

ORIGINAL RESEARCH REPORT

Quantity Over Quality? Reproducible Psychological Science from a Mixed Methods Perspective

Hollen N. Reischer and Henry R. Cowan

A robust dialogue about the (un)reliability of psychological science findings has emerged in recent years. In response, metascience researchers have developed innovative tools to increase rigor, transparency, and reproducibility, stimulating rapid improvement and adoption of open science practices. However, existing reproducibility guidelines are geared toward purely quantitative study designs. This leaves some ambiguity as to how such guidelines should be implemented in mixed methods (MM) studies, which combine quantitative and qualitative research. Drawing on extant literature, our own experiences, and feedback from 79 self-identified MM researchers, the current paper addresses two main questions: (a) how and to what extent do existing reproducibility guidelines apply to MM study designs; and (b) can existing reproducibility guidelines be improved by incorporating best practices from qualitative research and epistemology? In answer, we offer 10 key recommendations for use within and outside of MM research. Finally, we argue that good science and good ethical practice are mutually reinforcing and lead to meaningful, credible science.

Keywords: reproducibility; replicability; mixed methods; open science; ethics

Psychological science has benefited tremendously from the recent surge in attention to reproducibility.¹ Reproducibility has been defined in many ways; at its core, reproducibility relates to the broad problem of determining “whether scientific claims based on scientific results are true” and the strategies scientists use to estimate this “truth” (Goodman, Fanelli, & Ioannidis, 2016).² A growing metascience literature now offers valuable recommendations on research transparency (Miguel et al., 2014; Morey et al., 2016; Simmons, Nelson, & Simonsohn, 2011; Simons, Shoda, & Lindsay, 2017), replication studies (Asendorpf et al., 2012; Crandall & Sherman, 2016; Simons, 2014), data sharing (Goodman et al., 2014; Lindsay, 2017), and overall research design and execution (Chambers, 2017; Christensen, 2018; Wicherts et al., 2016).

However, this attention has largely focused on a certain type of psychology study: confirmatory, quantitative analyses, often of convenience samples. While this type of research is invaluable, it is not the only way to conduct meaningful, credible psychological science. As recent contributions in clinical science (Tackett et al., 2017) and qualitative psychology (Levitt et al., 2018) point out, methodologically diverse subfields of psychological

science must also find a place at the reproducibility table. In that vein, the current paper focuses on mixed methods (MM) psychological studies. MM studies belong to a large and diffuse category of research which combines quantitative and qualitative analyses. There are many different ways to conduct an MM study. Some MM study designs include, for example, qualitative methods as hypothesis-generation preceding hypothesis-testing; qualitative methods as illustration or follow-up of hypothesis-testing; quantitative analysis of qualitative data; or findings triangulated through simultaneous qualitative and quantitative analyses (see Survey Method below for an elaboration of these designs). It is unclear to what extent current best practices in reproducibility ought to apply to these frequently employed designs. There is a need for clear discussion about which best practices apply, which (if any) do not, and which should be modified in an MM context.

Moreover, qualitative research has a rich tradition of attending to issues such as researcher bias, context effects, generalizability, and ethical implications of research. Qualitative researchers often embrace a view of psychology as a *human* science (Giorgi, 1969). In this view, knowledge of the human experience guides data collection, colors interpretation, and motivates research questions which advance human and social welfare (Josselson, 2007; Mustakova-Possardt, Lyubansky, Basseches, & Oxenberg, 2014; Prilleltensky, 1996). Thus, a human science perspective directs researchers to attend

Northwestern University, Evanston, Illinois, US

Corresponding author: Hollen N. Reischer
(hollen@u.northwestern.edu)

to oft-ignored elements of research that may impede producing more objective, useful research. Quantitative researchers addressing these issues risk reinventing the wheel if they are not exposed to relevant qualitative best practices.

Although qualitative and quantitative psychological research have historically progressed apart, the lack of scholarly exchange between the two may harm reproducibility efforts for both. Rather than two methods at odds with one another, perhaps they are best seen as different cultures that would benefit from cross-cultural dialogue. In this view, mixed methodologists may be the most fluent code-switchers to spark this dialogue and improve methods in both types of research. We suggest there is a place in the “replicability crisis” for broader dialogue about what it means to do “good” science. Psychological science can only be strengthened by an understanding of diverse, rigorous methodologies.

In the current study we draw on extant literature, our own experiences, and feedback from 79 self-identified MM researchers to address two main questions: (a) how and to what extent do existing reproducibility guidelines apply to MM study designs; and (b) can existing reproducibility guidelines be improved by incorporating best practices from qualitative research and epistemology?

Method

Our first step was to increase our own knowledge of the state of affairs in reproducible psychological science, both broadly and in regard to our specific research methodology (i.e., MM). In addition to reviewing the literature, the first author participated in an intensive open science training (Berkeley Institute for Transparent Social Sciences, September, 2017) and other educational opportunities. She also presented early drafts of components of this paper and received feedback from faculty and graduate students. The second author co-organizes an ongoing reading group focused on the reproducibility movement.

Following these experiences, we identified a list of target issues and strategies which seemed most relevant to our questions about the movement’s relationship to MM research. We then conducted an empirical survey of MM researchers to understand how our peers conceived of these same issues and strategies. The method of this survey is detailed below, and the results of our literature review, personal experiences, and the study are presented together.

Survey Method

Data were collected via an online survey distributed to self-identified MM researchers; we provided our definition of “mixed methods” at the outset of the survey, as described below. We surveyed opinions and attitudes toward a variety of topics in reproducibility and open science; the survey is available on this study’s Open Science Framework (OSF) web page (<https://osf.io/fkxm7/>).

Definition of “mixed methods”

Our survey described four instantiations of “mixed methods” (MM), acknowledging there are others. These were presented in the survey as follows:

1. Qualitative methods as hypothesis-generation: Qualitative methods are used as exploratory analyses to develop specific hypotheses. These hypotheses are then tested using confirmatory quantitative methods. *E.g., a researcher organizes focus groups for mothers with post-partum depression to figure out which outcome variables to use in a future study on social support and post-partum depression.*
2. Qualitative methods as illustration or follow-up: Qualitative data are used for examples, illustrations, or supplementary analysis of quantitative findings. *E.g., a study finds a quantitative association between self-esteem and substance use attributions among drug users, and includes quotes to illustrate participants’ motivations for substance use.*
3. Quantitative analysis of qualitative data: Qualitative data are coded to generate quantitative variables. These quantitative variables are then analyzed using quantitative statistical methods. *E.g., trained coders rate an interview transcript for how much autonomy the interviewee describes (0 = no autonomy; 4 = strong autonomy). A researcher calculates the correlation between this autonomy code and a self-esteem questionnaire.*
4. Qualitative and quantitative methods in parallel: A study collects both qualitative and quantitative data, and triangulates findings through parallel quantitative data analyses. *E.g., a grounded theory³ study of adolescent mothers’ experiences in the health care system includes interviews and the Beck Depression Inventory. Researchers describe in detail a theme of “alienation” that appeared in 75% of participants’ interviews. Researchers note that these participants’ BDI scores were significantly higher than those of participants who did not discuss “alienation.”*

Respondents

Respondents were recruited through postings on psychology and related disciplines’ online message boards, mailing lists, and social media. The postings identified the researchers as graduate students working with MM research designs, seeking to collate opinions and best practices related to replicability and reproducibility in order to fill an identified gap in discourse about these issues. A copy of the recruitment language is posted on this study’s OSF page (<https://osf.io/fkxm7/>). Individuals were invited to participate if they were: (a) English speakers, (b) active scholars or scientists, and (c) self-identified as MM researchers. A total of 113 respondents opened the online survey. We excluded any responses that were less than 25% complete (defined a priori). This left a total of 79 responses (66 complete responses and 13 partial responses; partial responses were 50.9% complete on average, $SD = 21.2\%$).

Table 1 shows respondents’ demographic characteristics. A plurality of respondents identified as psychology researchers and a majority (77%) were graduate students, post-doctoral researchers, or professors. Roughly two thirds of respondents reported having completed a doctoral degree. These respondents tended to be early- to mid-career scholars, although 16 respondents (20%)

Table 1: Survey respondents' demographic information.

Variable	N (%) or Mean (SD)	
Field of study		
Psychology	45	(57%)
Other social science	8	(10%)
Other	19	(24%)
Prefer not to answer	7	(9%)
Position		
Professor/research scientist	33	(42%)
Post-doctoral researcher	10	(13%)
Graduate student	18	(23%)
Other	11	(14%)
Prefer not to answer	7	(9%)
Completed doctoral degree	51	(65%)
Years since doctoral degree	14.9	(14.6)
Tenured or tenure-track faculty	31	(39%)
Research-focused university	23	(74%)
Teaching-focused university/liberal arts college	8	(26%)
Methodological orientation (0 = all quant; 100 = all qual)	34.5	(25.7)

earned their doctorate at least 20 years prior to the survey. Most respondents worked in psychology but roughly one third (34%) worked in allied fields such as sociology or anthropology, reflecting the interdisciplinary nature of MM research and concern for reproducible best practices in these fields (see Barker & Pistrang, 2005; Freese & Peterson, 2017; Newman & Hitchcock, 2011).

Respondents rated their own research orientation on a slider that ranged from 0 = *entirely quantitative* to 100 = *entirely qualitative*. The distribution of these ratings was bimodal, with one peak at 50 and a second peak below 10. Only two respondents rated their research orientation above 80 on this scale, while twenty respondents rated their orientation as 10 or below. It is noteworthy that respondents offered wholehearted and nuanced arguments in favor of qualitative research principles (discussed in detail below), given that their own research orientations leaned toward quantitative methods.

Procedure

Respondents answered an online survey which provided working definitions of various reproducibility issues and strategies and asked a) how *relevant* the topic is to MM research; b) the respondent's *personal engagement* with the topic, that is, how much the respondent thinks about the topic in their own research; and c) for any free-response thoughts. Relevance and engagement were presented on a 1 to 5 scale, where 1 = *least agreement* and 5 = *most agreement*. When discussing reproducibility issues, we also asked participants to compare each issue's *prevalence* and *conceptual concern* in quantitative versus MM research. When discussing reproducibility strategies, we also asked

participants to compare each strategy's *practicality* and *desirability* in quantitative versus MM research. The order of various issues and strategies was randomized. The full survey can be found online at osf.io/g86qz/.

