

Why Do Some Psychology Researchers Resist Using Proposed Reforms to Research Practices?

A Description of Researchers' Rationales

Anthony N. Washburn

Brittany E. Hanson

Matt Motyl

Linda J. Skitka

Caitlyn Yantis

Kendal M. Wong

Jiaqing Sun

J. P. Prims

Allison B. Mueller

Zachary J. Melton

Timothy S. Carsel

University of Illinois at Chicago

In press at *Advances in Methods and Practices in Psychological Science*

Abstract

In response to the “replication crisis” many psychologists recommended that the field adopt several proposed reforms to research practices (PRRPs), like preregistration, to make research more replicable. However, reception of PRRPs is not well known. We wanted to know the rationales researchers had for not using them. We analyzed data of 1,035 researchers in social and personality psychology from Motyl et al. (2017) who asked them to explain when they thought it was acceptable to not use four specific PRRPs: preregistering hypotheses/methods, making data publicly available online, conducting formal power analyses, and reporting effect sizes. Our results suggest that (a) adoption and use of PRRPs is quite varied and (b) rationales for not using them reflect a need for more discussion and education about the utility and feasibility of the proposed reforms.

Keywords: replicability, research practices, methodology, meta-science, questionable research practices

Why Do Some Psychology Researchers Resist Using Proposed Reforms to Research Practices?

A Description of Researchers' Rationales

A recent string of heavily disseminated findings suggests that several published results in cognitive, social, and personality psychology fail to replicate (Open Science Collaboration, 2015). One explanation for the “replication crisis” is that many researchers engage in *questionable research practices* (QRPs; Simmons, Nelson, & Simonsohn, 2011). For example, in recent surveys, many psychologists admitted to deliberately engaging in QRPs including stopping data collection upon finding significant results and reporting unexpected findings as hypothesized a priori (i.e., hypothesizing after the results are known, or HARKing; Fiedler & Schwarz, 2016; John, Loewenstein, & Prelec, 2012; Motyl et al., 2017). As a response to these revelations and others regarding the tenuous nature of replicability in psychological science, leaders in the field created the Society for Personality and Social Psychology (SPSP) Presidential Task Force on Publication and Research Practices to promote, among other goals, researchers' use of proposed reforms to research practices (PRRPs) such as conducting formal power analyses and reporting effect sizes (Funder et al., 2014). Similarly, the Center for Open Science convened a committee of scholars to establish industry standards that would help promote transparency and openness in the social science, including preregistering study hypotheses and making data publicly available online (Transparency and Openness Promotion (TOP) guidelines; Nosek et al., 2015).

What remains unclear, however, is the extent to which researchers are implementing these standards. Initial evidence suggests that, although some have enthusiastically embraced PRRPs (e.g., Center for Open Science, 2013; Eich, 2014; Schweinsberg et al., 2016), others have argued for caution with regard to making certain practices prerequisites for publication, such as

preregistration and making data publicly available online (e.g., Finkel, Eastwick, & Reis, 2015).

For example, some argue that preregistering hypotheses and methods can stifle scientific discovery and may be less applicable to research that aims to replicate large field studies or study populations that are difficult to recruit (e.g., Goldin-Meadow, 2016a, 2016b; Scott, 2013).

Researchers may therefore resist using PRRPs for myriad reasons, ranging anywhere from concerns about practical implementation (e.g., IRB restrictions on sharing potentially identifiable information) to concerns about the utility and need for such practices.

The goal of the current research was to categorize and synthesize researchers' explanations for abstaining from using these practices. Although Motyl et al. (2017) provided some preliminary examples of justifications for not using PRRPs, the justifications provided were not meant to be a complete summary of the most frequent reasons why researchers do not always use these practices. Therefore, one cannot be sure whether the examples provided in Motyl et al. (2017) are representative of where the field stands as a whole or whether there are other rationales that might be critical in understanding how to proceed with the reforms to research practices discussion. The current research used inductive content coding to arrange all of the participant-generated justifications into broader categories. We then were able to take a closer look at how different categories of justifications are more or less prevalent in the lives of actual researchers in the field. In doing so, this research represents an important additional step in the ongoing conversation about how to remediate the replication crisis.

