

# Registered Reports and Replications in Attention, Perception, & Psychophysics

Published online: 21 June 2013  
© Psychonomic Society, Inc. 2013

*Attention, Perception, & Psychophysics* is launching a new type of research report: Registered Reports and Replications (RRR). Here is why we are trying this experiment. APP is always interested in strengthening the reliability and validity of the results in our science. However, there has been a recent rise in concern about several pitfalls on that road. These include the following:

1. *Publication bias*: Studies that yield statistically “significant” results get published, while those that don’t stay in the file drawer. The file drawer problem makes it difficult to evaluate how replicable a finding might be. If 20 people independently run the same study, 1 will get published; in the most dire case, 1 person might run 20 experiments and publish the one that beats  $p < .05$ . This can lead to a proliferation of nonreplicable effects in the literature.
2. *P-hacking* (Simmons, Nelson, & Simonsohn, 2011): This covers a range of dangerous practices designed to get the  $p$ -value under the magic  $p < .05$  line. This includes the common practice of adding some extra observers to a study that looks promising ( $p < .06$ ). This sounds fairly benign, but consider the analogy to a coin flip game: If the coin comes up “heads,” I win. If it comes up “tails,” we flip again. Do you want to play this game?
3. *Underreporting of methods and variables*: While failure to report variables might not be a vast problem in most of what gets reported in APP, if you test 20 variables and report the 1 that is significant ( $p < .05$ ) . . . again, we have a problem, especially if you don’t mention the other 19, nonsignificant variables.
4. *Multiple comparisons in statistical testing*: There are long, complicated, and controversial arguments to be made here. Nevertheless, we can probably all agree that with enough massaging of the data—dropping a “bad” subject here, adding a new post hoc analysis there, and doing a few dozen pairwise comparisons—eventually, that  $p$ -value that you are reporting is going to be a bit meaningless.
5. *Underpowered studies*: There have been some interesting recent articles on the perils of small sample sizes and low statistical power (e.g., Button et al., 2013). This is more of a problem for experiments where you might compare two groups of observers and you might have just one data point per observer and is less of a problem when you are collecting vast numbers of trials in a within-subjects design. However, given that we are often dealing with small effect sizes (a few tens of milliseconds here, a few percentage points in accuracy there), it would be foolish for us to be complacent on this issue. Underpowered studies tend to generate both spurious (and unreplicable; see below) positive results and invalid negative results. Admittedly, it is unclear how to do power calculations for many of our standard designs. Nevertheless, we should make every effort to ensure adequate statistical power.
6. *Replication*: We just don’t replicate each other enough. We may not even replicate ourselves enough. We would avoid many of the perils of  $p$ -hacking (for example) if, when we finally got that experiment to work after 20 tries, we turned around and did a clean replication of the study.

APP recognizes that we cannot solve all of these problems with this RRR proposal, but we hope that the addition of this format to our portfolio of articles will be useful. Here is the idea.

1. You tell us what you are going to do and why. Basically, you would be writing the background and method sections of a proposed paper. You would be very specific about your hypotheses and methods (details below).
2. We would send this half of a paper out for review. We would ask reviewers the following:
  - a. Is this experiment going to be answering a question worth answering in the pages of APP?
  - i. For a registered report: Does a significant theoretical issue hang on the outcome?

- ii. For a replication: Is there reason to question whether the original result replicates?
  - iii. For a replication: Is the original finding of sufficient interest to be worth this kind of formal replication?
  - b. Is the proposed study methodologically sound and adequately powered?
  - c. Are new experimental manipulations, beyond a pure replication, proposed? Reviewers will be encouraged to propose adjustments in the design, parameters, and so forth. Potential authors can, in turn, dispute the reviewers' suggestions.
3. Critically, if the proposed paper is approved (perhaps after a round of revision), APP would commit to publish the results of the experiment regardless of the outcome or the statistical significance of the results. Again, revision might be required. We don't promise to publish horrible meandering discussion sections just because the experiment was done well. Once they see the final version, reviewers might have questions about the meaning of the results that might require some changes in the discussion, but we would have all agreed in advance that the results of this experiment should be published in APP.

The RRR format would be appropriate for replications of important results from other labs. You might, for example, propose a straight replication plus a manipulation. Alternatively, the format would be appropriate for a study designed to distinguish between two hypotheses or models in the literature. It is less likely to be appropriate for experiments breaking into entirely new territory. Our standard Article and Short Report formats remain in place and will be the right way to report on many lines of research. We hope that, by adding RRR to the mix, we will provide the community with a new and useful tool. We will see how well it works, if and when you use the tool.

Our RRR initiative is inspired, in part, by a similar initiative at *Cortex* (Chambers, 2013). The detailed instructions for RRR submissions in APP are found below and will be posted on the journal Web site. Submission of registered reports and replications will be through the usual Web-based process (<http://mc.manuscriptcentral.com/pandp>).

Please send questions and comments to Jeremy Wolfe at <[jwolfe@partners.org](mailto:jwolfe@partners.org)>.

