

EDITORIAL

Science revolves around the data

Jelte M. Wicherts*

We as psychological researchers have no problems with sharing our ideas, criticisms, and empirical findings with our peers and the wider community, yet we seem surprisingly reluctant to share the raw data that underlie our scientific enterprise. The *Journal of Open Psychology Data* was established to change the “closed research culture” in psychology in which around 73% of corresponding authors fail to act upon a signed statement that they would share upon request data from their published papers¹, in which fraudsters like Diederik Stapel could go on for years without sharing their (fabricated) data with coauthors and peers (Wicherts et al. requested data from Diederik Stapel in the summer of 2005 but like many others he indicated he lacked the time to share the data)^{2,3,4}, in which a higher prevalence of statistical errors is associated with unwillingness to share data⁵, in which it has become evident that data analyses are prone to human errors and a great deal of bias^{6,7,8,9}, and in which replications of previous findings are often hard to publish¹⁰.

Data are often much more interesting than the dense summaries we read in research papers. Data can be submitted to secondary analyses that can be useful and theoretically relevant. For instance, differences in variance between conditions in a randomized experiment may reflect heterogeneity of an effect.

Moderation of effects due to demographic variables (age, sex) may only come out if we collate (raw) data from multiple experiments. Correlations between variables¹¹ (widely ignored in the experimental paradigm) may shed new light on individual differences. Such correlations are often required for meta-analyses, for instance to compute standard errors in within-subject designs or to summarize effect sizes from multiple dependent variables. Novel methods of analysis, theories, and empirical results may lead us to revisit older data. Newly developed psychometric models may shed light on psychological measurement and the nature of individual differences or of experimental inductions. Secondary analyses may shed new light on findings and re-analyses of data enable verification of statistical results and conclusions. And researchers may simply disagree on how to best analyze a given dataset, which should become part and parcel of scientific debates. For instance, when a field is confronted with diverging results¹², it is worthwhile to have the data of original studies and their replications available for further scrutiny and debate.

Sharing data in psychology is uncommon¹ and a survey conducted by the Data Archiving Network Services (DANS) among over 200 psychologists in The Netherlands¹³ highlighted a poor practice of archiving data. Results showed that many psychological researchers appear to think that saving a haphazardly documented data file on one's current computer amounts to archiving the data for posterity. Everyday many valu-

* Editor-in-Chief.
Department of Methodology and Statistics,
Tilburg School of Behavioral Sciences,
Tilburg University, The Netherlands.

able psychological data sets get lost simply because researchers move offices, replace their old computer, hire a new research assistant, update a statistical software package, or lose track of the data for other reasons. An explicit promise to share information or data upon request often does not work either; in a study in a related field only 44% of authors were able to share supplementary information as promised in their recent article¹⁴. Moreover, researchers may tend to think that they lose their competitive advantage if they share data that could be submitted to secondary analyses in follow-up work. But quite often researchers simply lack the time or expertise to run secondary analyses on their data set. A good option is to just publish it.

Researchers in psychology are often insufficiently aware of the values of sharing their data¹⁵. Sharing data is associated with higher impact, in the sense that papers from which data were shared garner relatively more citations^{16,17}. The *Journal of Open Psychology Data* is meant to further reward sharing of data with the publication of a paper in a peer-reviewed journal in which authors describe the data they have submitted to a repository. The journal therefore becomes a place to share interesting psychological datasets and to find useful data for novel research or educational purposes. Publishing data that are useful and interesting for future research means contributing to the literature in a novel way. Collectors of the data always have the most intimate knowledge of the data and so they are the first in line for any potential collaboration. More importantly, publishing one's data means behaving in accordance with the scientific norm of communality, to which the preponderance of scientists subscribe¹⁸. Publishing data represents a 21st century view on publishing scientific results¹⁹ in which we need no longer worry about the outdated notion of journal space that has so long restricted the amount of information we shared with

the scientific community when reporting empirical results.

