

Kent Academic Repository

Full text document (pdf)

Citation for published version

LeBel, Etienne P. and Borsboom, Denny and Giner-Sorolla, Roger and Hasselman, Fred and Peters, Kurt R. and Ratliff, Kate A. and Smith, Colin Tucker (2013) PsychDisclosure.org: Grassroots support for reforming reporting standards in psychology. *Perspectives on Psychological Science*, 8 (4). pp. 424-432. ISSN 1745-6916.

DOI

<https://doi.org/10.1177/1745691613491437>

Link to record in KAR

<http://kar.kent.ac.uk/36642/>

Document Version

Publisher pdf

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

PsychDisclosure.org: Grassroot Support for Reforming Reporting Standards in Psychology

Etienne P. LeBel
University of Western Ontario

Denny Borsboom
University of Amsterdam

Roger Giner-Sorolla
University of Kent

Fred Hasselman
Radboud University Nijmegen

Kurt R. Peters
ICF International

Kate A. Ratliff Colin Tucker Smith
University of Florida

There is currently an unprecedented level of doubt regarding the reliability of research findings in psychology. Many recommendations have been made to improve the current situation. This article reports results from PsychDisclosure.org, a novel open-science initiative that provides a platform for authors of recently published articles to disclose four methodological design specification details that are not required to be disclosed under current reporting standards, but which are critical for accurate interpretation and evaluation of reported findings. Grassroots sentiment -- as manifested in the positive and appreciative response to our initiative -- indicates that psychologists want to see changes made at the systemic level regarding disclosure of such methodological details. Almost 50% of contacted researchers disclosed the requested design specifications for the four methodological categories (excluded subjects, non-reported conditions and measures, and sample size determination). Disclosed information provided by participating authors also revealed several instances of questionable editorial practices, which need to be thoroughly examined and redressed. Based on these results, we argue that the time is now for mandatory methods disclosure statements for all psychology journals, which would be an important step forward in improving the reliability of findings in psychology.

Key words: Reporting standards, disclosure, methodological design specifications, methodology

The goal of science is to amass cumulative knowledge about how the world works. In psychology, this means accruing cumulative knowledge regarding the prediction and explanation of human behavior. Recently, there has been a growing concern regarding the non-replicability of research findings in psychology (Ioannidis, 2012; Pashler & Harris, 2012) and neuroscience (Button et al., 2013), with some even considering it a "crisis of confidence" (Pashler & Wagenmakers, 2012).¹ Researchers have considered many factors contributing to replication difficulties in psychology including the near-exclusive focus on publishing novel findings (Giner-Sorolla, 2012; Nosek, Spies, & Motyl, 2012), an over-emphasis on conceptual replications (LeBel & Peters, 2011), the use of unreliable measures (LeBel & Paunonen, 2011), and an extreme dearth of independent direct replications (i.e., less than 0.1% of published articles in the last 100 years; Makel, Plucker, & Hegarty, 2012). The use of questionable research practices (QRPs) -- or *researcher degrees of freedom* -- has also been highlighted as contributing to false-positive results in psychology (John, Loewenstein, & Prelec, 2012; Simmons, Nelson, & Simonsohn, 2011) and hence contributing to the non-replicability of findings.

In the wake of this crisis, a number of initiatives and recommendations have been made to improve the

reliability of psychological findings. One prominent example of such an initiative is a website (PsychFileDrawer.org/) where researchers can post failed (or successful) replications in an attempt to combat the oft-referenced file drawer problem in which non-significant (or redundant) findings typically fail to be published. Another example is a large-scale effort called the Reproducibility Project (Open Science Collaboration, 2012), in which researchers are investigating factors that predict the reproducibility of psychological findings.² The initiative described in the current work was inspired by these action-oriented approaches and by a recommendation that seemed of critical importance to us. Specifically, Simmons, Nelson, and Simonsohn (2012; see also Simmons, Nelson, & Simonsohn, 2011) recommended that authors submitting manuscripts include a 21-word disclosure statement stating that all exclusions, measures, experimental conditions, and sample size determination rules have been reported.³ This statement is intended to allow the reader (and reviewers) to rule out and/or account for *researcher degrees of freedom* that authors might have exploited

¹ Following Asendorpf et al. (in press), we distinguish *replicability* (obtaining the same finding using the same methods with a different random sample) from *reproducibility* (obtaining the same result when executing the same analyses on the original dataset).