### Instructions for submitting a registered report or replication to *Attention, Perception, & Psychophysics*

This format is intended to strengthen the reliability and validity of the results in our field of science. The goal is to increase the visibility of valuable replications (Roediger, 2012; Wagenmakers, Borsboom, van der Mass, & Klivet, 2012) and decrease the likelihood of publication bias (Fanelli, 2010; Francis, 2012; Nosek, Spies, & Motyl, 2012; Rosenthal,

1979). Authors will submit manuscripts in two stages. An initial submission is a proposal of one or more experiments that have not yet been conducted. These will be reviewed with emphasis on the value of the work to be done. After the manuscript is approved and once the experiments have been completed, the revised manuscript is submitted for fast-track publication and will be reviewed to verify that the work was completed appropriately. All submissions for a registered report are required to contain information for increasing transparency and credibility described by the checklist below (based on Fuchs, Jenny, & Fielder, 2012; Simmons, Nelson, & Simonsohn, 2011). Registered reports are limited to 3,000 words of main body text plus figures, although exceptions are possible if approved by the editors; supplementary material is encouraged. The cover letter to the editors should explain why this is appropriate as a registered report.

#### Part one: checklist for authors submitting a manuscript for approval as a registered report or replication

1. Does this manuscript describe the theory under investigation and the specific hypotheses that lead to the procedures proposed?

*The manuscript needs to justify the significance of the proposed registered report or replication. Reports of these types should be using established methodologies and should be testing existing theories (Nosek, Spies, & Motyl, 2012). This is not the place for methodological and/or theoretical innovations. Our standard Article and Short Report formats serve those roles. The RRR format is a mechanism for confirming or disconfirming prominent theories and findings in the field.*

2. Does the manuscript report the previous, related experiments, published or unpublished, conducted by the same or other researchers?

*If this work is an extension of other experiments conducted by the same researchers, authors are required to briefly describe these previous studies and their outcomes in the submitted manuscript. In particular, if the proposed experiment is the product of a process of exploratory investigations, a brief account should be given. In the case of a replication proposal, this would include unpublished failures to replicate the target study.*

3. Does the manuscript specify all the variables, both independent and dependent, in the experiment? Are all conditions to be tested clearly described?

*Detailed descriptions of the methods are crucial to appropriate evaluation. Even if some variables are not important for the agenda of the manuscript, they need to be described for review purposes.*

4. Does the manuscript address the issue of statistical power?

*In cases where standard power analysis is possible, authors should describe procedures that will achieve the*

power of .90 or higher. *A priori power analysis at this level is recommended to improve the reliability of results in the field (Tressoldi, 2012). All studies should give a clear justification for the combination of number of observers/participants and number of observations.*

5. Does the manuscript specify a clear rule for terminating data collection?

*In most cases, this will be determined by the number of participants that are needed to reach the required power level. Interim analyses, if planned, should also be described.*

6. Does the manuscript specify the data analysis procedures that will be used?

*Rules for data elimination, such as participant exclusion criteria and outlier trimming, must be prespecified. In the final report, it is possible to include additional post hoc analyses. In the RRR format, the reader must be able to clearly distinguish these from the registered analyses.*

7. Does the manuscript specify a plan for making the raw data publically available?

*This can be as simple as specifying the Web address to which the data will be posted.*

*[NOTE: Even though the initial submission will be referring to a future study, please use the past tense for all sections of the text, despite the procedures having not yet been conducted.]*

8. Does the project have ethics approval and all other necessary approvals, and is funding in place to start the research immediately?

Part two: checklist for authors submitting an approved registered report for publication

1. Does the manuscript describe the completed experiment(s) in a manner that complies with the approved report (methods and analyses)?
2. Do the authors certify that the data for the registered experiment were collected after receiving approval from AP&P?
3. Does the manuscript note any unforeseen changes in the approved methods and analyses? Sometimes there are procedural errors, data-coding errors, participant recruitment problems, and so forth. These should be acknowledged in the manuscript and flagged in the cover letter. This is not the time or place to report substantive changes in the conditions of the experiment. Such changes take the paper outside of the realm of a registered report (although, perhaps, a fine standard research paper).
4. Does the manuscript describe and justify all post hoc analyses?

5. Do the conclusions follow from the results?

Since this is a new format, there may be issues that we have not anticipated. We will work with authors to resolve such questions as they arise. Please consult the editors as needed.

## Deadlines

Once a registered report or replication is approved, the authors have 1 year to submit the actual manuscript with the results. That deadline can be extended by negotiation with the editor, but in general, after a year, the project requires new approval.

**Acknowledgments** I thank Adriane Seiffert for drafting the first version of this policy and the journal's associate editors for useful commentary.

## References

- 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.
- Chambers, C. D. (2013). Registered reports: a new publishing initiative at *Cortex*. *Cortex*, *49*, 609–610. <http://www.ncbi.nlm.nih.gov/pubmed/23347556>
- Fanelli, D. (2010). “Positive” results increase down the hierarchy of the sciences. *PLoS One*, *5*, e10068.
- Francis, G. (2012). Publication bias and the failure of replication in experimental psychology. *Psychonomic Bulletin & Review*, *19*, 975–991.
- Fuchs, H. M., Jenny, M., & Fielder, S. (2012). Psychologists are open to change, yet wary of rules. *Perspectives on Psychological Science*, *7*, 639–642.
- 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.
- Roediger, H. L., III (2012). Psychology's woes and a partial cure: The value of replication. *APS Observer*, *25*(2), 9, 27–29.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.
- 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.
- Tressoldi, P. E. (2012). Replication unreliability in psychology: Elusive phenomena or “elusive” statistical power? *Frontiers in Psychology*, *3*, 218.
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Mass, H. J. L., & Klievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, *7*, 632–638.