men, 139 women, 9 not reporting gender; 79.4% White, 8% Asian, 5% Hispanic, 1% Black, 3.8% Other) and were compensated with an Amazon gift card worth US$25.00. Participants reported having $M = 19.71$ years of experience ($SD = 9.40$).

Interventions. After giving informed consent, participants were randomly assigned to read one of five short (109–274 word) initial statements about research ethics. Those who were in the control condition read the federal definition of research misconduct as falsification, fabrication, and plagiarism (FFP; $n = 56$). One experimental intervention described the impact of QRPs on science ($n = 56$), emphasizing their adverse effects on replicability, false positive findings, and what is called "the reproducibility crisis." A second intervention, which constituted an "anti-rationalization" measure, emphasized recent efforts to improve the reliability of science such as data sharing, study preregistration, increased transparency ($n = 58$). A third intervention appealed to participants' desire to maintain consistency between their professional self-concepts and ethical identities ($n = 56$). The fourth intervention targeted researchers' presumed commitment to ethical research, given their status as federally funded researchers ($n = 61$; see additional file S1 for complete statements).

Following the initial statements, participants were asked 17 reaction questions assessing their attitudes toward the statement they had read. (The reaction questions were generated ad hoc as face-valid items for purposes of Study 1 by the second and third author.) Responses to the questions were given on 7-point Likert-type scales

Table 1. Study 1 Reaction Questions: Please Respond to the Following Items Regarding What You Just Read Using the Scale Provided Below.

Not at all						Very much
1	2	3	4	5	6	7

1. I was aware of the information presented.
2. I believe this information to be true.
3. I think questionable research practices are a genuine problem in my field.
4. I am willing to take steps outlined in this paragraph to prevent questionable research practices.
5. I feel motivated to prevent questionable research practices after reading this information.
6. I believe all scientists should follow the guidelines outlined in this paragraph.
7. I feel more informed about the current culture of scientific research ethics.
8. Based on information presented here, I would probably feel more comfortable reporting accurate data, even if they are not ideal.
9. Based on information presented here, I feel empowered to conduct my research in an ethical manner.
10. I feel that being an ethical researcher is integral to my identity as a scientist.
11. Based on reading this information, I do not feel there are many justifications for researchers to massage their data for significance in today's climate.
12. This information makes me feel motivated to restore the public's trust in science.
13. Based on the information presented here, I feel questionable research practices have a considerable impact on the field.
14. How positive do you feel about this message?
15. How negative do you feel about this message?
16. How encouraged do you feel after reading this message?
17. How motivating is this message to encourage scientific research?
What do you think should be done to reduce questionable research practices in your field? This could include what you want out of your institution, journals, or granting agencies.

(1 = *not at all*; 7 = *very much*), with higher scores indicating more agreement with each statement. This was followed by an open-ended response opportunity in which participants were asked to write about what they recommend to reduce QRPs in their respective fields (questions and writing prompt are presented in Table 1).

QRP defensibility and willingness. Participants then indicated the extent to which they found various QRPs ethically defensible and the extent to which they would be willing to engage in each of them. Specifically, participants evaluated 31 QRPs, presented in random order, deemed either unambiguously unethical (UU-QRP) or ambiguously unethical (AU-QRP) as determined and validated in previous research (Sacco et al., 2018). UU-QRPs consisted of 16 items, such as concealing data or results that contradict one's own previous research; AU-QRPs consisted of 15 items, such as deciding whether to include or exclude data after looking at the impact of doing so on the results. Responses were collected on 7-point Likert-type scales for both defensibility (1 = *completely indefensible*; 7 = *completely defensible*) and willingness (1 = *completely unwilling to engage in this behavior*; 7 = *completely willing to engage in this behavior*). Both the defensibility (UU-QRP α = .85; AU-QRP α = .88) and willingness components (UU-QRP = .80; AU-QRP = .87) attained acceptable reliabilities using Cronbach's alphas (see additional file S2 for QRP questionnaire).

Motives. Also following the protocol used in previous research, participants responded to a "motives questionnaire" consisting of a series of questions about why they would engage in the previously described QRPs were they to do so. In particular, participants evaluated the impact of QRPs on science (three items, α = .90), their ability to rationalize their behavior were they to engage in QRPs use (three items, α = .74), and perceptions of the risks involved with QRP use (six items, α = .81; see additional file S3 for Motives questionnaire). Participants then answered several demographics questions regarding their age, gender, race/ethnicity, academic field, years in the field, and career extramural funding received before being debriefed. Finally, participants were offered the opportunity to receive compensation for their participation by providing an email address from which they would receive their e-gift card(s). Email addresses were collected via a separate Qualtrics link to dissociate participants' responses from identifying information.