The *Journal of Open Psychology Data* publishes data papers concerning data from research that has been reported elsewhere (typically in a substantive journal) and data from relevant research that has not been previously published, including replication attempts of previous results. The goals of the *Journal of Open Psychology Data* are (1) to encourage a culture shift within psychology towards sharing of research data for verification and secondary analyses, (2) to reward the sharing of data via repositories by providing full article-level metrics and citation tracking, (3) to offer peer-review of the quality and re-use potential of data sets and the documentation thereof, (4) to enable rapid open-access publishing at a low cost (currently only €30), (5) to offer an online forum for discussion, reanalysis and verification of data, and (6) to facilitate publication of data from replication research. An increasing number of grant-giving organizations stipulate that data financed by public money should eventually be made available to the scientific community and JOPD offers a means of doing so in a manner that is subject to rigorous peer-review.

This is how publishing in JOPD works. After having collected data that are potentially interesting or after having published a paper on the dataset (in a substantive journal), the author submits the data to one of several high-standard data repositories listed on the JOPD website and subsequently writes a JOPD paper on the basis of the template. Papers are relatively short and include information on the origin and whereabouts of the data (with robust links), a description of the sample, variables in the dataset, methods, procedures, and measurements. If data were used in a previously published paper, authors should include specific references to that earlier work while keeping an eye on readability (e.g., by providing a glossary of procedures instead of stating that all procedures are described elsewhere). Also included are discussions of issues

related to research ethics and privacy, potential drawbacks, and the reuse potential. After a quick perusal by the editor of the suitability of the work presented (including a check of data access), the manuscript is sent out to two reviewers with expertise in the substantive area of interest, who assess it on (1) quality of the description in the data paper, (2) accessibility of the underlying data and completeness of documentation and meta-data in the repository, and (3) the reuse potential of the data (for research and education) or its value for replication research (when data concerns a replication). Decisions are made to accept, revise, or reject the manuscript. For reasons of transparency²⁰, in Summer 2013 we will begin to publish reviewer's reports alongside published papers, which we encourage reviewers to sign (although anonymity is allowed). In addition, the website offers possibilities to comment on published papers by the community. Also, we solicit submissions of data papers that concern replications, which have hitherto been notoriously hard to publish (especially in cases of "failed replications") and may play a key role in understanding of when effects do or do not occur. So we welcome submissions that involve data from both published and unpublished works. The core criterion is whether the data have the potential to be used in future work, which includes alternative analyses, novel types of analyses, and meta-analyses. Moreover, data may also be useful for educational purposes. For instance, data from published papers can be used in assignments in which students replicate the reported statistical analyses.

To date, major publishers and professional organizations have done little to change the current culture of secrecy concerning data in psychology²¹. But sometimes all we need is a good place to open up. I hope that JOPD will motivate researchers to share their data and help end the culture of secrecy that is so unbecoming of science.

Acknowledgements

Acknowledgments are due to Peter Doorn from DANS for his generous support in setting up the JOPD, the editorial board members, and Brian Hole from Ubiquity Press.

References

- Wicherts J M, Borsboom D, Kats K and Molenaar D** 2006 The poor availability of psychological research data for reanalysis. *American Psychologist*, 61: 726–728. DOI: <http://dx.doi.org/10.1037/0003-066X.61.7.726>
- Crocker J and Cooper M L** 2011 Addressing scientific fraud. *Science*, 334(6060): 1182–1182. DOI: <http://dx.doi.org/10.1126/science.1216775>
- Smithsohn U** 2013 Just post it: The lesson from two cases of fabricated data detected by statistics alone. *Psychological Science*, in press.
- Wicherts J M** 2011 Psychology must learn a lesson from fraud case. *Nature*, 480(7) DOI: <http://dx.doi.org/10.1038/480007a>
- Wicherts J M, Bakker M and Molenaar D** 2011 Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One*, 6(11)e26828 DOI: <http://dx.doi.org/10.1371/journal.pone.0026828>
- Bakker M, van Dijk A and Wicherts J M** 2012 The rules of the game called psychological science. *Perspectives on Psychological Science*, 7(6): 543–544. DOI: <http://dx.doi.org/10.1177/1745691612459060>
- Bakker M and Wicherts J M** 2011 The (mis) reporting of statistical results in psychology journals. *Behavior Research Methods*, 43(3): 666–678. DOI: <http://dx.doi.org/10.3758/s13428-011-0089-5>
- Simmons J P, Nelson L D and 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: <http://dx.doi.org/10.1177/0956797611417632>

