


Outstanding Data sets: A new category of articles that promotes modelling published in the Quantitative Methods for Psychology



Denis Cousineau ^a, Sebastien Helie ^b

^a École de psychologie, Université d'Ottawa

^b Department of Psychological Sciences, Purdue University

Abstract ■ Model comparison requires that the data set be fixed. Yet, it is common practice to replicate studies when a researcher wishes to test a new model. Consequences of this practice are (a) reduced comparability between model fits, and (b) unnecessary time and money invested in experimentation. The Quantitative Methods for Psychology journal opens a new category of articles, the Outstanding data sets, whose purpose is to publish research paper as well as research data. With this new category of paper, modeling cycle will take less time because it won't be necessary to rerun an experiment just to have a data set.

Keywords ■ Editorial, category of articles, outstanding data set.

 denis.cousineau@uottawa.ca ( <http://quibb.ca>)

Introduction

A fair amount of research in experimental psychology involves modeling. Modeling can be under the guise of statistical analyses, or using a model of psychological processes and mental events. This journal has published tutorials on many statistical models (e.g., Desjardins, 2005, Giguère, 2006, Chartier and Faulkner, 2008, Doey and Kurta, 2011). It has also published articles related to psychological models (e.g., Newell, Mayer-Kress and Liu, 2006, Rosenbloom, 2006, Harvey, 2011).

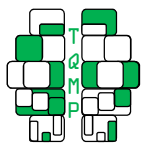
Model comparison is an approach by which two or more models can be compared to see which one provides the best account of the results. Tools for model comparisons are increasingly well-understood (see Hélie, 2006, Burnham & Anderson, 2004). When the models are nested, i.e., when one model is a restricted case of a second model, model comparison is straightforward: the decrement in fit for the restricted model follows a chi-square distribution with the degrees of freedom given by the number of constrained parameters (Pitt & Myung, 2002). When the models are not nested, they can still be compared but if the numbers of free parameters are not equal, the Aikake information criterion or the Bayesian information criterion must adjust fits to penalize more complex models. Model complexity can alternatively be assessed using Minimum Descriptive Length (Grünwald, 2000, Glover & Dixon, 2004). The problem is that all these

approaches assume that the same data are used to fit the different models. Unfortunately, this is rarely the case.

In a typical research proposing a new model, the researcher must have data, and therefore, a new experiment is run. The dataset being different, new fits of these data are not directly comparable with the fits previously reported in the literature. In addition, it is very rare that a researcher will run a perfect replication of the original experiment. The variations may not be minor; the researcher may even be unaware of the variations if the methodology in the original article is not sufficiently detailed.

For these reasons, there is a lack of compatibility between fits from existing literature and new fits based on new data sets. One of the most important premises of science is that we should change only one factor at a time. Yet, in the way we currently test models, we change two factors simultaneously: the model and the dataset.

The most common solution is to fit the original model on the new data set. Although it is quite a prerequisite to be able to fit a competitor's model before proposing alternatives, this solution is perilous: the researcher may be unable to reproduce fits similar to those reported in the original article. In this case, he or she will never be able to publish the newer or revised model. This is unfortunate because the inability



to reproduce a fit may not be his or her fault but may be caused by confounds or confusions in the original article. With the culture of publishing only positive results in psychology (Yong, 2012), a paper correcting the issue(s) of the original article may never be published (Ioannidis, 2012).

A second solution occasionally seen is to share a data set (e.g., Reynolds, Miller, 2009, who reanalyzed data from Cousineau and Shiffrin, 2004). However, this practice is quit uncommon. It is dependent on the author(s) of the original data results and they can withdraw their approval at any moment in the process. Also, the data may no longer be available if more than five years have elapsed since the original publication.

A third solution is to have a data repository. With this solution, replicating experiment is no longer required and if a fit cannot be reproduced exactly, the only explanation is that the original model is insufficiently specified.