Study 1 Results

General Reactions to the Interventions Statements

Participants' reactions to 17 questions about the four intervention conditions and control were analyzed initially using a multivariate analysis of variance (MANOVA). In terms differences between conditions for the main effects,

Table 2. Mean Scores (and Standard Deviations) for Reaction Items for Each Intervention.

Reaction	Impact	Anti-rationalization	Consistency	Commitment	Control	p value
Aware	5.38 (1.33)[a]	4.91 (1.85)[a]	6.45 (0.86)[b]	6.16 (1.28)[b]	6.46 (0.90)[b]	<.001
Believe	5.57 (1.43)[a]	5.46 (1.24)[a]	5.75 (1.49)[ab]	6.16 (1.11)[a]	6.08 (1.33)[a]	.023
Problem	4.96 (1.50)[a]	4.61 (1.54)[a]	4.75 (1.50)[a]	4.70 (1.47)[a]	3.88 (1.51)[b]	.006
Steps	5.87 (1.24)[a]	5.19 (1.45)[b]	6.13 (1.07)[a]	6.33 (1.00)[a]	5.96 (1.57)[a]	<.001
Motivated	5.36 (1.60)[a]	5.06 (1.47)[a]	5.57 (1.51)[a]	5.67 (1.27)[a]	4.96 (1.71)[a]	.070
All science	6.00 (1.27)[a]	5.09 (1.48)[b]	6.38 (0.98)[a]	6.46 (0.94)[a]	6.46 (1.20)[a]	<.001
Informed	4.30 (1.71)[a]	4.54 (1.65)[a]	3.83 (1.91)[a]	4.21 (1.64)[a]	3.60 (1.62)[a]	.046
Empowered	5.23 (1.85)[a]	4.91 (1.58)[a]	5.45 (1.46)[a]	5.37 (1.55)[a]	5.15 (1.82)[a]	.476
Identity	6.70 (0.97)[a]	6.72 (0.56)[a]	6.74 (0.83)[a]	6.77 (0.53)[a]	6.81 (0.57)[a]	.939
Massage	4.87 (1.95)[a]	4.43 (1.63)[a]	5.06 (1.82)[a]	5.23 (1.71)[a]	5.29 (1.58)[a]	.081
Restore	4.98 (1.77)[a]	4.81 (1.52)[a]	4.77 (1.77)[a]	4.86 (1.48)[a]	4.92 (1.62)[a]	.970
Impact	5.42 (1.28)[a]	4.93 (1.40)[a]	5.17 (1.49)[a]	4.82 (1.50)[ab]	4.27 (1.82)[c]	.003
Positive	3.62 (1.91)[a]	5.09 (1.32)[b]	4.98 (1.73)[b]	5.33 (1.34)[b]	4.58 (1.67)[b]	<.001
Negative	4.25 (1.85)[a]	2.63 (1.36)[b]	2.62 (1.57)[b]	2.44 (1.29)[b]	2.69 (1.54)[b]	<.001
Encouraged	3.55 (1.61)[a]	4.76 (1.13)[b]	4.64 (1.62)[b]	4.74 (1.19)[b]	4.08 (1.47)[b]	<.001
Motivating	4.08 (1.68)[a]	4.56 (1.32)[a]	4.62 (1.82)[a]	4.63 (1.53)[a]	4.21 (1.40)[a]	.220

Note. Superscripts indicate the degree of differences between conditions in a Bonferroni corrected post hoc test with the same letter indicating no difference between conditions.

participants in the impact and anti-rationalization conditions reported less awareness of QRPs than did the other three conditions. Compared with the control condition, participants in all four intervention conditions perceived QRPs to be more of a problem and more negatively impactful on research. Participants in the anti-rationalization condition were less willing to take steps to reduce QRPs compared with the other conditions and were the least interested in having all science follow proposed guidelines. The impact condition also made scientists feel the least positive and most negative about science compared with the other conditions (see Table 2 for the descriptive statistics for each item by condition, the significance levels for each individual analysis of variance [ANOVA], and the differences in the post hoc analyses; Bonferroni corrections were used for post hoc analyses to minimize the chance of Type I error for main effects).