Disclosures

Data, Materials, and Online Resources

Survey materials are available on the Open Science Framework (<https://osf.io/c4xtb>). Due to IRB restrictions, we are not able to publicly share any data from this project. A preprint version of Motyl et al. (2017) is available at <https://psyarxiv.com/qxaby/>.

Measures

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

Subjects

The research reported in this article was conducted under the approval of an institutional review board (protocol #2015-0946).

Conflicts of Interest

The authors declare that they have no conflicts of interest with respect to the authorship or the publication of this article.

Author Contributions

ANW and BEH drafted the manuscript. BEH and ANW analyzed the data. MM conceived of the idea. LJS and MM developed the survey. All authors developed the coding schemes and coded the PRRPs justifications. MM and LJS read, revised, and provided critical feedback on the manuscript. Authors starting with CY are listed in reverse alphabetical order. All authors approved the final version of the manuscript for submission.

Method

Data for this study come from a larger survey about research practices and perceptions of the status of social and personality science (see Motyl et al., 2017, for a description of the full dataset). The prevalence of using each PRRP described in the current paper was originally published in Motyl et al. (2017). However, the acceptability frequencies and inductive content

coding schemes that form the substance of the current research have not been previously published. We present the prevalence of each PRRP alongside our new findings to provide a clear picture of how our participants were thinking about using and not using these research practices.

Sampling and Participants

We contacted the members of the Society for Personality and Social Psychology (6,172 members), European Society for Social Psychology (1,200 members), and the Society of Australasian Social Psychologists (166 members) to participate in our study by sending e-mail solicitations.¹ We attempted to reach as many social and personality psychologists (faculty, postdocs, and graduate students) as possible, and our sample size consisted of those who replied to our solicitation (see Motyl et al., 2017 for more detailed information about response rates). All responses to our solicitation were collected between September 29, 2015 and January 11, 2016. Of the 1,414 people who opened the survey, between 1,016 and 1,035 responded to the relevant items related to the acceptability of PRRPs.

Of those who provided background information, most identified primarily as social (78%) or personality psychologists (8%). Our sample was 49% male and 47% female (the remainder preferred not to say; for reference SPSP's gender breakdown is 40% male and 55% female, see SPSP, 2016). Thirty-two percent of our sample were graduate students, 14% non-tenure track post-degree holders (e.g., adjuncts, post-docs), 19% assistant professors, 15% associate professors, and 20% full professors or above. Sixty-nine percent of our sample were

¹ We also contacted the Asian Association of Social Psychology, but they declined to distribute the survey to their members.

affiliated with a public university, 30% with a private university, and 1% did not have a university affiliation.

Measures

Prevalence of PRRPs. Among other questions, Motyl et al. (2017) asked researchers how often they used four PRRPs (i.e., reporting effect sizes, preregistering hypotheses, making data publicly available online, and conducting formal a priori power analyses) with a 5-point scale (*never, rarely, sometimes, often, always*).

Acceptability of NOT using PRRPs. Motyl et al. (2017) also asked researchers the extent to which they believed that NOT engaging in each of these PRRPs was acceptable with 7-point response scales ranging from *very unacceptable* to *very acceptable*.

Rationale for NOT using PRRPs. Motyl et al. (2017) asked researchers who indicated that NOT using a given PRRP was acceptable (i.e., slightly, moderately, or very acceptable) or that they were uncertain (the midpoint of the scale) to write a rationale for their answer in a provided textbox (e.g., “When is NOT reporting effect sizes an acceptable research practice?”).²

Content Coding

We employed an inductive coding strategy to identify themes in the rationales for not using each PRRP, separately (e.g., Boyatzis, 1998; Epstein & Martin, 2005). First, low-level themes based on similarities across rationales were identified and discussed by the authors as a group. Similar themes were collapsed together into higher-order categories that were specific enough to be distinct yet broad enough to capture sets of ideas contained in the lower-level themes (e.g., Joffe & Yardley, 2004). Anywhere from seven to fifteen discrete themes were