² A few journals now have mechanisms for incentivizing direct replications including new pre-registered replication manuscript categories at *Perspectives on Psychological Science* and *Cortex* and a special issue on direct replications of important findings in social psychology at *Social Psychology*.

³ Simmons et al.'s 21-word disclosure statement reads: *We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.*

in order to obtain the reported results. For instance, knowing that all measures/items were reported, readers and reviewers can assess the extent to which multiple testing contributed to achieving statistical significance of particular results. Likewise, by stating that data collection stopped once a predetermined sample size was reached, readers and reviewers can rule out that optional stopping was used (i.e., analyzing the data as they accumulate and halting the study whenever results reach statistical significance; Armitage, McPherson, & Rowe, 1969; Feller, 1940).

PsychDisclosure.org: Nature and Aims

In line with these observations, and to further stimulate the realization of an open and honest research culture, we launched PsychDisclosure.org on November 17, 2012. PsychDisclosure.org is a novel open science initiative intended to contribute to improving the reliability of psychological findings by providing a platform for authors to publicly disclose four categories of methodological design specifications (i.e., excluded subjects, non-reported conditions and measures, and sample size determination) that are not required to be disclosed under current reporting standards, but which are essential for accurate interpretation and evaluation of research findings. We use the term *design specification statements* to refer to the additional information disclosed by participating authors. Another goal of the initiative is to investigate the extent to which psychological journals' reporting standards might be suboptimal (i.e., to investigate the prevalence of important research design decisions not reported in actual publications). If such practices are indeed widespread, this suggests that disclosure statements – of the kind proposed by Simmons et al. – should in fact become a required component of the manuscript submission process. The current article reports the first results from the initiative, which unequivocally support a policy change for manuscript submissions.

PsychDisclosure.org: Authors contacted.

Using a bottom-up, grassroots approach, our initiative involved emailing a random 50% of all corresponding authors of articles published in 2012 and onward in the prominent journals *Psychological Science* (*PS*),

Journal of Personality and Social Psychology (*JPSP*), *Journal of Experimental Psychology: Learning, Memory, and Cognition* (*JEP:LMC*), and *Journal of Experimental Psychology: General* (*JEP:G*).⁴ We asked authors whether they were interested in publicly disclosing four details of the methods used to obtain findings for all studies reported (see Supplementary Materials for the email text). The content of their replies would then be posted on the PsychDisclosure.org website. Only a random 50% of corresponding authors in each issue were contacted for disclosure, to mitigate respondent and institutional concerns regarding the anonymity of non-respondents. This way, readers viewing the disclosed information on PsychDisclosure.org have no way of knowing if articles are absent because authors decided not to disclose the information or because authors were never invited.

PsychDisclosure.org: Results. As of March 5, 2013, 46.4% (i.e., 161 of 347) of contacted authors have replied to our email and publicly disclosed the requested information. As can be seen in Table 1, the response rates were similar across the four journals, though slightly higher for *JEP:LMC* and *JEP:G*.

In the actual published articles, rates for fully disclosing methodological design specifications varied widely across the four methodological categories. Overall, the vast majority of researchers stated that they had fully reported all excluded observations (88.8%) and fully reported all tested experimental conditions (88.2%). However, only 54.7% reported all assessed measures/items and a mere 11.2% reported their data collection stopping rule. These full disclosure rates were similar across the four journals, except for the *measures* category, in which full disclosure rates were markedly higher for *JEP:LMC* and *JEP:G*. (87.2% and 81.8%, respectively, vs. 46.7% and 20.0%; see Table 1).

For methodological design specifications that had not been reported in the published article, the reasons for failing to disclose these details varied substantially across the four categories (see Figure 1).⁵ Of particular note for conditions, 42% (8 out of 19) of the design specification statements stated that additional tested experimental conditions were not reported

Table 1. Response rates to our email and reported rates of full disclosure within the published articles in each journal.