Defensibility

We conducted a 5 (Condition: Impact, Anti-Rationalization, Consistency, Commitment, Control) $\times$ 2 (QRP: UU vs. AU) mixed-model custom ANCOVA with repeated factors over the latter factors and used number of years in one's respective field as a continuous moderator to test for interactive effects with the categorical predictors. A QRP main effect indicated that participants found AU-QRPs ($M =$ 3.01, $SD =$ 1.01) more defensible than UU-QRPs ($M =$ 1.61, $SD =$ 0.57), $F(1, 272) = 170.92, p < .001, \eta^2 =$ 0.386. No other main effects emerged for categorical predictors or interactions, Fs < 1.20, ps > .300.

Willingness

We conducted another 5 $\times$ 2 mixed-model custom ANCOVA for willingness. A QRP main effect indicates that participants found AU-QRPs ($M = 2.63, SD = 0.99$) more defensible than UU-QRPs ($M = 1.41, SD = 0.46$), $F(1, 272) = 133.86, p < .001, \eta^2 = 0.329$. No other main effects for categorical predictors or interactions emerged, Fs < 0.55, ps > .466.

Motives Questionnaires

Initially, we conducted a one-way custom MANCOVA for the three motives questionnaires as an omnibus analysis to reduce the family-wise error rate through a single omnibus analysis. In the instance of a significant interaction, we decomposed the interaction utilizing separate subordinate linear regressions to compare the efficacy of each intervention as a function of age. For each decomposition, we utilized the control condition as the comparison group to test each intervention against baseline in a simple slope analysis tested using *Interaction* (Soper, 2013), resulting in four post hoc analyses for a measure.

Rationalization. A main effect of Condition emerged for rationalization, $F(4, 258) = 2.64, p = .034, \eta^2 = 0.039$. Post hoc least significant difference (LSD) analyses indicated that participants in the anti-rationalization condition ($EMM = 3.03, SE = 0.20$) rationalized QRPs the most, which was marginally greater than for the impact condition (estimated marginal means [EMM] = 2.50, $SE = 0.19$), $p =$

.056; no other significant differences emerged in the contrasts ($ps > .200$). Effects for rationalization were qualified by a marginal interaction, $F(4, 258) = 2.37, p = .053, \eta^2 = 0.035$. We compared the efficacy of each intervention against the control for researchers whose careers are 1 standard deviation above and below the mean (i.e., comparing early- vs. later-career researchers). Compared with the control, later-career researchers reported marginally less rationalization toward QRPs in the Impact Condition ($b = 0.17$, $SE = 0.09, p = .061$); no such difference emerged in early-career researchers ($b = -0.13, SE = 0.09, p = .144$). Early-career researchers reported less rationalization toward QRPs in the anti-rationalization condition compared with those in the control condition ($b = -0.37, SE = 0.12, p = .004$), whereas no difference emerged for later-career researchers ($b = 0.15, SE = 0.13, p = .224$). Early-career researchers in the consistency bias condition reported marginally lower rationalization compared with those in the control condition ($b = -0.34, SE = 0.19, p = .086$), whereas no difference emerged for later-career researchers ($b = 0.24, SE = 0.20, p = .235$). No differences emerged comparing the commitment condition to the control condition for early- or later-career researchers ($ps > .154$).

Risk. A main effect of condition emerged for risk, $F(4, 258) = 2.76, p = .028, \eta^2 = 0.041$. However, post hoc LSD tests indicated no differences between conditions ($ps > .270$). Effects for risk were also qualified by a significant interaction, $F(4, 258) = 3.13, p = .015, \eta^2 = 0.046$. Early-career researchers reported perceptions of less risk for QRPs in the control condition compared with the impact condition ($b = 0 .14, SE = 0.07, p = .054$), whereas no difference emerged for later-career researchers ($b = -0.08$, $SE = 0.07, p = .260$). Later-career researchers perceived QRPs marginally riskier in the anti-rationalization condition ($b = -0.19, SE = 0.10, p = .065$), whereas early-career researchers found QRPs marginally less risky in the anti-rationalization condition ($b = 0.18, SE = 0.10, p = .065$). No differences emerged for early- or later-career researchers for the consistency bias condition ($ps > .320$). Early-career participants perceived QRPs as riskier in the commitment condition ($b = 0.72, SE = 0.30, p = .018$); no difference emerged for later-career researchers ($b = -0.29$, $SE = 0.30, p = .328$).

Impact. No main effect or interaction emerged for impact, thus prompting us to consider it no further, $Fs = 0.54$, $p > .450$.

Discussion

Compared with control, responses to all four interventions resulted in perceptions of QRPs as having greater negative impact on science. The results confirmed previous findings about ethical distinction between AU- and UU-QRPs, but the interventions themselves elicited no significant change in attitudes regarding QRP defensibility or willingness. Participants in the anti-rationalization condition expressed the least willingness to take steps to reduce QRPs use, indicating a potential defensive "backfire" effect of the anti-rationalization condition (Bohner et al., 2002). This interpretation is consistent with the fact that participants in the anti-rationalization condition indicated a marginally greater tendency to rationalize QRP use, as measured on the motives questionnaire.