Quantitative analysis was conducted using R version 3.2.4 (R Core Team, 2017) with packages broom (Robinson, 2015), lm.beta (Behrendt, 2014), psych (Revelle, 2018), and tidyverse (Wickham, 2016b). Figures were created using the package ggplot2 (Wickham, 2016a). For the qualitative analysis, both authors read all free responses, identified common themes for each topic, and agreed on the most representative themes and quotes through consensus. Formally, this thematic analysis was theoretical, semantic, and realistic (Braun & Clarke, 2006). We considered performing a quantitative analysis of the free response data (e.g., quantitatively coding relevant themes for statistical analysis), however, we chose not to pursue this analysis to ensure a more "purely" qualitative component to the study.

Results and Discussion

Results focus first on issues related to reproducibility and then on strategies for addressing these issues. Before each specific topic and in **Table 2**, we have highlighted ten specific recommendations to increase reproducibility both within and outside of MM research. These recommendations are a synthesis of results, relevant literature, and our own perspectives, as described below.

Issues Related to Reproducibility

MM research designs, by definition, include quantitative components. Given that many reproducibility issues have

Table 2: Summary of recommendations for psychological researchers across the methodological spectrum.

Concern or strategy	Recommendations	Study design	Data collection and analysis	Interpretation and dissemination
Questionable research practices (QRPs)	1. Be aware of qualitative QRPs in mixed methods research. At the same time, elements of a “qualitative mindset” (e.g., sensitivity to context and how data were generated) may protect against QRPs.		✓	✓
Power and precision	2. Plan sample size according to intended theoretical inferences: population-level (power analysis) or localized (purposive sampling).	✓		✓
Perverse incentives	3. Editors and reviewers should advocate for slow science, ethical discussion, and inclusion of reviewers with diverse methodological expertise.			✓
Replication studies	4. Consider which validity indices are most relevant to a given study; replication studies may not be appropriate.	✓		✓
Open science: Data sharing	5. Create a data-sharing plan before collecting data. Feel empowered to say “no” to a data-sharing request if the risk to confidentiality cannot be effectively managed.	✓		✓
Open science: Preregistration	6. Researchers should routinely consider preregistration; theoretical and iterative preregistration may be particularly relevant in mixed methods research.	✓	✓	✓
Open science: Reducing barriers to entry	7. Strive to include community members in the research process when possible.	✓	✓	✓
Open science: Reducing barriers to entry	8. Include the research of minorities and women when teaching psychology.			✓
Reporting practices	9. Report details of methodology and decisions made throughout the research process; do not assume shared understanding with readers.			✓
Reporting practices	10. Reflect on, discuss, and report personal and contextual biases that could affect research design or interpretation.	✓	✓	✓

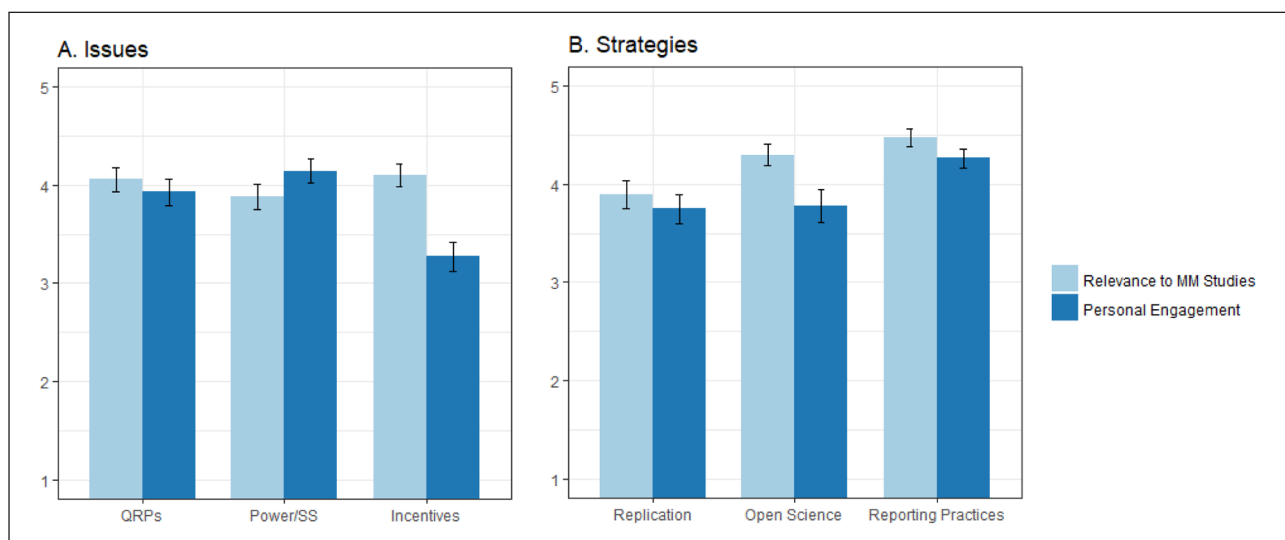
Note: Checkmarks indicate the stage(s) of the research process most relevant to each recommendation.

been identified in the quantitative literature, we wondered how and to what extent these issues apply to MM designs. We focused on three broad issues: questionable research practices (QRPs), power and sample size, and perverse incentives. QRPs refer to researchers' conscious or unconscious decisions, made at any point in the research process, that increase the likelihood of ending up with a false positive result or inflated effect size. Statistical power is the ability to correctly reject the null hypothesis; sample size is an important component in determining power. Finally, perverse incentives refer to academic incentive structures which reward individual research output (e.g., fast and frequent publishing of novel, positive results) over the evidentiary value of the field as a whole.

Table 3 shows participants' assessments of each issue's relevance in MM designs, and their own personal engagement with each issue. **Figure 1** displays these same data graphically. Survey respondents felt that all three issues were relevant to MM designs, with the median response on each issue being 4, or “very relevant.” Tukey's HSD test at $\alpha = .05$ indicated no significant differences between respondents' perceptions of each issue's relevance. However, Tukey's HSD test showed that respondents reported more personal engagement with QRPs and power/sample size than with perverse incentives. Differences between personal engagement with QRPs and power/sample size were not significant. Respondents also tended to report that perverse incentives

Table 3: Mixed methods researchers' ratings of relevance and personal engagement for various reproducibility issues and strategies.

	Relevance		Personal engagement	
	Mean	95% CI	Mean	95% CI
Issues				
Questionable research practices	4.06	[3.81, 4.31]	3.92	[3.65, 4.20]
Perverse incentives	4.10	[3.88, 4.33]	3.27	[2.97, 3.57]
Power and sample size	3.88	[3.62, 4.14]	4.14	[3.90, 4.38]
Strategies				
Replication studies	3.89	[3.61, 4.17]	3.75	[3.45, 4.05]
Open science practices	4.30	[4.08, 4.51]	3.78	[3.45, 4.51]
Improved reporting practices	4.47	[4.29, 4.65]	4.26	[4.06, 4.46]

**Figure 1:** Mixed methods researchers' ratings of relevance and personal engagement for various reproducibility issues (A) and strategies (B).

were more prevalent in quantitative than in MM designs (see **Table 4** and **Figure 2**). Interestingly, respondents believed that small sample size studies were both more prevalent and less of a conceptual concern in MM designs.

Questionable research practices (QRPs)

Recommendation 1: Be aware of qualitative QRPs in mixed methods research. At the same time, elements of a "qualitative mindset" (e.g., sensitivity to context and how data were generated) may protect against QRPs.

One of the evidentiary value movement's piercing insights has been to define a class of research practices which are not fraudulent and may not obviously be misleading, but which inflate the probability of finding a positive result (John, Loewenstein, & Prelec, 2012; Simmons, Nelson, & Simonsohn, 2011). For example, a researcher might report an unexpected finding as if it were predicted a priori, not report all dependent variables in a study, or test hypotheses before deciding whether to collect more data. QRPs are now widely seen as relevant

and concerning in quantitative studies. However, we were unable to find corresponding literature on QRPs in qualitative research, and we were unsure of how QRPs were being understood among MM researchers.

The survey respondents were quite aware of QRPs, indicating that QRPs were "very" relevant in MM studies and reporting high personal engagement in attending to QRPs (see **Table 3** and **Figure 1**). They also tended to think that QRPs were equally prevalent and equally concerning in both quantitative and MM designs (see **Table 4** and **Figure 2**). Although QRPs are typically discussed in a quantitative context, respondents identified specific QRPs that relate to an MM study's qualitative components. To the best of our knowledge, these qualitative QRPs have not been previously reported in the psychological metascience literature. They included: (a) cherry-picking examples to suit a hypothesis, over-interpreting unusual examples, or presenting unusual examples as representative; (b) modifying a coding scheme without reporting and justifying the modifications; and (c) failing to report

Table 4: Comparing the prevalence and conceptual concern of reproducibility issues in quantitative vs. mixed methods (MM) study designs.

	Prevalence				Conceptual concern			
	More prevalent in quantitative	About the same	More prevalent in MM	No answer	More concerning in quantitative	About the same	More concerning in MM	No answer
Questionable research practices	20 (30%)	29 (43%)	7 (10%)	11 (16%)	13 (20%)	38 (58%)	4 (6%)	10 (15%)
Perverse incentives	34 (50%)	23 (34%)	3 (4%)	8 (12%)	12 (31%)	30 (44%)	6 (9%)	11 (16%)
Small sample sizes	10 (14%)	9 (13%)	38 (55%)	12 (17%)	30 (45%)	22 (33%)	7 (11%)	7 (11%)

Note: Cells contains response counts and percentages of survey respondents who endorsed each option.