Journal	N	Response Rate	% Full Disclosure for Methodological Categories			
			Exclusions	Conditions	Measures	Sample Size
<i>Psychological Science</i>	131	45.8	88.3	88.3	46.7	15.0
<i>Journal of Personality and Social Psychology</i>	92	43.5	92.5	87.5	20.0	10.0
<i>Journal of Experimental Psychology: Learning, Memory, & Cognition</i>	79	49.4	84.6	84.6	87.2	5.1
<i>Journal of Experimental Psychology: General</i>	45	48.9	90.9	95.5	81.8	13.6
Overall	347	46.4	88.8	88.2	54.7	11.2

Note. All numbers represent percentages, except the N column, which indicates the number of articles for each journal selected for inclusion as of March 5, 2013. Full disclosure rates indicate the percentage of articles wherein authors answered "Yes", indicating they had fully reported the respective methodological design specifications in the published article.

because statistically significant differences did not emerge in those conditions. Also, one participating author (5% of design specification statements in this category) stated that not all conditions were reported due to an editorial request because the data in those conditions revealed a more complex pattern. Even though this was found for only 1 out of 19 statements in this category, such editorial requests are striking given that they mischaracterize the evidence.

due to psychometric problems, which we thought was surprising given we expected this to be higher (LeBel & Peters, 2011). Approximately 5% went unreported due to editorial requests, which again we found striking given the clear sub-optimality of such a practice.

Finally, for sample size, about 44% of design specification statements indicated that data collection halted once a predetermined sample size was reached and about 4% (6 of 143) collected more data due to editorial requests. Of note, about 5% of our respondents stated they continued collecting data until the relevant patterns were clear, which is surprising given that these statements amount to public admissions of optional stopping.

PsychDisclosure.org Benefits

PsychDisclosure.org has benefits for both the researchers reporting their results and for the field in general.

Researcher benefits. The main benefit of researchers publicly disclosing the four methodological design specifications on PsychDisclosure.org is to increase the information value of their published article. With the additional disclosed information, readers of the article in question can more accurately interpret the reported results in light of the claim that many researchers have engaged in so-called QRPs (John et al., 2012). John et al. found that the majority of researchers -- in a sample of over 2,000 psychologists -- admitted to not always reporting all of a study's dependent measures or to deciding to collect more data after looking to see whether the results were statistically significant. When the four categories of methodological design specifications are disclosed, however, readers can (a) effectively rule out the possibility that reported results are partly due to unreported

design specifications (in the case of full disclosure) or (b) arrive at a more nuanced interpretation of the findings given their judgment of the "principledness" of the reasons provided for not disclosing the methodological information in the published article. For example, if a corresponding author disclosed that all excluded observations, tested experimental conditions, and all assessed measures/items were reported in the published article and that data collection stopped once