The two interventions that showed the greatest promise for changing researcher attitudes in the desired direction were the consistency and commitment statements. Compared with other conditions, early-career researchers in the former condition indicated less rationalization, and early-career researchers in the latter condition found QRPs riskier. Although the effects were modest, these aspects of the results suggested modifying study procedures in three ways. First, given the efficacy of the consistency bias statement in reducing QRP endorsement, we focused primarily on that intervention's efficacy in Study 2. Second, the consistency and control statements were effectively streamlined and refined in an attempt to amplify their effect; that is, because of the promise of the consistency bias manipulation, we sought to modify the wording to ensure it could most optimally communicate encouragement to foster consistency between one's own identity and research ideals. This would afford us the opportunity to replicate and extend to the most robust finding of the previous study to determine how effective evoking a consistency bias would be. That is, whereas Study 1 identified an intervention with *clinical* significance in influencing attitudes, Study 2 sought to identify a procedure that could elicit greater *statistical* significance. Finally, given that UU-QRPs were largely considered more aversive than AU-QRPs across conditions, we focused entirely on AU-QRPs in Study 2.

Study 2 Method

Participants

After securing IRB approval for modifications, NIH- and NSF-funded researchers were invited to participate in exchange for US$5.00 in Amazon gift cards, as the length of this study was truncated considerably relative to Study 1. A total of 145 participants completed the survey with similar disciplinary representation as before (80 men, 58 women; $M_{age} = 49.92, SD = 10.80$; 85.5% White). The same recruitment process used in Study 1 occurred over the course of a month; to prevent practice effects, participants from Study 1 were not recontacted for Study 2. An a priori power analysis indicated that completed responses from 130 participants

would suffice to detect effects ($f^2 = 0.06$, $\beta = 0.80$). Respondents had spent an average of $M = 23.45$ years ($SD = 10.22$) in their respective fields.

Intervention. After giving informed consent, participants were randomly assigned to read one of two statements about research ethics. In one condition, participants read a statement, modified from Study 1, again designed to appeal to their presumed desire to maintain consistency between their professional self-concepts and ethical identities ($n = 69$). The other condition consisted of the same control (the federal definition of research misconduct; $n = 76$) used in Study 1 (See additional file S4 for complete statements).

QRP defensibility, willingness, and motives. As in Study 1, participants then indicated the extent to which they perceived AU-QRPs, presented in random order, as ethically defensible ($\alpha = .77$) and they indicated their willingness to engage in them ($\alpha = .78$). Participants then completed the motives questionnaire ($\alpha s < .72$) and previously-used demographics questions. Same as before, participants were then offered the opportunity to receive compensation for their participation by providing their email address (collected via a separate Qualtrics link) from which they would receive their e-gift card(s).

Study 2 Results

We submitted our data to Model 1 of PROCESS (Hayes, 2013) to test for moderation of years in the field to determine the efficacy of the intervention while considering the possibility for an interaction. A main effect of Condition emerged, such that participants in the consistency condition reported reduced perception of QRPs ($M = 3.05$, $SD = 0.83$) as defensible compared to the control condition ($M = 3.15$, $SD = 0.87$), $b = 0.78$, $SE = 0.36$, $p = .03$. However, effects were qualified by a significant interaction, $\Delta R_{\dot{c}}^2 = 0.03$, $b = -.03$, $SE = 0.01$, $p = .04$). Simple slope analyses indicate that experienced researchers ($+1$ SD years in the field) in the consistency condition did not differ from those in the control condition, $b = -0.20$, $SE = 0.20$, $p = .33$. Conversely, early-career researchers (-1 SD years) in the consistency bias condition reported marginally lower defensibility of QRPs than did those in the control, $b = 0.39$, $SE = 0.20$, $p = .054$. Viewed another way, in the control condition, more experienced researchers indicated less defensibility of QRPs than did less experienced researchers, $b = -0.03$, $SE = 0.01$, $p < .01$. No difference emerged in QRP defensibility, as a function of experience, in the consistency bias condition, $b = 0.00$, $SE = 0.01$, $p = .90$.

No conventionally significant main effects or interactions emerged for willingness or the three motives scales,

and we considered them no further ($p s > 0.053$). These data provide evidence of efficacy in utilizing ethical identity–based interventions to reduce QRP endorsement, particularly when fostering an ethical identity and particularly for early-career scientists (see Figure 1 for a graphical representation of effects).