One existing data repository was done under the initiative of A. Heathcote and the *Newcastle Cognition Laboratory* (<http://www.newcl.org/?q=node/4>). However, this repository never extended beyond this lab. J. Wolfe also opened a data repository specialized in visual search experiments, inviting others to contribute (Wolfe, 1998), but this site is now offline. Another similar endeavor is Russ Poldrack's Open fMRI Project (<http://openfmri.org/>), which stores and distributes fMRI data sets. However, this project is still in its infancy and relatively unknown and very few cognitive neuroscientists are using it. Other disciplines, such as machine learning, have been more successful. For instance, the University of California Irvine manages the UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/>), and fitting data from this repository is becoming a standard benchmark when new machine learning models are proposed.

This last solution is probably the best one, and is congruent with ethical guidelines which require that data be available for inspection five years after publication. In addition, the APA Certification of Compliance with APA Ethical Principles states clearly that "After research results are published, psychologists do not withhold the data on which their conclusions are based from other competent professionals who seek to verify the substantive claims through reanalysis..." (American Psychological Association, 2001, p. 396). In addition to the ability to replicate analyses and model fits, Wicherts, Borsboom, Kats, & Molenaar (2006) report five other arguments in favor of data repository.

A possible reason why these data repository initiatives were not successful is that there is not much incentive for researchers to do that. Our career is linked to publications, not to contributions to data repository. This is quite unfortunate because it implies that model fits across publications are not comparable and also that developing models implies devoting time and efforts in replicating studies. Although it is advisable that every psychology researcher be able to run experiments, there is in the present case unnecessary duplication of efforts.

To address these concerns, the *Quantitative Methods for Psychology* introduces a new category of article, the *Outstanding data sets* articles. The major purpose of these articles is to promote modeling, but using a common data set from experimental psychology. Using the data sets provided by these articles, we hope to promote modeling, inviting replies presenting new alternatives, or better modeling tools.

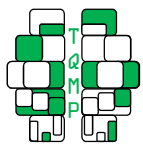
The *Outstanding data sets* present a data set, briefly put in context, and highlighting the agreed-upon effects and the problematic effects. To qualify for publication, the data set must be from a standard paradigm, manipulating accepted factors known to affect the results. It must also be exhaustive, covering a wide range of levels for the factors, presumably exploring them factorially whenever possible. Finally, the number of participants must be large enough, so that statistical power is demonstrably high.

The *Outstanding data sets* should also ideally present a first model. This model can be seen as a baseline model. The model must be well-explained so that its fits are replicable by the reader. If special tools are required, these must be explained or provided along with the article. If parameter search is done, search algorithm, settings and starting values must be provided explicitly. Ideally, the source code might be provided, although this is not a requisite.

By model, we specifically intend models referring to psychological constructs and enhancing the understanding of psychological phenomena. This excludes purely descriptive models such as general linear models used for statistical inference and null hypothesis testing.

The typical *Outstanding data sets* article contains those sections:

A brief theoretical context. The theoretical context can be short since, presumably, the paradigm is known and used by many researchers. It focuses on the typical findings, the factors that contribute most to changes in



performance. It also highlights the models currently put forth to explain such results, if there exists such models.

A summary of the methodology. Again, this section is brief, and can refer to other papers for the relevant details. This section explains the variety of factors explored and the reasons for their inclusion in the present experiment.

A summary of the results. The focus of this section is to confirm that typical results are replicated both in significance and in magnitude. Regarding non-significant results, a brief indication of the statistical power should be presented. Prior to this, outliers, if any, should be mentioned, as well as the criteria for their exclusion.

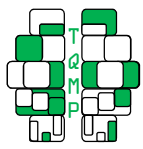
The model. If the author(s) chose to provide a baseline model, it is presented here (or a summary if it was presented in another paper). The parameters are listed explicitly in the text as well as in a summary table, and the indices of fit are given.