² Due to a programming error with the survey, seventy-one researchers who said that not reporting effect sizes was “moderately acceptable” were not given a text box to write a rationale.

identified in the rationales for each PRRP (e.g., “effect sizes do not affect research validity”, “making data available online could lead to misuse of public data”; see Tables 1 through 4 for substantive themes used to code rationales for not using each research practice). Once the themes were established, two different authors coded each rationale for each practice per the relevant themes. Because researchers in our sample sometimes gave multiple rationales for not engaging in a practice, the authors only coded the first rationale given. Inter-rater reliability in theme coding was acceptable for each practice (making data publicly available: $\kappa = .66$, preregistration: $\kappa = .65$; reporting effect sizes: $\kappa = .66$; conducting formal power analyses: $\kappa = .63$; see Landis & Koch, 1977 for Cohen’s kappa guidelines). Any disagreements in coding were discussed and resolved by the first and second authors.

Results

Prevalence of PRRPs

As originally reported in Motyl et al. (2017), reported use of each of the PRRPs varied considerably. According to researchers’ self-reports, almost all (99%) have reported effect sizes, most (87%) have conducted formal power analyses, just over half of (56%) have made their data publicly available online, and few (27%) have preregistered hypotheses. Not surprisingly, acceptance of not using these PRRPs was inversely related to the prevalence of using such practices (e.g., people were most accepting of not preregistering hypotheses). Although these data do not speak directly to day-to-day usage of these practices, they do at least provide a glimpse of how much researchers are taking these practices into consideration when designing or conducting their research. We next examined the acceptability of and rationales for not using each PRRP.

Preregistering Hypotheses

A majority of researchers (85.38% of our sample for this item; $n = 876$) indicated that not preregistering hypotheses was acceptable or reported being uncertain about the acceptability/unacceptability of this practice. Of these 876 researchers, 73.97% provided a rationale for their answer.

Researchers most commonly stated that not preregistering hypotheses was acceptable when conducting exploratory or descriptive research (e.g., conducting pilot studies, analyzing large scale data sets, conducting qualitative research, and analyzing secondary data analysis). The second most common rationale for not preregistering hypotheses was that there is currently no requirement for preregistration and/or the practice is not necessary. For example, many researchers stated that preregistration is not the norm and preregistering does nothing to increase the validity of the research. The third most common rationale for not preregistering hypotheses was that preregistration is unnecessary for honest, ethical researchers. In other words, if researchers are honest in their methods and results, then preregistering adds nothing to the integrity of the research. The next few most common rationales reflected either uncertainty about how to preregister in the first place or the burdensome and restrictive nature of preregistration (see Table 1 for frequencies and example rationales from each theme).

Table 1
Frequencies and Examples of Rationales for not Preregistering Studies

Rationale	%
Exploratory research -pilot studies, descriptive research, secondary data analysis	33.33
No current requirement and/or does not increase validity -no incentive to do so, not common practice	20.52
Research integrity -creates an atmosphere of distrust, most researchers are trustworthy	7.87
There are no resources and/or do not know how to -guidelines are unclear, journals have no way to preregister	7.87
Transaction costs and/or unnecessary burden -takes too much time, too much bureaucracy	6.48

Constrains researcher degrees of freedom and/or ignores serendipitous findings	5.40
-many interesting findings are unexpected, let the data speak	
Uncodeable	3.40
-unsure or misunderstood the question	
People can still engage in unethical research practices	3.24
-people can preregister after they run studies, can still engage in QRPs	
Concerns about being scooped	2.31
-other researchers may run preregistered studies	
Hypotheses are “registered” else where	2.01
-hypotheses are registered in IRB protocols or grant applications	
Other	1.70
-preregistering will bias the research	
Obvious hypotheses and/or programmatic research	1.23
-theory clearly points to specific hypotheses	
Not preregistering is fine as long as you replicate your findings	1.08
-when results are consistent or robust	
Teaching and/or research training	1.08
-graduate student projects are just to help students learn the process	
Snark	0.93
-preregistration is silly or idiotic	
Evaluation problems with reviewers and/or editors	0.93
-unexpected findings will be stigmatized	
Overvalues confirmatory research	0.62
-unregistered findings are not less valid than registered findings	

Note. “Snark” refers to comments that were sarcastic and did not meaningfully fit into any of the other codes.