Discussion

A salient finding in Study 2 is the overall impact of the consistency intervention at reducing researchers' attitudes toward the defensibility of ethically ambiguous QRPs. While the effectiveness of the intervention was rather modest, the findings nonetheless suggest some potential for ethical identity-affirming messages and the psychological appeal of adhering to professional ethical commitments as a means of changing the endorsement of borderline ethical practices in science.

A second notable result is the discrepancies in researchers' responses depending on their experience. While the consistency intervention had no significant impact on more experienced researchers, once adjusted for margin of error, it predicts a meaningful difference in reactions from less experienced scientists.

It is interesting to consider these experience-related differences in terms of conflicting indications regarding career stage and QRP use in the literature. Previous findings suggest that mid-career scientists are more likely to admit to having engaged in various QRPs than early-career researchers (Martinson et al., 2005), a finding that can be interpreted in multiple ways. Conversely, a more recent survey on researchers in health professions education found more QRP engagement by younger researchers (Maggio et al., 2019). This is more along the lines of what one might expect: Younger researchers have relatively more to gain and less to lose by skirting the margins of ethicality, an argument supported by mathematical models (Lacetero & Zirulia, 2011). Early-career researchers may be comparatively less committed to norms of good science and possess a less acute grasp of those norms. They are also more likely to be the authors of a retracted paper (Fanelli et al., 2015). Increased susceptibility of younger researchers to QRPs is also supported by recent findings from a meta-analysis which were consistent with the hypothesis that there is a positive association between early-career status and the risk that a scientist is reporting overestimated effect sizes (Fanelli et al., 2017).

A few limitations should be mentioned in connection with this research. Self-selection bias is always a possibility in research designs of the sort used in these two studies. Also important to recognize is the possibility of important disciplinary differences in attitudes and perceptions. Ethical concepts and descriptions of QRPs can have importantly different shades of importance and meaning depending on
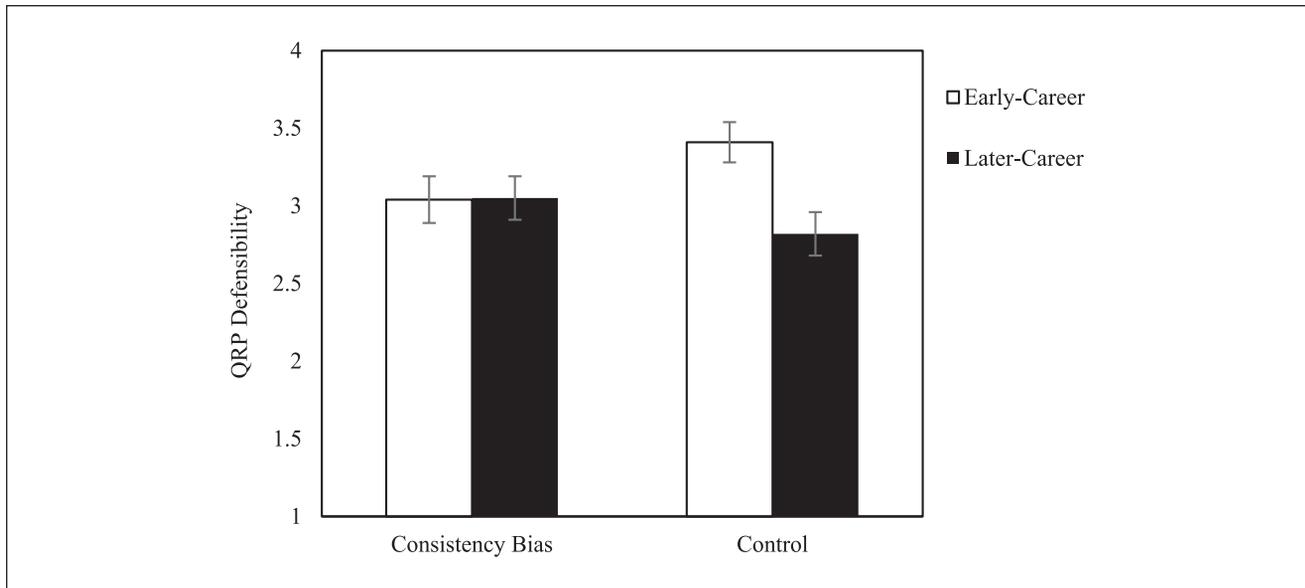
Figure 1. Comparison of consistency intervention and control as a function career experience, Study 2.
Note. Although later-career researchers' perceived defensibility of QRPs appears lower in the control condition, such differences are not statistically significant when considering the standard errors.

the kind of research a scientist does and the field in which he or she works.