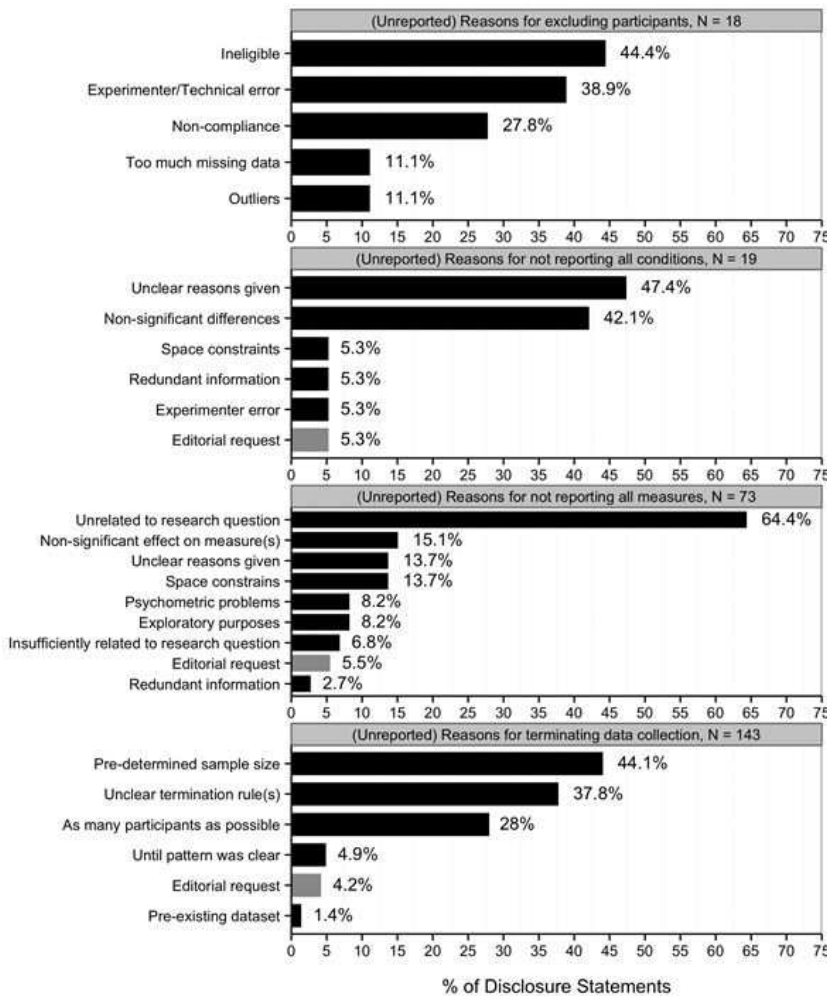


Figure 1. Reasons given for not including the methodological design specifications in the published article (later provided to PsychDisclosure.org). Within each disclosure category, totals are greater than 100% because more than one reason was sometimes mentioned. Ns indicate the number of design specification statements the percentages are based on (out of the 347 possible articles).

For measures, it is noteworthy that 15% (11 of 73) of design specification statements mentioned that assessed measures went unreported because no statistically significant differences emerged on those measures, which is clear evidence of *researcher degrees-of-freedom* (John et al., 2012; Simmons et al., 2011). About 8% of design specification statements indicated that additional measures went unreported

an a priori sample size was reached, then a reader can effectively rule out alternative hypotheses that may explain the results (e.g., selective reporting of measures/items, optional stopping). In short, clarifying that the methodology behind a finding was rigorous should justifiably increase confidence in that finding. Although the benefits of disclosing unreported methodological decisions might seem personally threatening to some researchers, we believe that doing so is actually compatible with self-interest. The disclosure in no way jeopardizes the published status of a paper or its place on one's CV because it is not fair to blame authors for playing by the implicit rules of the game as they stood when the paper was accepted. Moreover, as discussed below, these rules are sometimes enforced by reviewers and editors.

A second benefit for researchers of adopting our disclosure norms involves promoting and establishing more sensible publication criteria. It will likely become clear that many published findings that were subsequently replicated and extended had less-than-perfect support for the initially reported findings, by today's rigid significance standards. To reach back across the decades, one paper that was apparently published under norms of full disclosure was Festinger and Carlsmith's (1959) classic induced-compliance experiment in support of cognitive dissonance theory. Under close examination, only one of its three relevant outcome measures yielded a statistically significant effect of condition, with the other two yielding "marginal levels" of statistical significance. Yet as a result of further replications and modifications of this paradigm, the reality of induced compliance effects now stands beyond question.

Benefits for the field. An important benefit of PsychDisclosure.org for the field is that it contributes to a partial opening of the black box of peer-review and sheds light on the Questionable Editorial Practices (QEPs) that have recently been identified (Levelt Committee, Noort Committee, & Drenth Committee, 2012) as factors that contribute to the replicability problem in psychology. For example, it is now known that action editors and reviewers often require authors to selectively report experimental conditions, measures, or samples in order to "streamline" the article or to increase the compellingness of the story. (And indeed, as will be seen, this possibility is borne out in several of the design specification statements made by participating authors.) Given that most researchers will of course adhere to such QEPs, it is absolutely critical that such information be publicly disclosed and for awareness regarding QEPs to increase so that reforms in editorial practices can be implemented.