Making Data Publicly Available Online

Seven hundred and twenty-six researchers (70.35% of our sample for this item) indicated that not making data available online was acceptable or reported being uncertain about the acceptability/unacceptability of this practice. Of these 726 researchers, 74.38% provided a rationale for their response.

The most common rationale researchers gave for not making data publicly available online was themed around participant confidentiality. Researchers were particularly concerned about protecting sensitive (e.g., organizational) and more easily identifiable (e.g., close relationships) data. The second most common rationale revolved around intellectual property

issues. For example, researchers in our sample reported that they were worried about making data available to other researchers to use before they themselves had analyzed all the publishable data (e.g., getting scooped). The third most common rationale for not making data publicly available revolved around existing data sharing norms. Many in our sample stated that they would share data upon request from other researchers, which justified not sharing data publicly. The remaining rationales revolved around the ideas that making data publicly available just is not necessary, there is no standard way to do so, and/or there are too many institutional and other restrictions (e.g., IRB issues, participant consent; see Table 2 for frequencies and example rationales from each theme).

Table 2
Frequencies and Examples of Rationales for not Making Data Publicly Available Online

Rationale	%
Participant confidentiality issues	27.22
-need to protect participant identity for sensitive data	
Intellectual property issues	16.48
-the researcher is not finished analyzing all the data	
Sharing upon request	15.37
-will share with researchers upon request	
This proposed research practice reform is not needed	8.52
-not sharing data is always okay	
Restrictions on data sharing	8.33
-IRBs will not allow data sharing, participants did not consent	
Lack of education or modeling of the proposed research practice reform	5.93
-it is not the norm, unclear where to post data online	
Uncodeable	4.07
-unsure or misunderstood the question	
Transaction costs	4.07
-it is too time consuming to prepare data for online repository	
Other	3.70
-when data is uninformative, researcher is lazy	
Misuse of public data	2.96
-data are too complex for non-content experts to analyze	
This proposed research practice reform does not actually improve science	2.78
-cannot infer data quality from data posted online	

Snark

-to hide misdeeds

0.56

Note. “Snark” refers to comments that were sarcastic and did not meaningfully fit into any of the other codes.

Conducting Formal A Priori Power Analyses

Six hundred and fourteen researchers (59.90% of our sample for this item) indicated that not conducting an a priori power analysis was acceptable or reported being uncertain about the acceptability/unacceptability of this practice. Of these 614 researchers, 67.92% provided a rationale for their answer.

The most common rationale for not conducting power analyses reflected having no basis to estimate the true effect size in a study. For example, some researchers pointed out that power analyses are not possible when the effect sizes are unknowable (e.g., in new research paradigms or exploratory settings). The second most common theme reflected researcher experience and the ability to guess the appropriate sample size for a study without formal power analysis. For example, many in our sample stated that they already know how large their sample needs to be to find effects in programmatic research. The third most common response consisted of rationales that were uncodeable as a justification for not conducting power analyses. The majority of these rationales reflected uncertainty about why it was acceptable to not conduct power analyses or consisted of people who stated they were confused by the question. The next most common rationale for not conducting power analyses reflected concerns about limited sample sizes and resources. Many researchers stated that they had a finite number of subjects, thus precluding the need for a power analysis that might require more participants than would be possible to recruit. The remaining rationales included that high power is always achieved by having large enough samples or analyzing data where one has no control over the sample size (e.g., secondary data

analysis on publicly available datasets; see Table 3 for frequencies and example rationales from each theme).