In conclusion, it must be acknowledged that mere brief statements alone are unlikely to have a lasting impact on attitudes and behavior. The effectiveness of other ethical identity-affirming interventions has been mixed, but benefits tend to be greater and more lasting in contexts in which the changes in attitude are reinforced through various means and supported with adequate resources (Cohen & Sherman, 2014). This conclusion coheres with information from vaccination contexts. According to the recent comprehensive review of various strategies to increase vaccination rates, single interventions implemented individually almost always produce only small effects (Brewer et al., 2017). For larger effects, multiple strategies implemented together are more promising. Nonetheless, our data suggest that psychological interventions of the sort tested here have some promise as part of systematic efforts to improve research integrity.

Best Practices

Researcher attitudes toward QRPs and willingness to engage in them are almost certainly shaped by many factors: individual psychology, educational training, work environment, social context, and systemic incentives and structures. Many structural reforms and institutional initiatives currently under way are promising, but they are in their infancy and the desirability and possibility of further improvements and developments is all but assured. Strategies like those discussed above could be used in a

variety of ways and settings, and to the extent they prove effective, their greatest benefit is likely to occur when repeatedly reinforced and integrated with other measures.

Research Agenda

The extant literature on encouraging ethical behavioral through affirming a positive sense of individuals' identity shows mixed results. In some circumstances, consistency-like interventions can have a moral licensing or credentialing effect, though on other occasions, they can be useful. To our knowledge, ours is the first study to attempt a consistency intervention as it relates to perceptions of QRPs. As in other domains, the psychology involved is likely subtle and merits further study. Therefore, there is a need for more research into why and under what conditions consistency interventions promote research integrity.

Educational Implications

In light of the promising impact of the consistency intervention in Study 2 on less experienced researchers, if used repeatedly and successfully integrated into formal and informal aspects of graduate and early-career training and mentoring, they could be part of improved research integrity education practices. Similar interventions have demonstrated short-term efficacy in reducing early-career researchers' endorsement of QRPs through an in-person training module for graduate students by encouraging participants to foster attitudinal consistency between their identity as a research scientist and their ideals of conducting research with integrity

(Sacco & Brown, 2019). They could be used for example for written assignments, group journal club contexts, and research methods readings. Given the importance of message repetition in facilitating long-lasting attitudinal change (Cacioppo & Petty, 1979), written prompts of the sort tested in these studies are likely to have the greatest positive educational impact when used repeatedly and when the underlying pro-integrity message is conveyed through both face-to-face and written means of communication.

Acknowledgments

Declaration of Conflicting Interests

Funding

ORCID iDs

Samuel V. Bruton https://orcid.org/0000-0002-5455-125X

Donald F. Sacco https://orcid.org/0000-0001-6017-5070

Supplemental Material

Supplemental material for this article is available online. Raw data available from the corresponding author on request.

References

3rd World Conference on Research Integrity. (2013). *Montreal statement on research integrity in cross-boundary research collaborations*. https://wcrif.org/montreal-statement/file

Agnoli, F., Wicherts, J. M., Veldkamp, C. L. S., Albiero, P., & Cubelli, R. (2017). Questionable research practices among Italian research psychologists. *PLOS ONE*, *12*, Article e0172792.

Ariely, D. (2013). *The (honest) truth about dishonesty: How we lie to everyone—Especially ourselves*. HarperCollins.

Ariely, D., & Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological Science*, *13*, 219–224.

Bastardi, A., Uhlmann, E. L., & Ross, L. (2011). Wishful thinking: Belief desire, and the motivation evaluation of scientific evidence. *Psychological Science*, *22*, 731–732.

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, *116*, 116–126.

Bem, D. J. (1972). Self-perception theory. In M. Zanna & J. Olson (Eds.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). Academic Press.

Bohner, G., Ruder, M., & Erb, H. P. (2002). When expertise backfires: Contrast and assimilation effects in persuasion. *British Journal of Social Psychology*, *41*, 495–519.

Bouter, L. M. (2015). Commentary: Perverse incentives or rotten apples? *Accountability in Research*, *22*(3), 148–161.

Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & Ter Riet, G. (2016). Ranking major and minor research misbehaviors: Results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review*, *1*, Article 17.

Brewer, N. T., Chapman, G. B., Rothman, A. J., Lesak, J., & Kempe, A. (2017). Increasing vaccination: Putting psychological science in action. *Psychological Science in the Public Interest*, *18*(3), 149–207.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376.