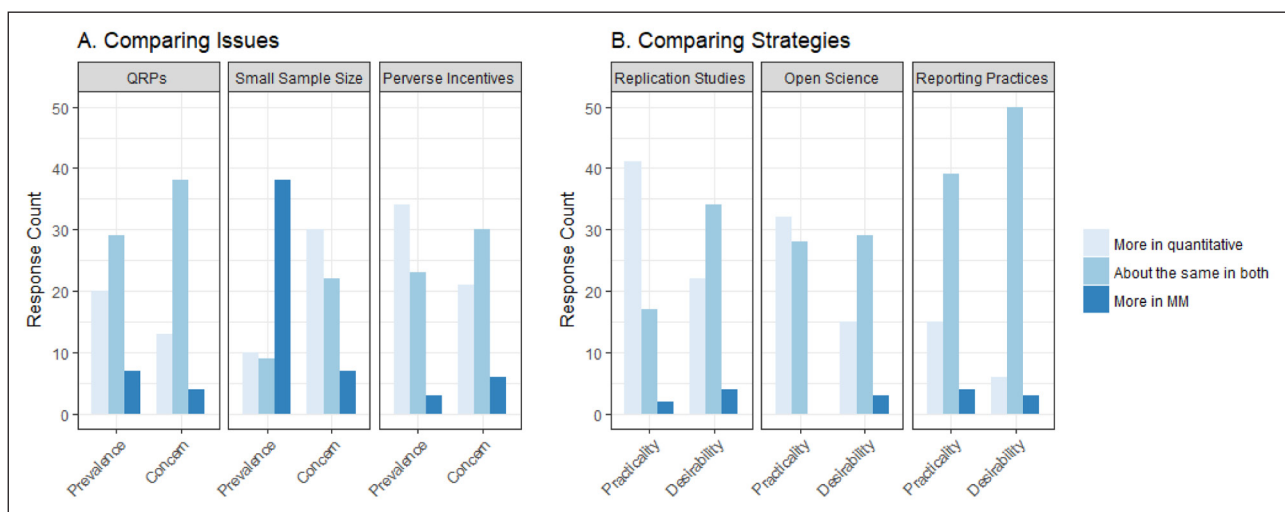


Figure 2: Mixed methods researchers' comparisons of (A) the prevalence and conceptual concern of various replicability issues, and (B) the practicality and desirability of various replicability strategies, in quantitative vs. mixed methods research.

the phases of an MM project, or how quantitative and qualitative data have been integrated. **Table 5** lists these and other examples of quantitative and qualitative QRPs.

Respondents also pointed out that at least some QRPs (e.g., *p*-hacking, falsifying or manipulating data) are difficult to envision in a qualitative context. Similarly, several respondents felt that the mindset required for qualitative research protects against QRPs because it is more descriptive, more sensitive to context and how the data were generated, less hypothesis-driven, less apt to assume that the researcher already understands the phenomena of interest, less wedded to statistical significance and “positive” results, and more ethically concerned with the “truthfulness” of the product.

Finally, respondents highlighted some tension between qualitative and quantitative epistemologies in their discussion of QRPs. Quantitative logic has tended not to question certain commonly-held assumptions. For example, Likert scale measures are often treated as if they represent objective truth about a theoretical construct, rather than the outcome of a chain of inference. If we

adopt a qualitative mindset, we might consider the many links in an inferential chain that starts with latent truth and ends with scale item responses. These might include: (a) the scale developer’s training, life experience, biases, and motivation; (b) specific item wordings and instructions; (c) the study setting; (d) participants’ comprehension, assumptions, and motivations; (e) the need for researcher and participant to understand items in the same way; and so on. Although these issues are sometimes considered in quantitative studies, qualitative logic encourages researchers to habitually foreground and discuss the data generation process and the assumptions of any given project.

However, we note this transparency does not necessarily protect against QRPs, as some participants seemed to imply. While it makes quantitative QRPs less likely, it also opens the door to qualitative QRPs. Researchers using quantitative, qualitative, and mixed methods are all prone to human biases and motivated reasoning. There seem to be equivalent numbers of decision points and “researcher degrees of freedom” in any study design. Thus,

Table 5: Examples of Questionable Research Practices (QRPs) in Quantitative and Qualitative Research.

Research phase	Quantitative	Qualitative
Data analysis	Testing hypotheses before deciding to collect more data Flexible stopping criteria for data collection Flexibly excluding outliers	Cherry-picking examples to suit a hypothesis Presenting unusual examples as representative Over-interpreting unusual examples
Study reporting	Failing to report all of a study's dependent measures Reporting an unexpected finding as if it had been predicted ahead of time Selectively reporting analyses or procedures that "worked"	Failing to report and justify modifications to a coding scheme Failing to report when concepts or codes have been dropped from a qualitative analysis Failing to report the phases of an MM project, or how quantitative and qualitative data have been integrated

a qualitative mindset may provide a false sense of security if a researcher is not alert to the possibility of qualitative QRPs such as those described above.

Power and precision

Recommendation 2: Plan sample sizes according to intended theoretical inferences: population-level (e.g., power analysis) or localized (e.g., purposive sampling).

Recommendations for quantitative reproducibility commonly include larger sample sizes and increased statistical power and precision (Bertamini & Munafò, 2012; Ioannidis, 2005; Simmons et al., 2011). There is an obvious conceptual tension in recommending larger sample sizes for MM studies because qualitative methods demand intense scrutiny of each participant. How do MM researchers make sense of this tension?

Interestingly, survey respondents felt that small-sample-size studies were both more prevalent and less of a concern in MM research compared to purely quantitative research (see **Table 4** and **Figure 2**). Respondents laid out several reasons that qualitative samples tend to be smaller than quantitative samples or even, in the words of one respondent, why "qualitative research is generally meant for [emphasis added] small samples." Respondents pointed to: (a) the amount of time and resources required to analyze each participant's data; (b) the practice of purposive rather than random or convenience sampling; (c) the focus on person-centered, complex, contextual analyses; and (d) the accompanying participant burden. These concerns are somewhat analogous to issues of sample size in longitudinal research or research with hard-to-recruit samples (see Tackett et al., 2017). A minority of respondents argued that the added depth of qualitative data makes large sample sizes in MM studies all the more valuable, despite the outsized investment required. For example, one respondent wrote:

Having such an in-depth approach to research powered at a high level, particularly incorporating mixed methods, is really valuable, and allows participant experiences of a concept to really come to the fore, without sacrificing the [rigor] of the work. Large sample qualitative studies really excite me in that sense, and I think we're seeing more and more of them since the "replication crisis."

In addressing these considerations, it is important to make a distinction between population-level inference and localized inference. Quantitative research typically (though not always) aims to maximize breadth and generalizability of inferences to an entire population. In contrast, qualitative research often aims "not to generalize to a population but to obtain insights into a phenomenon, individuals, or events" (Onwuegbuzie & Leech, 2007). In this type of design, the participants are treated as a fixed rather than a random effect (Judd, Westfall, & Kenny, 2012) and inferences are *localized*, that is, limited to the specific study participants.

Population-level inference follows the logic of statistical power and precision. Probability sampling and large sample sizes are beneficial, as they increase power to detect population-level effects and precision to estimate the sizes of these effects. A researcher can estimate a predicted or meaningful effect size based on previous literature, then conduct a power analysis to determine an adequate sample size. If resource investment is a concern, MM researchers have several options to reduce the analytic burden per participant. These include: (a) collecting smaller quantities of qualitative data from each participant; (b) analyzing a subset of the qualitative data (e.g., choosing a single pertinent interview question as the unit of analysis, rather than the entire interview); and (c) reducing the complexity of qualitative analyses. Another solution is to work with large collaborative teams to pool data collection and analysis resources around shared questions of interest (for a recent example of a large MM study using this approach, see McLean et al., 2019). Qualitative analyses may be simplified by using pre-existing coding schemes, training large teams of coders, or supplementing traditional coding with computerized coding (e.g., LIWC; Pennebaker, Boyd, Jordan, & Blackburn, 2015).

Localized inference does not necessarily follow the same logic, as it is a detailed understanding of phenomena within the context of the particular participants under study (Onwuegbuzie & Leech, 2007). For example, a researcher might administer a quantitative measure of depression to 500 participants, then conduct qualitative interviews with the 20 highest-scorers to investigate the life experiences that led these specific participants to report high depression scores. Localized qualitative

inferences would shed light on the phenomenology of depression for these particular participants. Some inferences about severely depressed people in this specific cohort or community might be warranted. However, further research would be required to determine the extent to which the qualitative findings generalized to other people who experience depression. Population-level inferences require an adequate sample size, defined by adequate statistical power and precision. However, with purposive sampling and localized inferences, a researcher can obtain important insights about a phenomenon, individuals, or events, even if a sample size is too small to permit population-level inferences. Researchers should therefore consider their intended theoretical inferences ahead of time, plan a sampling strategy accordingly, and avoid population-level inferences if a sample size cannot support such claims.

Perverse incentives

Recommendation 3: Editors and reviewers should advocate for slow science, ethical discussion, and inclusion of reviewers with diverse methodological expertise.

Perverse incentives refer to institutional “carrots and sticks” that systematically favor individuals’ research output (e.g., number of studies, citations, grants, and novel results) over the field’s scientific progress. Anecdotally, MM research is often assumed to be slower and more labor-intensive, leading to a slower publication rate. As Smaldino and McElreath (2016) and others have pointed out, academic incentives reward publication rate above all else, which motivates researchers to value expediency over rigor. We were curious, therefore, how MM researchers made sense of perverse incentives if their work tends to be slower and more laborious.

While survey respondents felt that perverse incentives were relevant in MM designs, they also felt that these incentives were less prevalent and less of a conceptual concern than in quantitative research (see **Table 4** and **Figure 2**). Respondents also reported thinking about incentives less frequently than other issues in their own research decision-making (see **Table 3** and **Figure 1**). Even so, survey respondents were consistent with the existing literature on this subject; although the survey defined perverse incentives broadly, respondents focused almost exclusively on publication rate. “Publication is the currency of higher ed,” one respondent wrote, “so one would be foolish not to think strategically” about publishing. Another respondent cited expectations that junior authors publish a certain number of articles each year, and concluded, “I can’t imagine a better way to fill the airwaves with mindless junk.”

Many respondents argued that this publication rate is harder to meet when working in MM. Practically, MM designs often take longer to complete, have difficulty fitting within journal word limits, and are not suitable for piecemeal publication. Conceptually, valuable components of MM designs such as ecological validity, reflexivity, philosophical and ethical discussion, qualitative rigor, and tolerance of ambiguity seem not to be valued

by many journal editors and reviewers. Journals’ lack of interest in MM research then becomes a self-fulfilling prophecy, as the incentive to publish pushes early-career researchers away from MM studies. One respondent described this cycle:

I suspect the majority of research that’s published in psychology is quantitative-only...so fewer people understand and recognize the value of qualitative data. I have observed bias within my field, and my own department, where hard-core quantitative is viewed as better (more intellectual, rigorous, valuable) than qualitative (seen as weak, easy, and less important).