Table 3
Frequencies and Examples of Rationales for not Conducting Formal Power Analyses

Rational	%
No basis for estimating effect size and/or conducting exploratory research -new research paradigm where effect size is not known	30.94
Experience and/or can guess appropriate sample size without formal analyses -programmatically research where one knows how big the sample should be	15.11
Uncodeable -unsure or confused about the question	13.67
Limited resources and/or other constraints on increasing sample sizes -sample limited by subject pool size	10.55
Not normative or required -journals do not require power analyses	8.39
Other -intentionally running underpowered studies	8.39
Planned or actual sample sizes are huge -sample sizes are large enough to detect small effects	5.52
Power analyses are ridiculous and/or not valid -researchers can arbitrarily pick effect size for power analysis	5.04
Doing secondary data analyses and/or no control over sample size -sample size is unchangeable	1.20
Snark -when researchers do not care about quality science	0.72
Qualitative research -formal effect sizes are hard to determine	0.48

Note. “Snark” refers to comments that were sarcastic and did not meaningfully fit into any of the other codes.

Reporting Effect Sizes

Three hundred and sixty-two researchers (35.63% of our sample for this item) indicated that not reporting effect sizes was acceptable or reported being uncertain. Of these 362 researchers, 56.08% provided a rationale for their answer.

The most common rationale for not reporting effect sizes was the belief that effect sizes are redundant with other reported statistics and could be inferred by the reader. For example,

many researchers stated that effect sizes can be computed from the test statistics directly (e.g., F or t tests and degrees of freedom) and, therefore, reporting them on their own is unnecessary.

The second most common rationale captured the belief that effect sizes are unimportant, uninformative, or irrelevant for certain types of research questions. Some researchers reported that psychologists are more interested in process, and not intervention, therefore, effect sizes are unimportant and potentially misleading. Others noted that manipulations and measures, particularly in social psychology, are arbitrary and therefore effect sizes are similarly arbitrary and uninformative because they are not measuring the size of a “real life” effect. The third most common rationale was the idea that reporting effect sizes is not standard nor the norm and is not often required by journals. Finally, some in our sample argued that methods for calculating effect sizes vary greatly and that effect sizes themselves do not speak to the validity of the research (see Table 4 for frequencies and example rationales from each theme).

Table 4
Frequencies and Examples of Rationales for not Reporting Effect Sizes

Rationale	%
Effect sizes are redundant with other reported statistics and/or should be calculated by readers -all the information is there for the reader to calculate effect sizes	25.12
Effect sizes are not important, uninformative, and/or irrelevant to the field or specific research questions -psychologists are interested in process, not intervention	21.67
Effect size reporting is not a standard or norm or journal requirement in our field -some journals, reviewers, editors do not ask for effect sizes	15.76
Uncodeable -unsure or misunderstood the question	11.82
Proper methods for calculating effect size are not available or are unclear -it is difficult to calculate effect sizes for multi-level analyses	11.82
Effect sizes do not affect research validity -effect sizes say nothing about the integrity of the research	9.85
Other -not sure how to interpret effect sizes	3.94

Discussion

Our results suggest that the rationales for not adopting PRRPs in psychology are varied. Many researchers identified possible limitations of the current discussion on research practice reform. For example, most of the researchers in our sample provided rationales that revealed concerns about how to implement PRRPs under specific conditions (e.g., preregistering exploratory research, IRB restrictions on data sharing). Concerns about how to implement PRRPs indicate that the discussion needs to more specifically address how to apply these practices with methods outside of traditional laboratory research, such as field designs or research on sensitive populations (see Finkel et al., 2015 for a discussion). Furthermore, guidelines for how to incorporate these practices into IRB protocols and consent documents may help researchers participate in open science while still respecting participants' privacy, particularly as federal regulations change (e.g., Jaschik, 2017). In addition to these practical considerations, some researchers provided rationales that reflect open questions in the research practice reform discussion (e.g. do researchers maintain authorship rights over their publicly available data?) that will require continued discussion to establish field norms.