A second benefit of our initiative for the field is to raise awareness regarding ways in which our journals'

current reporting standards may be ineffective. At present, these standards do not require authors to disclose information concerning the aforementioned categories of important methodological design specifications. By continually contacting a randomly sampled 50% of corresponding authors of recently published articles in new issues of *PS*, *JPSP*, *JEP:LMC*, and *JEP:G*, our initiative is raising awareness among the community of psychologists about our journals' ineffectual reporting standards. The ultimate goal is that this increased awareness among the community of psychologists motivates journal editors to change editorial policies, such that the four categories of methodological details disclosed on PsychDisclosure.org become a required component for manuscript submission. Indeed, there is already evidence that our initiative is having an effect in this regard.

Normative Demand for Mandatory Disclosure Statements

Several aspects of the scientific community's response and reaction to PsychDisclosure.org suggest a strong normative demand for mandatory disclosure statements for journals in psychology. First, the current raw response rate of 46%, which is higher than we anticipated before the initiative was launched, clearly suggests that a large proportion of researchers believe such disclosure information to be important. It is crucial to note also that our response rate should be interpreted as a lower bound estimate (to the response rate anticipated if such disclosure was mandatory), given the total effort participating authors had to expend digging up methodological design specifications for all studies reported in the relevant paper, in some cases involving studies run 4-5 years earlier.

As surprised as we were by the raw response rate, we were even more surprised by the positive and appreciative nature of the feedback we received from participating authors. A large number of respondents (over 35) explicitly stated that (a) they fully supported our initiative, (b) this kind of initiative was long overdue, and (c) they were extremely grateful and appreciative for the important service we are providing to the field. Furthermore, the few negative reactions to our initiative involved concerns that journals, rather than an independent group of researchers, should be soliciting such information. Such researchers supported our cause but disagreed with the means. Also, several researchers disclosed the requested information but expressed concern that they were now in an unfair theoretical position whereby theoretical opponents could use some of their disclosed methodological details against them. These researchers felt strongly that we should have contacted 100%, rather than a random 50%, of

corresponding authors -- compelling evidence that these researchers believe such information needs to be disclosed for *all* published articles in psychology.

Taken together, we argue that the high response rate and the nature of the feedback to our initiative demonstrate that the community of psychologists desires the implementation (in some shape or form) of mandatory disclosure statements. Our results are consistent with a survey of over 1,200 psychologists (sent in December 2011) who overwhelmingly supported changes to improve research practices in psychology in relation to Simmons et al.'s (2011) recommendations (Fuchs, Jenny, & Fiedler, 2012). The vast majority of psychologists responding to that survey agreed that researchers should report all assessed measures, tested experimental conditions, and their data collection termination rule. Fuchs et al. also found that a near-majority of the psychologists agreed that it should be a *publication condition* that authors disclose eliminated observations (52% agreement), all tested experimental conditions (66%), all assessed measures (46%), and the data collection termination rule (46%). We believe our response rate and positive feedback provide stronger evidence than Fuchs et al.'s findings for requiring such information as a publication condition, given that researchers had to actively reply to our email, find and collate the information, and agree to go on public record regarding the disclosure of the unreported methodological information. That is, the act of actually disclosing the information carries much more weight than simply *saying* that we should be doing this. After all, actions speak louder than words.

Concerns Regarding Mandatory Disclosure Statements

We now address various concerns that have been raised regarding requiring disclosure of the four categories of methodological design specifications as a mandatory requirement for publication in psychology journals.

It will lead to unfair evaluations. Because of the highly competitive nature of peer-reviewed publishing, esthetics of data perfection often seem to prevail to the detriment of an overall view of the evidence (Giner-Sorolla, 2012). One concern is that mandatory disclosure would lead to unfair evaluations of imperfect-looking findings, even if these are to be expected statistically when testing a true hypothesis; that the rigid standards applied now would be applied to more honest reporting, and that more papers would be rejected because of this. The obvious answer is for reviewers and editors to take a more comprehensive view of the evidence, and for authors to help this view along through meta-analytic or multivariate approaches showing an overall consistency of results

even if some, taken individually, are not statistically significant.