In sum, respondents felt that publication rate was the key incentivized metric and felt that current incentives favor fast and frequent publication of purely quantitative work. What does this mean for perverse incentives in MM research? The time and effort required to collect and process qualitative data suggests that MM researchers may be especially motivated to extract publishable findings—as one respondent observed, “I also have a general heuristic that says that the harder data are to collect and process/code, the stronger these perverse incentives become.” On the other hand, several respondents noted that “publishable” in an MM context does not require positive, novel, or surprising findings. The status of a null finding is often conceptually different in a qualitative context because “mixed models use qualitative data and exploratory analyses which are often considered interesting either way” (survey respondent). While a quantitative null is absence of evidence, in a well-designed study a qualitative null can be evidence of absence. Moreover, rich qualitative data allows researchers to “explore whether (and why) the quantitative outcomes match (or don’t) with perceived experiences. That is, a strong qualitative component in mixed-methods studies can provide researchers with more to go on than statistical significance” (survey respondent).

Given the potential disadvantages of time-intensive qualitative analyses for publication rates (and therefore researchers’ livelihoods), we suggest registered reports may be particularly useful to MM researchers. In this format, researchers submit a study proposal prior to data collection, the target journal reviews the proposal, and if the editorial board accepts the study in principle and the researchers adhere to their proposal, the journal is committed to publishing the completed study regardless of the findings. Similarly, we suggest journal editors may deliberately be more open-minded to MM studies and ensure they have relationships with qualitative and MM experts who are able to evaluate submissions’ qualitative rigor.

Finally, survey respondents touched on deeper issues in the publication process. As discussed above, respondents discussed the characteristics of MM studies that make them less easily publishable in top journals. Consider a partial list of these characteristics: (a) a deliberate, measured

timeline; (b) complete rather than piecemeal publication; (c) discussions of ethics; (d) ecological validity; and (e) designs in which null results are meaningful. These are not unique to MM designs. Rather, they are characteristics of good science. It is therefore all the more troubling that many researchers feel they must submit these studies to lower-prestige journals, accept a lower rate of publication, and expect damage to their career prospects as a result—not because this is harmful to the researchers themselves, but because it speaks to pervasive issues in a publishing system that favors a rapid pace of positive, novel, and surprising findings. The publishing industry’s “natural selection of bad science” (Smaldino & McElreath, 2016) hinders many rigorous study designs, not only those that employ MM.

Strategies for Improving Psychological Science

We considered three sets of strategies to address reproducibility issues: replication studies, open science, and reporting practices. Replication studies test whether a different research team and/or a different sample will produce similar results to an original study. Open science encompasses a suite of ideas and practices that promote increased transparency and efficiency in all areas of scientific research. Finally, reporting practices refer to norms and guidelines for the reporting of research, typically in peer-reviewed publications.

Respondents’ ratings of relevance and personal engagement are shown in **Table 3**, and ratings of each strategy’s desirability and practicality are shown in **Table 6**. **Figures 1** and **2** display these same data graphically. Survey respondents felt that all three strategies were “very” to “extremely” relevant to MM designs. Tukey’s HSD tests at $\alpha = .05$ showed that replication studies were seen as less relevant to MM than open science practices and reporting practices. There was no difference in relevance between the latter two. Tukey’s HSD also showed that respondents reported more personal engagement with reporting practices than either open science practices or replication studies.

Respondents tended to report that all three strategies were equally desirable in MM and quantitative research. Respondents indicated that comprehensive reporting practices were equally practical in MM and quantitative research, while both replication studies and open science practices were more practical in quantitative research.

Replication studies

Recommendation 4: Consider which validity indices are most relevant to a given study; replication studies may not be appropriate.

The reproducibility movement has extensively promoted replication studies. Generally, these studies recruit a new participant sample, re-run an original study’s procedures, and statistically test whether the new study’s results “replicate” the original study’s results by reaching statistical significance or showing a similar effect size.⁴ Yet purely qualitative researchers often argue that their research is not meant to be replicable according to the rules and conventions of inferential statistics (e.g., Freeman, 2011; Rappaport, 2005). From that perspective, qualitative or MM researchers may justifiably wonder whether direct replication is relevant to their work. Although quantitative, qualitative, and MM researchers agree on the need for scientific validity, replication studies do not apply equally well in these research domains, and more diverse strategies are needed (Newman & Hitchcock, 2011). As a first step in defining the role of replication in MM research, we wondered how MM researchers currently interpret and apply replication in their own scientific practices.

As described above, survey respondents found that replication studies were relevant to MM, but less relevant and less practical in MM than either open science or reporting practices. This stance reflects the epistemological implications of each research method. Many respondents articulated that while replicability is relevant to the quantitative components of MM research, it is generally an inappropriate benchmark for qualitative research. As one respondent stated:

Table 6: Comparing the practicality and desirability of reproducibility strategies in quantitative vs. mixed methods (MM) study designs.

	Practicality				Desirability			
	More practical in quantitative	About the same	More practical in MM	No answer	More desirable in quantitative	About the same	More desirable in MM	No answer
Replication studies	41 (65%)	17 (27%)	2 (3%)	3 (5%)	22 (35%)	34 (54%)	4 (6%)	3 (5%)
Open science practices	32 (50%)	28 (44%)	4 (6%)	0 (0%)	15 (25%)	39 (64%)	3 (5%)	4 (7%)
Improved reporting practices	15 (23%)	39 (61%)	4 (6%)	6 (9%)	6 (9%)	50 (78%)	3 (5%)	5 (8%)

Note: Cells contains response counts and percentages of survey respondents who endorsed each option.

If we believe...the world is neat, and that science is a matter of uncovering consistencies that exist in a fixed world, well, replicability will be the hallmark of science...However if we think of the world as not so neat...we may not expect to step in the same river twice.

Indeed, as another respondent wrote, "Qualitative approaches and critical epistemologies would question even the possibility that one would find the same results when it is tested in two different contexts, with two different samples, at two different times." This respondent continued, "considering different ways of understanding validity beyond the ability to reproduce the same results twice should be informing this conversation." Similarly, Finkel, Eastwick, and Reis (2015) have highlighted features of quality science (i.e., discovery, internal validity, external validity, construct validity, consequentiality, and cumulativeness) that ought not be lost at the expense of a narrow focus on direct replication. Although direct replication is one standard of validity, quality science can—and should—consider other standards as well.

Beyond direct replication, how can MM researchers enhance their studies' validity? Perhaps the most important method is thorough and specific reporting practices (see Reporting Practices below). In addition, other methods include:

1. Ensuring high interrater reliability among multiple coders through appropriate training and regular check-ins and reliability testing. See Hallgren (2012) for a starting point.
2. Tracking each individual coder's intra-rater reliability across time to identify any conceptual drift or laxity, which may call for additional training.
3. Triangulation with other analysts, data sources, or methods. Triangulation is generally agreed to provide a more robust understanding of the phenomenon of interest, and it may contribute to increased validity of findings if results converge. MM approaches are particularly useful in methodological triangulation (Johnson & Onwuegbuzie, 2004; Newman & Hitchcock, 2011).

Open science

We understand open science as including many concurrent efforts to reform the production and dissemination of scientific knowledge by increasing collaboration, efficiency, transparency, accountability, and accessibility to the public. On the survey, we provided several examples of open science, including data sharing, preregistration, and identifying and reducing barriers to entry to the research process. Researchers who work with particularly sensitive, hard-to-collect, complex, or longitudinal data have good reason to be skeptical of data sharing and preregistration. We sought to tease out and reconcile the benefits of these practices with the real risks they raise.

Furthermore, the qualitative elements of MM emphasize lived experience and suggest that we should examine not only open science's formal, dispassionate best practices, but also its interpersonal and social implications. Open

science is an ethical as well as a practical goal. Merton (1973) named *communalism* as one of the four structural norms of science: the fruits of scientific labor belong not only to the scientist, but to the public domain. In fact, the right to "enjoy the benefits of scientific progress and its applications" has been recognized by most nations as a human right for over half a century ("International Covenant on Economic, Social and Cultural Rights," 1966). As one respondent stated simply, "It is *unjust* [emphasis added] to hide scientific progress behind paywalls and encrypted hard drives." Open science practices bring us closer to fulfilling the value of communalism. Therefore, not only does open science facilitate replication, but it also fulfills science's ethical obligations. Good science and good ethics go hand-in-hand.

Data sharing

Recommendation 5: Create a data-sharing plan before collecting data. Feel empowered to say "no" to a data-sharing request if the risk to confidentiality cannot be effectively managed.

Researchers are often reluctant to share data for a variety of reasons, including confidential and/or disorganized data, mistrust or uncertainty around requesters' intentions, the substantial resource costs of preparing and sharing data, and skepticism of secondary data analysis (Broom, Cheshire, & Emmison, 2009; Martone, Garcia-Castro, & VandenBos, 2018; Polanin & Terzian, 2019; Wicherts, Borsboom, Kats, & Molenaar, 2006). Privacy concerns are particularly important in MM research, with qualitative research having "vastly more challenging issues around participant confidentiality" (study respondent). Many MM and qualitative studies work with small or unusual samples that may be easily reidentified (e.g., Saunders, Kitzinger, & Kitzinger, 2015). One respondent mentioned that "there are also ethical concerns about publishing personal experiences, even if de-identified, for all to see." A robust debate continues around the ethics and feasibility of qualitative data sharing (for a particularly useful exchange, see DuBois, Walsh, & Strait, 2018; Guishard, 2018; McCurdy & Ross, 2018; Roller & Lavrakas, 2018).

We believe that balancing these concerns are substantial benefits to data sharing. For example, it could be frustrating for a researcher that has dedicated years and many thousands of dollars to a research project, only for another investigator to turn out a quick publication using the data. However, that quick publication may be beneficial to the primary investigator. It might move a primary research question forward or open a new and unexpected line of research. MM studies generally produce a wealth of data that can support multiple studies and re-analyses. The original researcher may get more "bang" for their "research buck" by allowing others to delve into their hard-won data.

One respondent pointed out data sharing ought to be "considered at the design/pre-ethical approval stage... [it's] not something that can be left to consider later on." It is preferable to assess risks to confidentiality and plan for deidentified data collection ahead of time, rather than attempt to redact confidential details later (Joel, Eastwick, & Finkel, 2018; Kaiser, 2009; Sturdy, Burch, Hanson, & Molyneaux, 2017; Zandbergen, 2014). Furthermore,

any intentions to share data and for what purpose (e.g., education or research) should be made explicit as part of participants' informed consent.

One tricky issue is how to respond to data-sharing requests from other researchers. Even if data sharing was explicitly addressed during the consent process, it can be unclear whether a particular data-sharing request is reasonable. It is appropriate for qualitative data sharing to be negotiated between the original researcher(s) and requesting researcher(s). A data request form can be a useful starting point for conversations about sharing sensitive data (see Appendix).

