findings suggest that some sex disparities in career preferences are consequences rather than causes of sex disparities in the professions. We believe that the same social dynamics influence college men's and women's expressed preferences for different kinds of careers.

Because preferences and choices are sensitive to social conditions, it is difficult to determine whether biologically based sex differences in preferences, motives, and styles also exist and whether such differences make men better suited to careers in mathematics and science. Contrary to these claims, male and female infants have been found to show equal interest in people and objects in almost every well-controlled study, but some studies of older children reveal differing interests whose biological and cultural roots are difficult to disentangle. Some evolutionary psychologists suggest that men's and women's differing roles in reproduction led members of the two sexes to pursue different roles in the hunter-gatherer societies in which modern humans evolved (Pinker, 2002). Even if one accepts the controversial claim that differing sex roles in Paleolithic human societies produced differing preferences and temperaments today, however, it is far from clear how these differences would impact on the career choices of aspiring scientists. Because Paleolithic societies contained no scientists, we can only guess whether better science would come from people with the personal qualities of a good hunter or those of a good gatherer. Moreover, we do not know whether the preferences, motives, and styles that inclined early humans toward hunting versus gathering would incline today's humans toward biology versus physics. The immense changes in men's and women's work and lives over the course of human history suggest that people are highly flexible in their interests as well as their abilities.

Despite the great variation in human lives and cultures over space and time, humans have a strong tendency to attribute their current configuration of social roles to constant and necessary aspects of human nature. Moreover, personality traits that are *typical* of a given profession often are mistakenly thought to be *necessary* to the practice of the profession. Winston (1998) discussed a compelling example of this confusion in the letters written by E.G. Boring on behalf of students seeking positions in academic psychology. In the first half of the 20th century, the academic faculties of U.S. universities were overwhelmingly Christian, but Boring had a number of talented Jewish students. How unfortunate, he wrote of one student, that his considerable talents for psychological theory and experimentation could never flourish in an academic career because he shared "the defects of his race": a brash, passionate, assertive manner incompatible with rational academic inquiry. From today's perspective, it appears that Boring mistakenly assumed that the typical mannerisms of his Harvard colleagues were necessary for success in science. When commentators suggest today that fewer women than men have aptitude for science because few women have the thinking style or the assertive, competitive, or aggressive personalities needed for success in science, we suspect that Boring's mistake has resurfaced.

## REFERENCES


Ackerman, P. L. (2006). Cognitive sex differences and mathematics and science achievement. *American Psychologist, 61,* 722–723.

American Medical Association Women Physicians Congress. (2004). *Women in medicine: An AMA timeline*. Retrieved April 28, 2006, from http://www.ama-assn.org/ama1/pub/upload/mm/19/wimtimeline.pdf

Bradt, S. (2006, February 23). High school AP courses do not predict college success in science. *Harvard University Gazette*. Retrieved April 28, 2006, from http://www.news.harvard.edu/gazette/2006/02.23/05-ap.html

Dai, D.Y. (2006). There is more to aptitude than cognitive capacities. *American Psychologist, 61,* 723–724.

Gridley, M.C. (2006). Cognitive styles partly explain gender disparity in engineering. *American Psychologist, 61,* 724–725.

Klopfenstein, K. (2004). Advanced Placement: Do minorities have equal opportunity? *Economics of Education Review, 23,* 115–131.

Morgan, P. & Ramist, L. (1998, February). *Advanced placement students in college: An investigation of course grades at 21 colleges*. Unpublished statistical report (No. SR-98–13) from the Educational Testing Service. Retrieved April 24, 2006, from http://apcentral.collegeboard.com/repository/ap01.pdf.in_7926.pdf

Neumayer, L., Kaiser, S., Anderson, K., Barney, L., Curet, M., Jacobs, D., Lynch, T., & Gazak, C. (2002). Perceptions of women medical students and their influence on career choice. *The American Journal of Surgery, 183,* 146–150.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York: Viking.

Reed, V., & Buddeberg-Fischer, B. (2001). Career obstacles for women in medicine: An overview. *Medical Education, 35,* 139–147.

Sacchetti, M. (2006, February 18). Study hints AP classes overrated. *The Boston Globe,* p. B1.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist, 60,* 950–958.

Winston, A. S. (1998). "The defects of his race": E. G. Boring and antisemitism in American psychology, 1923–1953. *History of Psychology, 1*(1), 27–51.



Thanks to Leo Kamin and Kristin Shutts for valuable advice and to Drew Faust and the Radcliffe Institute for Advanced Study for financial support of this project.