Description of the data file(s). Finally, the data file or files are described. They must be preferably in plain text or else, in a non-proprietary format. Each column must be named and if there are relevant information regarding the columns, they must be specified (e.g. type of scale, signification of the numbers if coded using integers, etc.). The data file must contain the raw data, not summary data.

The *Outstanding data sets* articles are invitation to provide new models, or better modeling practice. The *Quantitative Methods for Psychology* journal will consider replies as long as they describe a new model which has psychological validity and bring some new and interesting knowledge. It could be under the form of a better-fitting model (improved fit, taking into account complexity issues), comparable fit but from a totally distinct perspective, opening new venues of inquiries, contributing better practice or better techniques (for example, faster) for models of this type, or providing a more complete explanation. If accepted, they will be published with a subtitle "A follow-up on " and the name of the authors and the year of publication of the original paper.

References

- American Psychological Association. (2001). Ethical standards for the reporting and publishing of scientific information. *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association. pp. 387-396.
- Burnham, K. P., & Anderson, D. R. (2004) Multimodel inference, *Sociological Methods & Research*, 33, 261-304.
- Chartier, S. & Faulkner, A. (2008) General Linear Models: An Integrated Approach to Statistics. *Tutorials in Quantitative Methods for Psychology*, 4, 65-78.
- Cousineau, D. & Shiffrin, R. M. (2004) Termination of a visual search with large display size effect. *Spatial Vision*, 17, 327-352.
- Desjardins, J. (2005) L'analyse de régression logistique. *Tutorials in Quantitative Methods for Psychology*, 1, 35-41.
- Doey, L. & Kurta, J. (2011) Correspondence Analysis applied to psychological research. *Tutorials in Quantitative Methods for Psychology*, 7, 5-14.
- Giguère, G. (2006) Collecting and analyzing data in multidimensional scaling experiments: A guide for psychologists using SPSS. *Tutorials in Quantitative Methods for Psychology*, 2, 26-38.
- Glover, S., Dixon, P. (2004) Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review*, 11, 791-806.
- Grünwald, P. (2000) Model Selection based on Minimum Description Length. *Journal of Mathematical Psychology*, 44, 133-152.
- Harvey, L. (2011) Hidden Markov models and learning in authentic situations. *Tutorials in Quantitative Methods for Psychology*, 7, 32-41.
- Hélie, S. (2006) An Introduction to Model Selection: Tools and Algorithms. *Tutorials in Quantitative Methods for Psychology*, 2, 1-10.
- Newell, K. M., Mayer-Kress G., & Liu, Y-T. (2006) *Tutorials in Quantitative Methods for Psychology*, 2, 66-76.
- Pitt, M. A., & Myung, I. J. (2002) When a good fit can be bad, *Trends in Cognitive Sciences*, 6, 421-425.
- Reynolds, A., & Miller, J. (2009) Display size effects in visual search: Analyses of reaction time distributions as mixtures. *Quarterly Journal of Experimental Psychology*, 62, 988-1009.
- Rosenbloom, P. S. (2006) A cognitive odyssey: From the power law of practice to a general learning mechanism and beyond. *Tutorials in Quantitative Methods for Psychology*, 2, 43-51.
- Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006) The poor availability of psychological research data for reanalysis, *American Psychologist*, 61, 726-728.
- Wolfe, J. M. (1998) What can 1 million trials tell us



2014 ■ vol. 10 ■ no. 1

about visual search? *Psychological Science*, 9, 33-39.
Yong, E. (2012) Bad copy. *Nature*. 485, 298-300.

Citation

Cousineau, D., & Hélie, S. (2014). Outstanding data sets: A new category of articles that promotes modelling published in the quantitative methods for psychology. *The Quantitative Methods for Psychology*, 10 (1), 1-4.

Copyright © 2014 Cousineau and Hélie. This is an open-access article distributed under the terms of the *Creative Commons Attribution License (CC BY)*. The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 5/05/13