It is important to note that sharing even de-identified or aggregated data will sometimes be too high-risk for participant safety or confidentiality. Organizations such as Responsible Data (<https://responsibledata.io>) and the UK Data Service (<https://www.ukdataservice.ac.uk/>) provide detailed, concrete guidelines for risk assessment. If after a thorough risk assessment the original researcher feels that it is not appropriate to share raw data due to confidentiality concerns, they should feel comfortable saying no to a data-sharing request, with reasons for doing so. We also encourage creative solutions to data-sharing. For example, original researchers may invite requesting researchers to visit their institution and work with the data directly, analogous to a historian visiting archives.

Regardless of whether raw data are deemed too sensitive to share, there are many other opportunities for MM researchers to share useful information. It is now de rigueur to share computer code used for quantitative data analysis (Nuijten et al., 2017). In addition to this practice, researchers might also share "lab notebooks," coding manuals, and vivid contextual information to assist other researchers in assessing a study's credibility and planning possible replications.

Lab notebooks would include, for example, field notes and coding memos. Coding memos are free-flowing reflections written primarily for the researcher's analytic process, but also intended for sharing. Memos keep close track of researcher's observations, preserving important analytical evolution and nuance that might otherwise be lost as the study progresses. Qualitative coding manuals often include not only the concrete rating system used to thematically group or score text, but also its driving theory, logic, and how it has been modified over time. Such manuals operationalize otherwise ambiguous thematic labels and coding processes. For example, one respondent described their frustration when they "contacted an author...to ask how they transformed their qualitative data into quantitative [survey] items" and the author responded, "Books on scale construction should help you figure this out." As the respondent reflected, "Without knowing their specific process, nobody knows how to replicate it. [Without transparency], everyone will try to re-invent the wheel." Useful guidelines for making qualitative research transparent are provided by Roberts, Dowell, and Nie (2019) and Aguinis and Solarino (2019); Zörgő and Peters (2019) provide a tool in their Reproducible Open Coding Kit (ROCK) book.

Finally, sharing "a rich account of the field site or population of interest," as one respondent put it,

highlights important factors such as interviewers' demographics and local cultural context, data that are sometimes lost in replication studies (e.g., Anderson et al., 2016; Gilbert, King, Pettigrew, & Wilson, 2016). When the original study provides rich contextual details, a comparison of those details with the present research context may reveal behavioral effects of sociocultural and demographic differences between the studies (Greenfield, 2017). When studies' theoretical underpinnings, methods, and context are transparent, future researchers can more easily conduct close replications, meta-analyses, and evaluate findings in the context of the bigger picture, even when it not possible to access original data.

Preregistration

Recommendation 6: Researchers should routinely consider preregistration; theoretical and iterative preregistration may be particularly relevant in MM research.

Preregistration and preanalysis plans have been widely touted as a strong, if imperfect, guard against QRPs (e.g., Lindsay, Simons, & Lilienfeld, 2016; van 't Veer & Giner-Sorolla, 2016). Several survey respondents reported skepticism or uncertainty about pre-registration of MM research that is exploratory in nature. However, there are substantial benefits to pre-registering such studies, particularly using more flexible forms of pre-registration. Pre-registration allows for clear and deliberate definitions of an exploratory project, freeing researchers from expectations or assumptions related to confirmatory testing (Miguel et al., 2014). And, while the static nature of a pre-registration appears on its face to oppose the dynamic nature of many MM study designs, pre-registration can in fact be a transparent way to document MM or qualitative research processes while enhancing the credibility and even dissemination of the research (Haven & Van Grootel, 2019). As summarized by Kern and Gleditsch (2017, p. 6):

A transparent archiving of research questions, expected data generating processes, and intended designs could help in providing insights into the procedure of qualitative inference...Scholars could illustrate their inferential process, and credibly point to surprising findings and hypotheses generated through the analysis of data or experiences in the field.

Note that this is not a classic "one-and-done" preregistration. MM preregistrations need not be yoked to a single time-stamped data collection and analysis plan. For example, relationship scientists Finkel et al. (2015) promote *theoretical preregistration*, which is the preregistration of theoretical propositions to be investigated, for research that tends to be especially intensive, longitudinal, and/or non-laboratory in nature. In clinical science, Tackett et al. (2017) call for *iterative pre-registration*, which records and time-stamps researchers' original intentions as well as ongoing project modifications and amendments. Political scientists Piñeiro and Rosenblatt (2016) similarly describe a style of pre-registration that emphasizes the iterative interplay between theory, empirical research, and interpretation. These pre-registration styles are

well-suited to MM projects, which are often designed to accommodate new questions and hypotheses as a study unfolds.

Although every MM project is different, researchers should consider some combination of theoretical and iterative pre-registration. For example, let's consider a hypothetical researcher, Dr. S., interested in the relationship between maternal social support and postpartum depression (PPD). Dr. S. pre-registers theoretical assumptions and initial questions about this relationship. Following outreach to potential research partners, Dr. S. decides to run focus groups through a local social service agency which serves women with PPD. Before data collection, Dr. S. registers a theoretical justification for choosing this method, previous knowledge of the sample, hypotheses, and a draft of focus group questions. Following data collection, Dr. S. records personal reflections and a preliminary theory of maternal social support and PPD. Dr. S. now determines which outcome variables are most relevant to examining this relationship, and pre-registers measures and hypotheses for quantitative data collection. This process of iterative pre-registration can continue throughout a study's life cycle.

Various tools are available to assist in this process. For theoretical and qualitative preregistration, templates and guidelines are available (Kern & Gleditsch, 2017; Piñeiro & Rosenblatt, 2016). For iterative pre-registration, version control management systems (e.g., GitHub), project workflow tools (e.g., Open Science Framework), and interactive online notebooks (e.g., Jupyter) are extremely helpful. Finally, pre-registration techniques for in-depth and longitudinal research are rapidly evolving, and it is likely that future insights from the subfields cited here will be helpful for MM researchers. To preregister the current study (<https://osf.io/fkxm7>), we followed a combination of traditional quantitative and proposed qualitative preregistration guidelines (Kern & Gleditsch, 2017; Piñeiro & Rosenblatt, 2016; van't Veer & Giner-Sorolla, 2016).

Reducing barriers to entry

Recommendation 7: Strive to include community members in the research process when possible.

Recommendation 8: Include the research of minorities and women when teaching psychology.

In considering the ways that individuals from underrepresented backgrounds encounter barriers to participating in and conducting research, a few respondents commented on the ways that power dynamics in research can create "a strong distinction between researcher and researched, between 'knower' and 'known.' This [strong distinction] is incredibly dangerous in quantitative research, but is also in some way anathema to mixed-methods research" (survey respondent). Another respondent agreed that:

Quantitative research can be guilty of colonizing [participants] by going in and forcing them to reflect their experiences in the language/understanding/epistemology/ontology

of the researcher, which may or may not reflect those of the participant, and which may or may not produce data that actually maps on to the lived reality of the participant.

As a caveat, the respondent continued, "Many quantitative researchers take it upon themselves to critically examine their role as researcher and colonizer in this process, and they take steps to address these issues. But quantitative research does not *require* [emphasis added] this of its practitioners." It is certainly possible to fall short of appropriately examining these issues in qualitative research, but in quantitative research it is not expected that they will be addressed at all.

The epistemological framework of MM designs tends to assume a co-creation of knowledge between researcher and researched, lessening power imbalances between the two. This benefits both individual studies and the field as a whole (Freire, 1970; Medin, 2017; Nzinga et al., 2018). One research design which seeks to co-create knowledge is community-based participatory research, which reduces barriers to entry by having participants participate in the research beyond their role as "subjects." Community members can be engaged at any or every stage of the research process, from hypothesis generation to data collection to data analysis to writing. This practice allows scientists to ask questions that are more relevant and useful to participants, ask such questions more appropriately and effectively, and increase trust among participants. There are also benefits for community members, who may gain scientific savvy and skills, develop interest in and preparation for pursuing research careers, and experience a mitigated power differential between researcher and participant.

Flowing from this, we suggest sustained attention to increasing diversity at every level of research, from participant to principal investigator. For example, instructors may wish to consider content produced by diverse scientists when constructing psychology courses. Because psychologists who have been historically excluded from mainstream scholarly communities are more likely to have published work in lower-impact factor journals, this research may need to be actively sought (Nzinga et al., 2018). When such research is included in curricula and literature reviews, researchers can draw from a more complete scientific literature, enhancing their ability to conduct thorough, credible, and reproducible research.

Conducting community-based participatory research and developing diverse curricula are but two examples of ways psychological researchers can leverage institutional power to broaden scientific perspectives, increase scientific literacy, and improve the conditions of oppressed people while simultaneously increasing the credibility of the science.

Reporting practices

Recommendation 9: Report details of methodology and decisions made throughout the research process; do not assume shared understanding with readers.

Recommendation 10: Reflect on, discuss, and report possible personal and contextual biases that could affect research design or interpretation.

Research reporting guidelines serve to enhance the transparency and thoroughness of research reports, enabling other researchers to scrutinize, build upon, and reproduce published results. For instance, these might include: (a) reporting all hypotheses, variables, interview questions, and analyses; (b) discussing constraints on generality; (c) providing contextual information about the researchers and the research context; and (d) employing meta-analytic thinking. Consistent with the other strategies we studied, calls for increased transparency have thus far focused on objective study elements. This contrasts with traditions within qualitative and, to a lesser extent, MM, research which also highlight the need to report subjective elements of research, such as researcher bias. We suspected that respondents with experience in qualitative methods would likely have much to offer to expand mainstream thinking on transparency and reporting guidelines.

Participants reported being more personally engaged with these strategies than any others we queried (see **Table 3** and **Figure 1**). One respondent wrote:

The key is credibility. Why should I believe you?... Where did the research question come from? What questions were asked to participants and why? What strategies were used to validate the questions before or during data collection, and what questions might have still remained confusing to participants? Were participants given a chance to “talk back” to the research process through open-ended questions...?

The handful of guidelines developed for MM research recognize that purely quantitative reporting standards fall short for this purpose (e.g., Leech & Onwuegbuzie, 2010; Moravcsik, 2010). These combine traditional “quantitative” guidelines with “qualitative” practices—most notably theoretical/methodological specificity and reflexivity (Levitt et al., 2018; O’Brien, Harris, Beckman, Reed, & Cook, 2014; Tong, Sainsbury, & Craig, 2007). Easily accessible reporting guidelines and toolkits are also available online (e.g., EQUATOR; www.equator-network.org).