A related concern is that mandatory disclosure may hamper exploratory research. Researchers may fear that they will be penalized for including exploratory measures or manipulations. Here, we think an emphasis on disclosure will actually be fair, rather than unfair. It will be fair to authors who honestly distinguish between exploratory and confirmatory research stages in their writing, rather than pretending they knew it all along (e.g., Kerr, 1998). It will be fair to authors who take the time and effort to choose confirmed, well-tested methods over improvised, ad-hoc ones – including validated measures and manipulations, and high-powered sample sizes – because such methods increase the ability to interpret results, even null results. Combining these two standards, correctly applied disclosure standards will be fair to authors who follow-up exploratory work with principled validation. Authors must show that discarded methods have lost out because of independently verifiable issues with validity or reliability (as indicated in 6 of 73 design specification statements), rather than because of the circular reasoning that they happened to show no connection between predictor and outcome. For example, disclosing that two experimental conditions were dropped, due to independent evidence that most of their participants failed to understand task instructions, would be a principled reason for exclusion.

In a final fairness concern, studies involving expensive and time-consuming designs or hard-to-reach populations, such as longitudinal, developmental or cross-cultural research, often maximize the opportunity of data collection by testing many hypotheses at the same time. Would requiring authors to disclose all assessed measures lead to unfair evaluations of such designs – or, perhaps, to indigestible papers groaning with pages and pages of unrelated material? Again, the key to fairness is to avoid applying disclosure criteria mechanically, but to see them as an opportunity for principled reasoning. For one, theoretically unrelated measures can be described in broad terms, and even in footnotes or supplementary materials, as long as editors and reviewers are convinced from this description that they bear no relevance to the question at hand, and so do not present a validity problem for the conclusions of the paper. Given the importance of research context, such inclusions deserve to be evaluated before publication in any case, without necessarily interrupting the flow of the final article's argument. What would also be unfair (to science) is using the effortful nature of one's research to justify drawing conclusions that are supported by only one out of five

conceptually related measures included without mentioning the failures.

To strongly ensure that a focus on some measures/manipulations over others can be justified in any given write-up, researchers can honor and support particularly intensive studies by identifying focal, peripheral, exploratory and confirmatory research hypotheses a priori (which can be done via pre-registration; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). The effortful nature of a study's data collection makes it even more important to make sure to construct each hypothesis so that an effect in either direction -- or even a null effect (e.g., Mehl, Vazire, Ramirez-Esparza, Slatcher, & Pennebaker, 2007) -- is theoretically interesting. The latter goal can be achieved by using methodologically sound manipulations and measures (LeBel & Peters, 2011) and stronger forms of hypothesis testing (Fiedler, Kutzner, & Krueger, 2012; Meehl, 1967, 1978; Platt, 1964).

It will be ineffective, due to too many exceptions. Some are concerned that having rigid mandatory reporting requirements may not be effective given the wide variety of research designs and types of empirical investigations published in psychology journals. We agree that there will inevitably be challenges in implementing mandatory disclosure statements. For instance, standardized 4- or 5-category disclosure statements may not be completely relevant for all research designs. Also, standardized disclosure statements might not be able to rule out -- for all kinds of empirical investigations -- that questionable research practices were used to obtain the reported findings. We believe, however, that these situations will occur in the minority of cases (Simmons et al., 2012). More important, changes and improvement can (and inevitably will) be made over time to improve the mandatory disclosure statements so as to maximize their utility in promoting sounder research practices. As Simmons et al. pointed out by way of analogy, public health benefited immensely when U.S. politicians in 1906 implemented *imperfect* food regulations, such that all manufacturers were required to disclose the ingredients used in the production of all foods. By the same token, mandatory disclosure statements need not be perfect for real and substantial scientific benefits to be realized.

People will lie. There likely will be a non-zero rate of lying in disclosure statements, just as there is a non-zero rate of data fabrication and falsification. However, we believe that lying will be rare in practice for the simple reason that there is a huge psychological difference between a sin of omission, which is not reporting such information, and a sin of commission, which entails making actual false statements regarding the methods used to obtain the reported results

(Cushman, Young, & Hauser, 2006; DeScioli, Christner, & Kurzban, 2011). Indeed, a case can be made that making false statements in such disclosure statements would actually qualify as scientific fraud (American Psychological Association, 2010).