Cacioppo, J. T., & Petty, R. E. (1979). Effects of message repetition and position on cognitive response, recall, and persuasion. *Journal of Personality and Social Psychology*, *37*(1), 97–109.

Casadevall, A., & Fang, F. C. (2018). Making the scientific literature fail-safe. *Journal of Clinical Investigation*, *128*(10), 4243–4244.

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Obstetrics & Gynecology*, *114*, 1341–1345.

Chiu, K., Grundy, Q., & Bero, L. (2017). "Spin" in published biomedical literature: A methodological systematic review. *PLOS Biology*, *15*(9), Article e2002173.

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*, 333–371.

Conway, P., & Peetz, J. (2012). When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin*, *38*, 907–919.

Crutzen, R., & Peters, G.-J. Y. (2017). Targeting next generations to change the common practice of underpowered research. *Frontiers in Psychology*, *8*, Article 1184.

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLOS ONE*, *4*, Article e5738.

Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, *90*, 891–904.

Fanelli, D. (2013). Only reporting guidelines can save (soft) science. *European Journal of Personality*, *27*, 124–125.

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*(6), Article e0127556

Fanelli, D., Costas, R., & Ioannidis, J. P. A. (2017). Meta-assessment of bias in science. *Proceedings of the National*

Academy of Sciences of the United States of America, *114*, 3714–3719.

Feynman, R. P. (1974). *Cargo cult science: Caltech's 1974 commencement address*. http://calteches.library.caltech.edu/51/2/CargoCult.htm

Fischhoff, B. (1975). Hindsight ≠ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Fong, E. A., & Wilhite, A. W. (2017). Authorship and citation manipulation in academic research. *PLOS ONE*, *12*(12), Article e0187394.

Gerrits, R. G., Jansen, T., Mulyanto, J., van den Berg, M. J., Klazinga, N. S., & Kringos, D. S. (2019). Occurrence and nature of questionable research practices in the reporting of messages and conclusions in international scientific Health Services Research publication: A structured assessment of publications authored by researchers in the Netherlands. *BMJ Open*, *9*, Article e027903.

Gini, F., & Ariely, D. (2012). The dark side of creativity: Original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, *102*, 445–459.

Gino, F., Norton, M. I., & Weber, R. A. (2016). Motivated Bayesians: Feeling moral while acting egoistically. *Journal of Economic Perspectives*, *30*, 189–212.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*, 472–482.

Hayes, A. F. (2013). *Methodology in the social sciences: Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Herron, D. M. (2012). Is expert peer review obsolete? A model suggests that post-publication reader review may exceed the accuracy of traditional peer review. *Surgical Endoscopy*, *26*, 2275–2280.

Ikeda, A., Xu, H., Fuji, N., Zhu, S., & Yamada, Y. (2020). Questionable research practices following pre-registration. *Japanese Psychological Review*, *62*, xxx-xxx. Preprint doi: 10.31234/osf.io/b8pw9

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, *2*, Article e124.

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

Joynson, C., & Leyser, O. (2015). The culture of scientific research. *F1000 Research*, *4*, Article 66.

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., . . . Errington, T. M. (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. *PLOS Biology*, *14*, Article e1002456.

Lacetero, N., & Zirulia, L. (2011). The economics of scientific misconduct. *Journal of Law, Economics, and Organization*, *27*, 568–603.

Maggio, L. A., Dong, T., Driessen, E. W., & Artino, A., Jr. (2019). Factors associated with scientific misconduct and questionable research practices in health professions education. *Perspectives on Medical Education*, *8*, 74–82.

Martinson, B. C., Anderson, M. S., & de Vries, R. (2005). Scientists behaving badly. *Nature*, *435*, 737–738.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*, 633–644.

Mazar, N., & Ariely, D. (2015). Dishonesty in scientific research. *Journal of Clinical Investigation*, *125*, 3993–3996.

Motyl, M., Demos, A. P., Carsel, T. S., Hanson, B. E., Melton, Z. J., Mueller, A. B., Prims, J. P., Sun, J., Washburn, A. N., Wong, K. M. & Yantis, C. (2017). The state of social and personality science: Rotten to the core, not so bad, getting better, or getting worse? *Journal of Personality and Social Psychology*, *113*, 34–58.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., . . . Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behavior*, *1*, 1–9.

Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, *29*, 85–95.

Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin*, *34*, 913–923.

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 2600–2602.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*, 615–631.

Nuzzo, R. (2015). How scientists fool themselves—And how they can stop. *Nature*, *526*, 182–185.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, Article 943.