Reflexivity emerged as a significant theme for survey participants. Researcher *reflexivity* refers to the process of reflecting on and disclosing researchers’ perspectives, context, and potential biases throughout the research process. Several respondents pointed out that reflexivity is an important tool for researchers regardless of their preferred methods, as no researcher can escape the personal biases they bring to research design, data collection, analysis, and interpretation. As one respondent described:

The researchers themselves act as the research instruments. Therefore, a different researcher would never be able to perfectly replicate a study.... factors such as the researcher’s age, sex, language skills, visible queerness or disability, familiarity

with participants’ pop culture references, etc. can all impact data collection.

Though the reflexive idea of the researcher-as-instrument may seem somewhat foreign to quantitative researchers, it is an appropriate response to the role of experimenter effects that have been well-documented at least since Clever Hans (Pfungst, 1965). The mere presence of a certain researcher can affect a study’s outcome, even when this is an unintended and unexplored factor (Rosenthal, 1966). This phenomenon has been exploited in experimental design, for example by varying confederates’ attractiveness (Barnes & Rosenthal, 1985) or race (Marx & Goff, 2005). Reflexive attention to the researcher-as-instrument provides a framework to think through and discuss these researcher effects. This approach can have substantial scientific benefits. As another respondent pointed out:

We’re encouraged to play down the individual motivation in our work, which leads to a complete lack of transparency...Additionally, there are elements of research that can be complemented by our own experiences and motivations, not inhibited by them. A great example of this is minority demographic group research, driven, or at least developed in consultation with, members of that group themselves!

For researchers who would like to be more reflexive in their reporting, several strategies are available. The most common strategies are: (a) keeping a research journal to track personal reactions, potential biases, and limitations; (b) discussing these issues as part of research teams’ standard operating procedure; and (c) explicitly reflecting on them in research reports. These actions all contribute to the goal of “making the familiar strange,” that is, calling our attention to aspects of our perspective that seem generic or universal but might appear idiosyncratic or privileged to others. Reflexive reporting communicates this information to colleagues and readers and gives them insight into the role the researcher-as-instrument might play in a study’s findings.

In the context of this paper, each author brings his or her own set of potential biases. While both authors are from culturally WEIRD (Western, educated, industrial, rich, democratic) backgrounds (Henrich, Heine, & Norenzayan, 2010), they differ somewhat in their life experiences and intellectual commitments. The first author has a strong interest and professional experience in social justice work, which informs her perspectives and emphasis on epistemology and values. The second author is committed to empiricism and pragmatism. His interest in MM research stems from the belief that current quantitative measures are functionally inadequate for capturing the qualia of subjective human experience. While the first author prefers to strike a balance between description and prescription (e.g., on ethics and values), the second author is most comfortable with innocuous descriptions of empirical findings. The two authors have discussed

these biases openly throughout the process of designing and reporting the current study.

Epistemology Drives Ethics

During data analysis, a strong theme connecting epistemology and ethics emerged. As one study respondent put it, the current moment in scientific progress has the potential to be an awakening “around how knowledge is constructed [and] the origins, nature, and limits of knowledge.” Quantitative methods seek universal knowledge through measurement, prediction, logic, and falsification. Qualitative methods, on the other hand, call upon the lived experiences of individuals to generate a *verisimilitude* (roughly, “lifelikeness”) that reflects knowledge about the nature of lived experience (Bruner, 1985; McAdams, 2012). As one respondent wrote, they would “want for people in the group being studied... to be able to look at the findings and say, *‘I can see that’* [emphasis added].”

What are the implications of research findings that resonate with participants’ lived experiences? Such findings demonstrate both external validity and ethical research. Consider, for example, that the vast majority of psychology studies published in leading journals observe and analyze a minority of the world’s population (Arnett, 2008; Henrich et al., 2010; Peterson, 2001). This is an issue not only because it hampers our ability to generalize such findings (which can be addressed in “constraints on generality” statements; see Simons et al., 2017), but also because underserved people around the world would be better served by more inclusive psychological research. One survey respondent offered another new norm for the field that speaks more directly to ethics:

I’d recommend that researchers (of all methodological leanings) ask “who benefits?” from the research they conduct, and “how?” Then move towards methodological choices that directly benefit participants and the public. The focus on replicability could then shift toward methods and research practices that tend to have a positive impact, and how to improve and expand upon those methods in working with diverse populations in diverse contexts.

This suggested move has dual purposes: first, bettering science so that it is comprised of more robust, diverse evidence; and second, adhering to psychology’s formal and principled imperatives to work toward the betterment of humanity (American Psychological Association, 2017a, 2017b; Josselson, 2007; Tindi & Sales, 2016). Because actions often have unexpected consequences, even psychological research that does not explicitly deal with ethical matters still has ethical implications (Louis, Mavor, La Macchia, & Amiot, 2014; Mustakova-Possardt, Lyubansky, Basseches, & Oxenberg, 2014; Toporek, Kwan, & Williams, 2012). There is much opportunity in the rapidly iterating open science movement to integrate ethical foci into dialogue and practice (already, see Christensen, 2018; Nosek et al., 2015; Wicherts et al., 2016). When we critically examine the epistemological assumptions that

drive our research, we challenge ourselves to be better scientists and challenge our science to do more good in the world.

Study Limitations

Respondents did not hesitate to point out one limitation of the current study; as one respondent observed, “the survey seems focused on a quantitative conception of mixed methods research.” Another cautioned that our approach was an example of researchers “trying to shoehorn non-quantitative, interpretive, meaning-centered methods into the old-fashioned experimentalist hypothesis-testing paradigms.” We agree with the spirit of these critiques, and also note that this stance was by design. Our first research question focused on MM applications of reproducibility best practices which were designed for quantitative studies. Thus, questions about specific issues and strategies were necessarily presented from a quantitative perspective. We believe that an even more qualitative approach to examining issues of reproducibility in MM research would be highly valuable, and we hope that the current study will serve as a launching point for more qualitatively-oriented studies. Similarly, our survey was not exhaustive. We did not survey respondents about many relevant topics such as reliability and validity, or specific data analysis tools. Moreover, one respondent noted that the complex nature of these concepts and their associated jargon made the survey wording unwieldy. This may have also been the case for other respondents.

Most importantly, the sample size of our survey was small, particularly compared to previous survey studies on psychological reproducibility which have reported sample sizes over 1,000 (John et al., 2012; Washburn et al., 2018). However, at its core, this project was theoretical rather than empirical. Throughout, we have treated survey results as illustrative rather than definitive. Moreover, the prevalence of MM research in psychology is low. A study comparing prevalence of research designs across disciplines found that only 7% of published psychology studies employed MM (Alise & Teddlie, 2010). A survey targeting a small fraction of a field should be expected to generate a small fraction of broader surveys’ responses. It is entirely possible that the current sample is not representative of all people who identify as MM psychological researchers. Indeed, our recruitment materials were designed to appeal to MM researchers who may have observed a lack of MM researchers’ voices in the reproducible social science debate, perhaps resulting in a sample more interested in the topic than the average MM researcher. Nevertheless, we believe the respondents’ insights offer a valuable window into the experiences and concerns of MM researchers.

Conclusion

There is a conspicuous lack of discussion and information around best practices for increasing the credibility of psychological science in the context of MM. In an effort to fill this gap, we surveyed social science researchers who self-identified as mixed methodologists on issues and strategies related to reproducibility. We agree with Poortman and Schildkamp (2012) that different criteria

for good science in quantitative, qualitative, and MM research are not needed. Instead, methodologists of all stripes would do well to learn each other's languages, and realize their own methods' unique advantages and limitations (Osbeck, 2014). In an effort to facilitate this methodological and epistemological exchange, we reported the most salient insights and best practices that emerged from the unique perspectives of our MM survey respondents (see **Table 2** for our ten recommendations). By including MM and qualitative approaches in the reproducibility conversation, psychological scientists of diverse methodological traditions can only enhance their ability to produce meaningful, credible science.

Appendix

Sample data-sharing request template

Project title

Principal investigator and institutional information

Date:

Requester/institutional affiliation:

Data requested. *Be as specific as possible (e.g., which interview questions or self-report scales?)*

Purpose of request. *What will the data be used for? Describe the study aims, hypotheses, and what role the data will play.*

Data storage and dissemination. *Where and how will the data be stored? Who will have access? Will the data be shared with any collaborators? How do you intend to publish the data?*

Protecting participant confidentiality. *How will risks to participant confidentiality be managed? Will any special precautions be taken?*

Data Accessibility Statement

This study was preregistered with the Open Science Framework (OSF) at <https://osf.io/fkxm7/>. The preregistration adheres to the disclosure requirements of OSF. The preregistration and all data in the manuscript is available at this link.

Notes

¹ The terms “replicability” and “reproducibility” have generated much definitional debate. In the current paper, we use the broad term “reproducibility” following Goodman et al. (2016) and consider replication studies as a specific, narrower topic (analogous to “results reproducibility” in Goodman’s terminology).

² The status and nature of psychology as a science has been the focus of entire literatures (e.g., Fanelli, 2010; Ferguson, 2015; Hatfield, 2002; Lilienfeld, 2012; Meehl, 1978, 1986; Pérez-Álvarez, 2018). For the purpose of this article, we assume psychology is

an empirical science of some kind(s) and attempt to synthesize perspectives of psychology as natural science and psychology as a human science.

³ *Grounded theory* is a qualitative data analysis technique in which researchers begin with no a priori hypotheses about the construct of interest, and instead use induction to generate theory directly from the data collected. For a description of grounded theory methodology, see for example Charmaz and Henwood (2017).

⁴ In our online survey, this category included both replication (i.e., running a study a second time) and reproduction (i.e., re-analyzing a study’s original data). Because respondents focused on replication studies (or “results reproducibility” in the terminology of Goodman et al., 2016), and to avoid confusion with the overall topic of the paper, here we refer to the category as “replication studies.”

Competing Interests

The authors have no competing interests to declare.