It will lengthen journal articles. Though this is technically true, making a disclosure statement and disclosing the recommended methodological information should take no more than a few sentences in the Methods section. For instance, Simmons et al.'s (2012) voluntary disclosure statement involves a brief 21-word sentence. And disclosing the additional methodological information (e.g., additional conditions, measures) should take no more than a few additional sentences because we are proposing only that the existence of additional measures and conditions be revealed, not that they be given a full statistical reporting. In the case where numerous additional measures were assessed, these can be described in online supplementary materials.

It is not feasible. A final and more general concern with having journals implement mandatory methodological disclosure statements is that it might impose overly burdensome infrastructure costs for journals. This is simply not the case given that journals could just request that authors include a version of Simmons et al.'s (2012) 21-word disclosure statement directly in their cover letters. Editors and reviewers could then confirm that all relevant methodological details were in fact disclosed in the manuscript and/or supplementary materials. As additional evidence supporting the feasibility of mandatory disclosure policies, we note that prominent journals in other areas have already implemented similar disclosure policies (e.g., the journals *Management Science* and *Marketing Science*; Desai, 2013).⁶

A Push for More Openness

Another possible implication of our PsychDisclosure.org results -- especially given the high response rate and positive feedback -- is that psychologists are ready for and desire scientific openness more broadly. This hunch is based on several aspects of the information voluntarily disclosed to us by corresponding authors. First, participating authors often disclosed *more* information than requested. For instance, many researchers provided very detailed methodological information even though we explicitly instructed them to provide *brief* design specification statements (with a few going as far as sending multiple-page documents as attachments with detailed information). Many authors also disclosed the

⁶ Notably, prominent journal *Nature* recently introduced editorial measures to improve the reliability of published findings by giving more space to methods sections, requiring that key methodological details are disclosed, and encouraging authors to make their raw data available (Campbell, 2013).

actual names of each of the additional measures not reported in the original article (our request was simply whether they had reported all assessed measures and, if not, to provide the reasons for not reporting them). Most compellingly, 32% of authors (i.e., 6 out of 19 design specification statements in the conditions category) actually disclosed that additional studies were executed but not reported in the published article. This is surprising given we did not actually request this information.⁷ Taken together, these observations suggest that many researchers desire broader disclosure of the research process in general.

We see this potential development as very exciting and wish to comment briefly on existing infrastructure at OpenScienceFramework.org (OSF) that will appeal to researchers interested in moving toward open science practices. The OSF is a website characterized by several important features that promote greater scientific openness, including pre-registration of studies prior to data collection and the archiving of study materials, data, and syntax files to facilitate internal and external direct replications and meta-analytic efforts (for more details see the Supplementary Materials). We strongly encourage psychologists to consider using the OSF to facilitate broader disclosure and transparency, which we believe will go a long way in improving the reliability of findings in psychology.

Conclusion

There is currently an unprecedented level of doubt regarding the reliability of research findings reported in psychology (Pashler & Wagenmakers, 2012) and neuroscience (Button et al., 2013) journals, which has further tarnished the credibility of our field (Carpenter, 2012; Kahneman, 2012) and threatens to induce further cuts to funding for psychological research. Our position is that the reliability of findings in psychology can be improved by requiring the disclosure of four important methodological design specifications, which are required for accurate interpretation and evaluation of reported research findings. Indeed, our investigation revealed that such important methodological details were often not reported, particularly the data collection termination rule and whether additional measures/items were assessed. We argue that the time is now to make the disclosure of such information mandatory at submission given the positive grassroots sentiment reflected in the high response rate to our requests (almost 50%) and appreciative reaction by many to our initiative, which suggest that psychologists want to see editorial policy changes made in this regard at the systemic level. Our initiative also revealed several instances of questionable editorial

practices, which need to be thoroughly examined and redressed. Open disclosure of the four targeted methodological categories will foster needed, action-oriented discussions regarding the soundness of various researcher and editorial practices. Ultimately – and ironically – it is our hope that this initiative, essential in the current climate, will eventually become unnecessary, being replaced by mandatory methods disclosure statements for all submissions to psychology journals.