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, *37*, 1915–1926.

Redlawsk, D. P., Civettini, A. J., & Emmerson, K. M. (2010). The affective tipping point: Do motivated reasoners ever "get it"? *Political Psychology*, *31*, 563–593.

Redman, B. K., & Caplan, A. L. (2016). Limited reproducibility of research findings: Implications for the welfare of research participants and considerations for institutional review boards. *IRB: Ethics & Human Research*, *38*, 8–10.

Resnik, D. B., Neal, T., Raymond, A., & Kissling, G. E. (2015). Research misconduct definitions adopted by U.S. research institutions. *Accountability in Research*, *22*, 14–21.

Rusbult, C. E., & Buunk, B. P. (1993). Commitment processes in close relationships: An interdependence analysis. *Journal of Social and Personal Relationships*, *10*, 175–204.

Sacco, D. F., Brown, M., & Bruton, S. V. (2018). Grounds for ambiguity: Justifiable bases for engaging in questionable research practices. *Science and Engineering Ethics*, *25*, 1321–1337.

Sacco, D. F., & Brown, M. (2019). Assessing the efficacy of a training intervention to reduce acceptance of questionable research practices in psychology graduate students. *Journal of Empirical Research on Human Research Ethics*, *14*, 209–218.

Sacco, D. F., Bruton, S. V., & Brown, M. (2018). In defense of the questionable: Defining the basis of research scientists' engagement in questionable research practices. *Journal of Empirical Research on Human Research Ethics*, *13*, 101–110.

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., & Kassam, K. S. (2015). Bias blind sport: Structure, measurement, and consequences. *Management Science*, *61*, 2468–2486.

Sheeran, P., Harris, P. R., & Epton, T. (2014). Does heightening risk appraisals change people's intentions and behavior? A meta-analysis of experimental studies. *Psychological Bulletin*, *140*, 511–543.

Shu, L. L., Mazar, N., Gino, F., Ariely, D., & Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 15197–15200.

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

Soper, D. S. (2013). *Interaction: Windows software for graphing and analyzing statistical interactions* [computer software]. https://www.danielsoper.com/Interaction/

Tijdink, J. K., Verbeke, R., & Smulders, Y. M. (2014). Publication pressure and scientific misconduct in medical scientists. *Journal of Empirical Research on Human Research Ethics*, *9*, 64–71.

Tsang, J. (2002). Moral rationalization and the integration of situational factors and psychological processes in immoral behavior. *Review of General Psychology*, *6*, 25–50.

Vazire, S. (2016). Editorial. *Social Psychological and Personality Science*, *7*, 3–7.

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. (2011). Why psychologists must change the way they analyze their data: The case of psi. *Journal of Personality and Social Psychology*, *100*, 426–432.

Ware, J. J., & Munafò, M. R. (2015). Significance chasing in research practice: Causes, consequences, and possible solutions. *Addiction*, *110*, 4–8.

Washburn, A. N., Hanson, B. E., Motyl, M., Skitka, L. J., Yantis, C., Wong, K. M., . . . Carsel, T. S. (2018). Why do some psychology researchers resist adopting proposed reforms to research practices? A description of researchers' rationales. *Advances in Methods and Practices in Psychological Science*, *1*, 166–173.

Wislar, J. S., Flanagin, A., Fontanarosa, P. B., & DeAngelis, C. D. (2011). Honorary and ghost authorship in high impact biomedical journals: A cross sectional survey. *British Medical Journal*, *343*, Article d6128.

World Medical Association. (2013). *World medical association declaration of Helsinki ethical principles for medical research involving human subjects*. https://www.wma.net/wp-content/uploads/2016/11/DoH-Oct2013-JAMA.pdf

Author Biographies

Samuel V. Bruton is a professor at The University of Southern Mississippi in the School of Humanities and director of USM's Office of Research Integrity. He specializes in applied ethics, ethical theory, and the philosophy of law. His role in preparation of the manuscript included corecipient of grant funding for the project, study design, drafting most of the text, interpreting results.

Mitch Brown is a postdoctoral fellow at Fairleigh Dickinson University in the School of Psychology. His primary focus in research is social perception, including an interest in perceptions of others' ethical behaviors. His role in preparation of the manuscript included study design, statistical analysis, drafting of the results, interpreting results, and review of the manuscript.

Donald F. Sacco is an associate professor at The University of Southern Mississippi in the School of Psychology. He is a social psychologist who specializes in research ethics and motivation. His role in preparation of the manuscript included corecipient of grant funding for the project, study design, interpreting results, and review of the manuscript.