- Wagenmakers E J, Wetzels R, Borsboom D and van der Maas H L J** 2011 Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem. *Journal of Personality and Social Psychology*, 100(3): 426–432.
- Asendorpf J B, Conner M, Fruyt F D, Houwer J D, Denissen J J A, Fiedler K, Fiedler S, Funder D C, Kliegl R, Nosek B A, Perugini M, Roberts B W, Schmitt M, van Aken M A G, Weber H and Wicherts J M** 2013 Recommendations for increasing replicability in psychology. *European Journal of Personality*, in press.
- Cronbach L J** 1957 The two disciplines of scientific psychology. *American Psychologist*, 12(11): 671–684.
- Shanks D R, Newell B R, Lee E H, Balakrishnan D, Eklund L, Cenac Z, Kavvadia F and Moore C** 2013 Priming Intelligent Behavior: An Elusive Phenomenon. *PLoS One*, 8(4)e56515 DOI: <http://dx.doi.org/10.1371/journal.pone.0056515>
- Vorbrood C** 2010 Archivering, beschikbaarstelling en hergebruik van onderzoeksdata in de psychologie (Archiving, sharing, and reusing of psychological research data). The Hague, The Netherlands: Data Archiving and Networked Services (DANS);
- Krawczyk M and Reuben E** 2012 (Un)Available upon Request: Field Experiment on Researchers' Willingness to Share Supplementary Materials. *Accountability in Research: Policies and Quality Assurance*, 19(3): 175–186. DOI: <http://dx.doi.org/10.1080/08989621.2012.678688>
- Wicherts J M and Bakker M** 2012 Publish (your data) or (let the data) perish! Why not publish your data too? *Intelligence*, 40(2): 73–76. DOI: <http://dx.doi.org/10.1016/j.intell.2012.01.004>
- Piwowar H A, Day R S and Fridsma D B** 2007 Sharing detailed research data is associated with increased citation rate. *PLoS ONE*, 2(3)e308 DOI: <http://dx.doi.org/10.1371/journal.pone.0000308>
- Piwowar H A and Vision T J** 2013 Data reuse and the open data citation advantage. *PeerJ PrePrints*, 1e1v1 DOI: <http://dx.doi.org/10.7287/peerj.preprints.1v1>
- Anderson M S, Martinson B C and De Vries R** 2007 Normative dissonance in science: Results from a national survey of US scientists. *Journal of Empirical Research on Human Research Ethics*, 2(4): 3–14.
- Nosek B A, Spies J and Motyl M** 2012 Scientific Utopia: II - Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, 7(6): 615–631. DOI: <http://dx.doi.org/10.1177/1745691612459058>
- Wicherts J M, Kievit R A, Bakker M and Borsboom D** 2012 Letting the daylight in: reviewing the reviewers and other ways to maximize transparency in science. *Frontiers in Computational Neuroscience*, 6: 20–20. DOI: <http://dx.doi.org/10.3389/fncom.2012.00020>
- Wicherts J M and Bakker M** 2009 Sharing: guidelines go one step forwards, two steps back. *Nature*, 461: 1053–1053. DOI: <http://dx.doi.org/10.1038/4611053c>

How to cite this article: Wicherts, J M 2013 Science revolves around the data. *Journal of Open Psychology Data*, 1(1)e1, pp. 1-4, DOI: <http://dx.doi.org/10.5334/jopd.e1>

Published: 7 June 2013

Copyright: © 2013 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License (CC-BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/3.0/>.

ju [The *Journal of Open Psychology Data* is a peer-reviewed open access journal published by Ubiquity Press

OPEN ACCESS 