Authors' Note

We would like to thank Brian Nosek and Jeffrey Spies for conceptual help with preliminary ideas that eventually led to the development of PsychDisclosure.org and Sampo Paunonen for help improving the disclosure questions and wording of the email sent to corresponding authors. Thanks to Paul Conway and Sampo Paunonen for valuable feedback on an earlier version of the manuscript. Thanks also to Christian Battista for technical help with Python coding and web development and Amanda Abado for administrative help corresponding with authors. This research was supported by a Social Science and Humanities Research Council (SSHRC) post-doctoral fellowship to the first author.

References

- American Psychological Association. (2010). *American Psychological Association ethical principles of psychologists and code of conduct*. Retrieved February 28, 2013, from <http://www.apa.org/ethics/code/>
- Armitage, P., McPherson, C. K., & Rowe, B. C. (1969). Repeated Significance Tests on Accumulating Data. *Journal of the Royal Statistical Society. Series A (General)*, 132, 235-244.
- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A., Fiedler, K. et al. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, 27, 108-119.
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14, 365-376.
- Campbell, P. (2013, April 24). Reducing our irreproducibility [Editorial]. *Nature*, 496, 398.
- Carpenter, S. (2012). Psychology's bold initiative. *Science*, 335, 1558-1561.
- Cushman, F. A., Young, L., & Hauser, M. D. (2006). The Role of Reasoning and Intuition in Moral Judgments: Testing three principles of harm. *Psychological Science*, 17, 1082-1089.

⁷ We considered requesting this information, but decided against it thinking the task would be too onerous.

- Desai, P. S. (2013). Marketing Science replication and disclosure policy. *Marketing Science*, 32, 1-3.
- DeScioli, P., Christner, J., & Kurzban, R. (2011). The omission strategy. *Psychological Science*, 22, 442-446.
- Feller, W (1940). Statistical aspects of ESP. *Journal of Parapsychology*, 4, 271-298.
- Festinger, L., & Carlsmith, J.M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203-210.
- Fiedler, K., Kutzner, F., & Krueger, J. (2012) The Long Way From Alpha-Control to Validity Proper: Problems with a Shortsighted False-Positive Debate. *Perspectives on Psychological Science*, 7, 661-669.
- Fuchs H., Jenny, M., & Fiedler, S. (2012). Psychologists are open to change, yet wary of rules. *Perspectives on Psychological Science*, 7, 634-637.
- Giner-Sorolla, R. (2012). Science or art? How esthetic standards grease the way through the publication bottleneck but undermine science. *Perspectives on Psychological Science*, 7, 562-571.
- Ioannidis, J. P. A. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, 7, 645-654.
- John, L., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth-telling. *Psychological Science*, 23, 524-532.
- Kahneman, D. (2012). *A proposal to deal with questions about priming effects*. Available at: http://www.nature.com/polopoly_fs/7.6716.1349271308!suppinfoFile/Kahneman%20Letter.pdf
- LeBel, E. P., & Paunonen, S. V. (2011). Sexy but often unreliable: Impact of unreliability on the replicability of experimental findings involving implicit measures. *Personality and Social Psychology Bulletin*, 37, 570-583.
- LeBel, E. P., & Peters, K. R. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in modal research practice. *Review of General Psychology*, 15, 371-379.
- Levelt Committee, Noort Committee, & Drenth Committee (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*. Retrieved 14 February 2013 (<http://www.tilburguniversity.edu/nl/nieuws-en-agenda/finalreportLevelt.pdf>).
- Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, 7, 531-536.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 3642.
- Makel, M. C., Plucker, J. A., & Hagerty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science*, 7, 537-542.
- Meehl, P. E. (1967). Theory testing in psychology and physics: A methodological paradox. *Philosophy of Science*, 34, 103-115.
- 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, 806-834.
- Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are Women Really More Talkative Than Men? *Science*, 317, 82.
- Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, 7, 615-631.
- Open Science Collaboration (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science*, 7, 657-660.
- Pashler, H., & Wagenmakers, E.-J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, 7, 528-530.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359-1366.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 Word Solution (October 14, 2012). Available at SSRN: <http://ssrn.com/abstract=2160588>.
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J. & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, 7, 627-633.