Some researchers in our sample seem to resist using PRRPs because they think they are not needed or that the transaction costs associated with them create an unnecessary burden on researchers (e.g., effect sizes are redundant, preregistering is another bureaucratic hurdle). These rationales suggest that some researchers are resistant to certain PRRPs. Thus, efforts should focus on combating resistance through clarifying the utility and benefit of using such practices. For example, even if researchers are honest in their endeavors, preregistration (and other PRRPs) may still protect them from unwittingly conducting questionable research practices (e.g., HARKing; Kerr, 1998). In addition to clarifying the utility of preregistration, our results indicate

that increasing the ease of use of preregistration platforms (and therefore reducing transaction costs) may also encourage more wide-spread adoption of this practice.

Finally, some rationales indicated that there is still a great deal of confusion about what PRRPs are and how one goes about using them. Some of the researchers in our sample indicated they were unaware that preregistering studies and making data available online, for example, were considered “best practices” or did not know how or where to preregister or make their data publicly available. Therefore, it would be beneficial to more broadly inform researchers in psychology about open science initiatives that provide platforms for storing preregistrations, study materials, and/or data online (e.g., Open Science Framework, The Dataverse Project, AsPredicted.org). These types of responses suggest that the field still has a way to go to educate all researchers—including those who may not be able to routinely attend conferences, for example—about still quite new platforms for preregistering and data sharing.

In summary, improving psychological science with PRRPs will require not only breaking down perceived barriers for researchers who want to use these reforms but do not think they can or, simply, are not sure how, but also addressing the fact that some researchers do not see a need to reform research practices in the first place.

References

- Boyatzis, R. E. (1998). *Transforming qualitative information*. London: Sage.
- Center for Open Science. (2013). *Badges to acknowledge open practices*. Retrieved from <https://openscienceframework.org/project/TVyXZ/>
- Eich, E. (2014). Business not as usual. *Psychological Science*, 25(1), 3-6.
- Epstein, L., & Martin, A.D. (2005). Coding variables. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 1, pp. 321–327). Oxford: Elsevier Academic Press.
- Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological & Personality Science*, 7(1), 45-52.
- Finkel, E. J., Eastwick, P. W., & Reis, H. T. (2015). Best research practices in psychology: Illustrating epistemological and pragmatic considerations with the case of relationship science. *Journal of Personality and Social Psychology*, 108(2), 275-297.
- Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Sansone, C., Vazire, S., & West, S. G. (2014). Improving the dependability of research in personality and social psychology: Recommendations for research and educational practice. *Personality and Social Psychology Review*, 18(1), 3–12.
- Goldin-Meadow, S. (2016a). Why preregistration makes me nervous. *APS Observer*, 29(7).
- Goldin-Meadow, S. (2016b). Preregistration, replication, and nonexperimental studies. *APS Observer*, 29(8).
- Jaschik, S. (2017, January 19). New ‘common rule’ for research. *Inside Higher Ed*. Retrieved from <https://www.insidehighered.com/news/2017/01/19/us-issues-final-version-common-rule-research-involving-humans>

- Joffe, H., & Yardley, L. (2004). Content and thematic analysis. In D. Marks & L. Yardley (Eds.), *Research methods in clinical and health psychology* (pp. 56–68). London: Sage.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524–532.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196-217.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.
- Motyl, M., Demos, A. P., Carsel, T. S., Hanson, B. E., Melton, Z. J., Mueller, A. B., ... Skitka, L. J. (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*.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... Yarkoni, T. (2015). Promoting an open research culture. *Science, 348*(6242), 1422–1425.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), 253–267.
- Schweinsberg, M., Madan, N., Vianello, M., Sommer, S. A., Jordan, J., Tierney, W., ... Uhlmann, E. L. (2016). The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline. *Journal of Experimental Social Psychology, 66*, 55-67.
- Scott, S. (2013). Pre-registration would put science in chains. *Times Higher Education*. Retrieved from <http://www.timeshighereducation.co.uk/comment/opinion/pre-registration-would-put-science-in-chains/2005954.article>

Simmons, J. P., Nelson, L. N., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant.

Psychological Science, 22(11), 1359-1366.

Society for Personality and Social Psychology. (2016). *Membership diversity statistics, SPSP,*

Inc. (12/31/2016). Retrieved from

<http://spsp.org/sites/default/files/Membership%20Diversity%20December%202016.pdf>