Author Contributions

- Contributed to conception and design: HNR, HRC
- Contributed to acquisition of data: HNR, HRC
- Contributed to analysis and interpretation of data: HNR, HRC
- Drafted and/or revised the article: HNR, HRC
- Approved the submitted version for publication: HNR, HRC

References

- Aguinis, H., & Solarino, A. M.** (2019). Transparency and replicability in qualitative research: The case of interviews with elite informants. *Strategic Management Journal, 40*(8), 1291–1315. DOI: <https://doi.org/10.1002/smj.3015>
- Alise, M. A., & Teddlie, C.** (2010). A continuation of the paradigm wars? Prevalence rates of methodological approaches across the social/behavioral sciences. *Journal of Mixed Methods Research, 4*(2), 103–126. DOI: <https://doi.org/10.1177/1558689809360805>
- American Psychological Association.** (2017a). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017). <http://www.apa.org/ethics/code/index.html>
- American Psychological Association.** (2017b). *Multicultural Guidelines: An Ecological Approach to Context, Identity, and Intersectionality*. Retrieved from <http://www.apa.org/about/policy/multicultural-guidelines.pdf>
- Anderson, C. J., Bahník, Š., Barnett-Cowan, M., Bosco, F. A., Chandler, J., Chartier, C. R., ... Zuni, K.** (2016). Response to Comment on “Estimating the reproducibility of psychological science.” *Science, 351*(6277), 1037. DOI: <https://doi.org/10.1126/science.aad9163>
- Arnett, J. J.** (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602–614. DOI: <https://doi.org/10.1037/0003-066X.63.7.602>

- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K., ... Wicherts, J. M.** (2012). Recommendations for increasing replicability in psychology. *European Journal of Personality*, *119*, 108–119. DOI: <https://doi.org/10.1002/per.1919>
- Barker, C., & Pistrang, N.** (2005). Quality criteria under methodological pluralism: Implications for conducting and evaluating research. *American Journal of Community Psychology*, *35*(3–4), 201–212. DOI: <https://doi.org/10.1007/s10464-005-3398-y>
- Barnes, M. L., & Rosenthal, R.** (1985). Interpersonal effects of experimenter attractiveness, attire, and gender. *Journal of Personality and Social Psychology*, *48*(2), 435–446. DOI: <https://doi.org/10.1037/0022-3514.48.2.435>
- Behrendt, S.** (2014). lm.beta: Add standardized regression coefficients to lm-Objects. Retrieved from <https://cran.r-project.org/package=lm.beta>
- Bertamini, M., & Munafò, M. R.** (2012). Bite-size science and its undesired side effects. *Perspectives on Psychological Science*, *7*(1), 67–71. DOI: <https://doi.org/10.1177/1745691611429353>
- Braun, V., & Clarke, V.** (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*. DOI: <https://doi.org/10.1191/1478088706qp0630a>
- Broom, A., Cheshire, L., & Emmison, M.** (2009). Qualitative researchers' understandings of their practice and the implications for data archiving and sharing. *Sociology*, *43*(6), 1163–1180. DOI: <https://doi.org/10.1191/1478088706qp0630a>
- Bruner, J.** (1985). Narrative and paradigmatic modes of thought. *Learning and Teaching: The Ways of Knowing*, *84*, 97–115. DOI: <https://doi.org/10.1016/j.jconhyd.2010.08.009>
- Chambers, C.** (2017). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton University Press. DOI: <https://doi.org/10.1515/9781400884940>
- Charmaz, K., & Henwood, K.** (2017). Grounded theory methods for qualitative psychology. In C. Willig & W. S. Rogers (Eds.), *The SAGE Handbook of Qualitative Research in Psychology* (pp. 238–256). London: SAGE Publications. DOI: <https://doi.org/10.4135/9781526405555.n14>
- Christensen, G.** (2018). *Manual of best practices in transparent social science research*. Berkeley, CA. Retrieved from <https://github.com/garretchristensen/BestPracticesManual/blob/master/Manual.pdf>
- Crandall, C. S., & Sherman, J. W.** (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, *66*, 93–99. DOI: <https://doi.org/10.1016/j.jesp.2015.10.002>
- DuBois, J. M., Walsh, H., & Strait, M.** (2018). It is time to share (some) qualitative data: Reply to Guishard (2018), McCurdy and Ross (2018), and Roller and Lavrakas (2018). *Qualitative Psychology*, *5*(3), 412–415. DOI: <https://doi.org/10.1037/qup0000092>
- Fanelli, D.** (2010). “Positive” results increase down the hierarchy of the sciences. *PLoS ONE*, *5*(4), e10068. DOI: <https://doi.org/10.1371/journal.pone.0010068>
- Ferguson, C. J.** (2015). “Everybody knows psychology is not a real science”: Public perceptions of psychology and how we can improve our relationship with policymakers, the scientific community, and the general public. *American Psychologist*, *70*(6), 527–542. DOI: <https://doi.org/10.1037/a0039405>
- 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. DOI: <https://doi.org/10.1037/pspi0000007>
- Freeman, M.** (2011). Toward poetic science. *Integrative Psychological and Behavioral Science*, *45*(4), 389–396. DOI: <https://doi.org/10.1007/s12124-011-9171-x>
- Freese, J., & Peterson, D.** (2017). Replication in social science. *Annual Review of Sociology*, *43*(1), 147–165. DOI: <https://doi.org/10.1146/annurev-soc-060116-053450>
- Freire, P.** (1970). *Pedagogy of the oppressed*. New York: Continuum. DOI: <https://doi.org/10.1086/642915>
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D.** (2016). Comment on “Estimating the reproducibility of psychological science.” *Science*, *351*(6277), 1037–1037. DOI: <https://doi.org/10.1126/science.aad7243>
- Giorgi, A.** (1969). Psychology: A human science. *Social Research*, *36*(3), 412–432. Retrieved from <http://www.jstor.org/stable/40969976>
- Goodman, A., Pepe, A., Blocker, A. W., Borgman, C. L., Cranmer, K., Crosas, M., ... Slavkovic, A.** (2014). Ten simple rules for the care and feeding of scientific data. *PLoS Computational Biology*, *10*(4), e1003542. DOI: <https://doi.org/10.1371/journal.pcbi.1003542>
- Goodman, S. N., Fanelli, D., & Ioannidis, J. P. A.** (2016). What does research reproducibility mean? *Science Translational Medicine*, *8*(341). DOI: <https://doi.org/10.1126/scitranslmed.aaf5027>
- Greenfield, P. M.** (2017). Cultural change over time: Why replicability should not be the gold standard in psychological science. *Perspectives on Psychological Science*, *12*(5), 762–771. DOI: <https://doi.org/10.1177/1745691617707314>
- Guishard, M. A.** (2018). Now's not the time! Qualitative data repositories on tricky ground: Comment on Dubois et al. (2018). *Qualitative Psychology*, *5*(3), 402–408. DOI: <https://doi.org/10.1037/qup0000085>
- Hallgren, K. A.** (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34. <https://doi.org/10.1016/j.biotechadv.2011.08.021>. Secreted. DOI: <https://doi.org/10.20982/tqmp.08.1.p023>
- Hatfield, G.** (2002). Psychology, philosophy, and cognitive science: Reflections on the history and philosophy of experimental psychology. *Mind*

- and *Language*, 17(3), 207–232. DOI: <https://doi.org/10.1111/1468-0017.00196>
- Haven, T. L., & Van Grootel, L.** (2019). Preregistering qualitative research. *Accountability in Research*, 26(3), 229–244. DOI: <https://doi.org/10.1080/08989621.2019.1580147>
- Henrich, J., Heine, S. J., & Norenzayan, A.** (2010). The weirdest people in the world? *Behavioral and Brain Sciences*. DOI: <https://doi.org/10.1017/S0140525X0999152X>
- International Covenant on Economic, Social and Cultural Rights**, New York, 16 December 1966, *United Nations Treaty Series*, vol. 993, No. 14531, available at: https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtmsg_no=IV-3&chapter=4&clang=en
- Ioannidis, J. P. A.** (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), 0696–0701. DOI: <https://doi.org/10.1371/journal.pmed.0020124>
- Joel, S., Eastwick, P. W., & Finkel, E. J.** (2018). Open sharing of data on close relationships and other sensitive social psychological topics: Challenges, tools, and future directions. *Advances in Methods and Practices in Psychological Science*, 1(1), 86–94. DOI: <https://doi.org/10.1177/2515245917744281>
- 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. DOI: <https://doi.org/10.1177/0956797611430953>
- Johnson, R. B., & Onwuegbuzie, A. J.** (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*. DOI: <https://doi.org/10.3102/0013189X033007014>
- Josselson, R.** (2007). The ethical attitude in narrative research: Principles and practicalities. In *Handbook of Narrative Inquiry: Mapping a Methodology* (pp. 537–566). DOI: <https://doi.org/10.4135/9781452226552>
- Judd, C. M., Westfall, J., & Kenny, D. A.** (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, 103(1), 54–69. DOI: <https://doi.org/10.1037/a0028347>
- Kaiser, K.** (2009). Protecting respondent confidentiality in qualitative research. *Qualitative Health Research*, 19(11), 1632–1641. DOI: <https://doi.org/10.1177/1049732309350879>
- Kern, F. G., & Gleditsch, K. S.** (2017). *Exploring Pre-registration and Pre-analysis Plans for Qualitative Inference*. Retrieved from https://www.researchgate.net/publication/319141144_Exploring_Pre-registration_and_Pre-analysis_Plans_for_Qualitative_Inference
- Leech, N. L., & Onwuegbuzie, A. J.** (2010). Guidelines for conducting and reporting mixed research in the field of counseling and beyond. *Journal Of Counseling & Development*, 88(1), 61–70. DOI: <https://doi.org/10.1002/j.1556-6678.2010.tb00151.x>
- Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C.** (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 26–46. DOI: <https://doi.org/10.1037/amp0000151>
- Lilienfeld, S. O.** (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist*, 67(2), 111–129. DOI: <https://doi.org/10.1037/a0023963>
- Lindsay, D. S.** (2017). Sharing data and materials in psychological science. *Psychological Science*, 28(6), 699–702. DOI: <https://doi.org/10.1177/0956797617704015>
- Lindsay, D. S., Simons, D. J., & Lilienfeld, S. O.** (2016, December). Research preregistration 101. *Observer*, 14–17. Retrieved from https://www.psychologicalscience.org/redesign/wp-content/uploads/2016/12/December_OBS_web_12-2-2016.pdf
- Louis, W. R., Mavor, K. I., La Macchia, S. T., & Amiot, C. E.** (2014). Social justice and psychology: What is, and what should be. *Journal of Theoretical and Philosophical Psychology*, 34(1), 14–27. DOI: <https://doi.org/10.1037/a0033033>
- Martone, M. E., Garcia-Castro, A., & VandenBos, G. R.** (2018). Data sharing in psychology. *American Psychologist*, 73(2), 111–125. DOI: <https://doi.org/10.1037/amp0000242>
- Marx, D. M., & Goff, P. A.** (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology*, 44(4), 645–657. DOI: <https://doi.org/10.1348/014466604X17948>
- McAdams, D. P.** (2012). Exploring psychological themes through life-narrative accounts. In *Varieties of narrative analysis* (pp. 15–32). DO: <https://doi.org/10.4135/9781506335117.n2>
- McCurdy, S. A., & Ross, M. W.** (2018). Qualitative data are not just quantitative text data with text but data with context: On the dangers of sharing some qualitative data: Comment on Dubois et al. (2018). *Qualitative Psychology*, 5(3), 409–411. DOI: <https://doi.org/10.1037/qup0000088>
- McLean, K. C., Syed, M., Pasupathi, M., Adler, J. M., Dunlop, W. L., Drustrup, D., ... McCoy, T. P.** (2019). The empirical structure of narrative identity: The initial big three. *Journal of Personality and Social Psychology*. DOI: <https://doi.org/10.1037/pspp0000247>
- Medin, D. L.** (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*, 12(4), 669–674. DOI: <https://doi.org/10.1177/1745691616687746>
- Meehl, P. E.** (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46(4), 806–834. DOI: <https://doi.org/10.1037/0022-006X.46.4.806>

- Meehl, P. E.** (1986). What social scientists don't understand. In D. W. Fiske & R. A. Shweder (Eds.), *Metatheory in Social Science: Pluralisms and Subjectivities* (pp. 315–338). Chicago: University of Chicago Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=_btWEg4oqsMC&pgis=1
- Merton, R. K.** (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago: The University of Chicago Press. DOI: <https://doi.org/10.2307/2063284>
- Miguel, E., Camerer, C., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., ... Van der Laan, M.** (2014). Promoting transparency in social science research. *Science*, *343*(6166), 30–31. DOI: <https://doi.org/10.1126/science.1245317>
- Moravcsik, A.** (2010). Active citation: A precondition for replicable qualitative research. *PS: Political Science & Politics*, *43*(01), 29–35. DOI: <https://doi.org/10.1017/S1049096510990781>
- Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., ... Zwaan, R. A.** (2016). The Peer Reviewers' Openness Initiative: Incentivizing open research practices through peer review. *Royal Society Open Science*, *3*(1), 150547. DOI: <https://doi.org/10.1098/rsos.150547>
- Mustakova-Possardt, E., Lyubansky, M., Basseches, M., & Oxenberg, J.** (Eds.). (2014). *Toward a socially responsible psychology for a global era*. New York: Springer. DOI: <https://doi.org/10.1007/978-1-4614-7391-6>
- Newman, I., & Hitchcock, J. H.** (2011). Underlying agreements between quantitative and qualitative research. *Human Resource Development Review*, *10*(4), 381–398. DOI: <https://doi.org/10.1177/1534484311413867>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... Chambers, C. D.** (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. DOI: <https://doi.org/10.1126/science.aab2374>
- Nuijten, M. B., Borghuis, J., Veldkamp, C. L. S., Dominguez-Alvarez, L., Van Assen, M. A. L. M., & Wicherts, J. M.** (2017). Journal data sharing policies and statistical reporting inconsistencies in psychology. *Collabra: Psychology*, *3*(1), 31. DOI: <https://doi.org/10.1525/collabra.102>
- Nzinga, K., Rapp, D., Leatherwood, C., Easterday, M., Rogers, L., Gallagher, N., & Medin, D. L.** (2018). Should social scientists be distanced from or engaged with the people they study? *PNAS*, *115*(45).
- O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A.** (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, *89*(9), 1245–1251. DOI: <https://doi.org/10.1097/ACM.0000000000000388>
- Onwuegbuzie, A. J., & Leech, N. L.** (2007). A call for qualitative power analyses. *Quality and Quantity*. DOI: <https://doi.org/10.1007/s11135-005-1098-1>
- Osbeck, L. M.** (2014). Scientific reasoning as sense-making: Implications for qualitative inquiry. *Qualitative Psychology*, *1*(1), 34–46. DOI: <https://doi.org/10.1037/qup0000004>
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K.** (2015). The development and psychometric properties of LIWC2015. *UT Faculty/Researcher Works*. DOI: <https://doi.org/10.15781/T29G6Z>
- Pérez-Álvarez, M.** (2018). Psychology as a science of subject and comportment, beyond the mind and behavior. *Integrative Psychological and Behavioral Science*, *52*(1), 25–51. DOI: <https://doi.org/10.1007/s12124-017-9408-4>
- Peterson, R. A.** (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*. DOI: <https://doi.org/10.1086/323732>
- Pfungst, O.** (1965). *Clever Hans: The horse of Mr. Von Osten*. Oxford, England: Holt, Rinehart & Winston, Inc.
- Piñeiro, R., & Rosenblatt, F.** (2016). Pre-analysis plans for qualitative research. [Planes Pre-Análisis para la investigación cualitativa]. *Revista de Ciencia Política*, *36*(3), 785–796. DOI: <https://doi.org/10.4067/S0718-090X2016000300009>
- Polanin, J. R., & Terzian, M.** (2019). A data-sharing agreement helps to increase researchers' willingness to share primary data: results from a randomized controlled trial. *Journal of Clinical Epidemiology*, *106*, 60–69. DOI: <https://doi.org/10.1016/j.jclinepi.2018.10.006>
- Poortman, C. L., & Schildkamp, K.** (2012). Alternative quality standards in qualitative research? *Quality and Quantity*, *46*(6), 1727–1751. DOI: <https://doi.org/10.1007/s11135-011-9555-5>
- Prilleltensky, I.** (1996). Human, moral, and political values for an emancipatory psychology. *The Humanistic Psychologist*, *24*(3), 307–324. DOI: <https://doi.org/10.1080/08873267.1996.9986859>
- R Core Team.** (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org>
- Rappaport, J.** (2005). Community psychology is (thank God) more than science. *American Journal of Community Psychology*, *35*(3–4), 231–238. DOI: <https://doi.org/10.1007/s10464-005-3402-6>
- Revelle, W.** (2018). psych: Procedures for Personality and Psychological Research (Version 1.8.3). Evanston, IL: Northwestern University. Retrieved from <https://cran.r-project.org/package=psych>
- Roberts, K., Dowell, A., & Nie, J. B.** (2019). Attempting rigour and replicability in thematic analysis of qualitative research data; A case study of codebook development. *BMC Medical Research Methodology*, *19*(1), 1–8. DOI: <https://doi.org/10.1186/s12874-019-0707-y>
- Robinson, D. G.** (2015). broom: An R package for converting statistical analysis objects into tidy data frames. Retrieved from <https://cran.r-project.org/package=broom>
- Roller, M. R., & Lavrakas, P. J.** (2018). A total quality framework approach to sharing qualitative research data: Comment on DuBois et al. (2018). *Qualitative Psychology*, *5*(3), 394–401. DOI: <https://doi.org/10.1037/qup0000081>

- Rosenthal, R.** (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.
- Saunders, B., Kitzinger, J., & Kitzinger, C.** (2015). Anonymising interview data: challenges and compromise in practice. *Qualitative Research, 15*(5), 616–632. DOI: <https://doi.org/10.1177/1468794114550439>
- 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*(11), 1359–1366. DOI: <https://doi.org/10.1177/0956797611417632>
- Simons, D. J.** (2014). The value of direct replication. *Perspectives on Psychological Science, 9*(1), 76–80. DOI: <https://doi.org/10.1177/1745691613514755>
- Simons, D. J., Shoda, Y., & Lindsay, D. S.** (2017). Constraints on Generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123–1128. DOI: <https://doi.org/10.1177/1745691617708630>
- Smaldino, P. E., & McElreath, R.** (2016). The natural selection of bad science. *Royal Society Open Science, 3*(9), 160384. DOI: <https://doi.org/10.1098/rsos.160384>
- Sturdy, J., Burch, S., Hanson, H., & Molyneaux, J.** (2017). Opening up evaluation microdata: Balancing risks and benefits of research transparency. *BITSS Preprints*. DOI: <https://doi.org/10.31222/osf.io/s67my>
- Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., ... Shrout, P. E.** (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science, 12*(5), 742–756. DOI: <https://doi.org/10.1177/1745691617690042>
- Tindi, A., & Sales, J.** (2016). Social justice and ethics: Dilemmas and opportunities for psychologists. Retrieved July 18, 2018, from http://www.mnpsych.org/index.php?option=com_dailyplanetblog&view=entry&category=event+recap&id=160:social-justice-and-ethics-dilemmas-and-opportunities-for-psychologists
- Tong, A., Sainsbury, P., & Craig, J.** (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*(6), 349–357. DOI: <https://doi.org/10.1093/intqhc/mzm042>
- Toporek, R. L., Kwan, K.-L. K., & Williams, R. A.** (2012). Ethics and social justice in counseling psychology. In *APA Handbook of Counseling Psychology* (Vol. 2). DOI: <https://doi.org/10.1037/13755-013>
- van 't Veer, A. E., & Giner-Sorolla, R.** (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology, 67*, 2–12. DOI: <https://doi.org/10.1016/j.jesp.2016.03.004>
- 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*(2), 166–173. DOI: <https://doi.org/10.1177/2515245918757427>
- Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D.** (2006). The poor availability of psychological research data for reanalysis. *American Psychologist, 61*(7), 726–728. DOI: <https://doi.org/10.1037/0003-066X.61.7.726>
- Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M.** (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid *p*-hacking. *Frontiers in Psychology, 7*(NOV), 1–12. DOI: <https://doi.org/10.3389/fpsyg.2016.01832>
- Wickham, H.** (2016a). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag. ISBN 978-3-319-24277-4, Retrieved from <https://ggplot2.tidyverse.org>
- Wickham, H.** (2016b). tidyverse. *R package version 1.0.0*. Retrieved from <https://cran.r-project.org/package=tidyverse>
- Zandbergen, P. A.** (2014). Ensuring confidentiality of geocoded health data: Assessing geographic masking strategies for individual-level data. *Advances in Medicine, 2014*, 1–14. DOI: <https://doi.org/10.1155/2014/567049>
- Zörgő, S., & Peters, G. Y.** (2019). *The ROCK book: The Reproducible Open Coding Kit* (1st ed.). DOI: <https://doi.org/10.31234/osf.io/stcx9>

How to cite this article: Reischer, H. N., & Cowan, H. R. (2020). Quantity Over Quality? Reproducible Psychological Science from a Mixed Methods Perspective. *Collabra: Psychology, 6*(1): 26. DOI: <https://doi.org/10.1525/collabra.284>

Senior Editor: Simine Vazire

Editor: Simine Vazire

Submitted: 21 August 2019

Accepted: 15 April 2020

Published: 14 May 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.