Correspondence concerning this comment should be addressed to Elizabeth S. Spelke, Department of Psychology, Harvard University, 1130 William James Hall, 33 Kirkland Street, Cambridge, MA 02138. E-mail: spelke@wjh.harvard.edu



DOI: 10.1037/0003-066X.61.7.726


# The Poor Availability of Psychological Research Data for Reanalysis


Jelte M. Wicherts, Denny Borsboom, Judith Kats, and Dylan Molenaar
*University of Amsterdam*


The origin of the present comment lies in a failed attempt to obtain, through e-mailed requests, data reported in 141 empirical articles recently published by the American Psychological Association (APA). Our original aim was to reanalyze these data sets to assess the robustness of the research findings to outliers. We never got that far.

This is what happened. In June 2005, we contacted the corresponding author of every article that appeared in the last two 2004 issues of four major APA journals: *Journal of Personality and Social Psychology, Developmental Psychology*, *Journal of Consulting and Clinical Psychology,* and *Journal of Experimental Psychology: Learning, Memory, and Cognition*. We chose to contact these authors because their articles had been published in prominent journals, which would ensure that the articles were of high scientific quality and that the authors were outstanding researchers (all of these journals have rejection rates of over 70%). Also, the fact that the articles were published recently meant that most authors probably still had access to their data and would be able to send them electronically. A final reason for contacting these authors was that, because their articles had been published in APA journals, we were certain that all of them had signed the APA Certification of Compliance With APA Ethical Principles, which includes the principle on sharing data for reanalysis. This principle is as follows:

> 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 and who intend to use such data only for that purpose, provided that the confidentiality of the participants can be protected and unless legal rights concerning proprietary data preclude their release. (American Psychological Association, 2001, p. 396)

Note that our study fit this description adequately: We had no other plans for the data except doing simple reanalyses with various techniques to verify the robustness of substantive conclusions to the influence of outliers. In our e-mail to the corresponding authors, we explicitly stated our mission in these terms, so that the text of the e-mail matched that of the APA guideline closely. Little could go wrong—or so we thought.

## The Problem

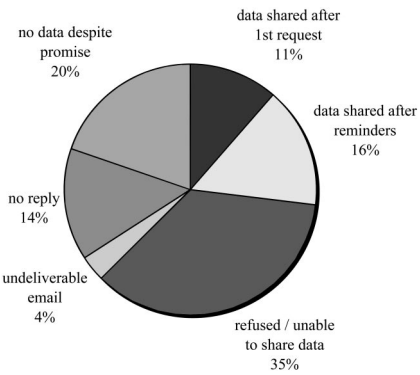
As the 141 articles included a total of 249 studies, we considered acquiring 90 to 100 data sets a realistic aim. We reasoned that adding a follow-up request after the original e-mail would take us a long way in that direction. Unfortunately, 6 months later, after writing more than 400 e-mails—and sending some corresponding authors detailed descriptions of our study aims, approvals of our ethical committee, signed assurances not to share data with others, and even our full resumes—we ended up with a meager 38 positive reactions and the actual data sets from 64 studies (25.7% of the total number of 249 data sets). This means that 73% of the authors did not share their data.

Interestingly, the current response rate shows a remarkable similarity to the response rate (24%) that Wolins (1962) reported over 40 years ago after his student had requested data from authors of 37 articles in several APA journals. Moreover, in a time when data were not electronically available, not readily copied, and not sent easily by e-mail, Craig and Reese (1973) received 38% of the data sets they requested from research published in four APA journals. One of those journals happened to be the *Journal of Personality and Social Psychology* (*JPSP*). Thirty years ago, despite a lack of APA guidelines concerning storage and sharing of data, 75% of data sets published in *JPSP* were shared for reanalysis (Craig & Reese, 1973). Nowadays, one would think that sending data is but a few mouse clicks away. However, as of 2005, only 22% of the data sets published in this journal were available for reanalysis. For the *Journal of Consulting and Clinical Psychology* and the *Journal of Experimental Psychology: Learning, Memory, and Cognition*, we received data from 22% and 41% of the studies, respectively. Of a total of 52 studies reported in the two latest 2004 issues of *Developmental Psychology*, only data from 8 studies (15%) were sent to us. In other words, we were not able to verify by reanalysis whether the conclusions of 85% of the studies in this journal would have been different absent the effects of outliers.

Figure 1 shows a detailed overview of the distribution of corresponding authors over several response categories. These results strike us as dramatic. Of course, accidents happen—data get lost, e-mail addresses become inactive, and computers crash—but a nonresponse of 73% signals a very serious problem even on its most favorable interpretation.

We think that the root of the problem lies in the following aspects of the researchers' situation. Sharing data involves a lot of work for most researchers; in our study, many of the authors who did share their data took considerable time and effort to provide us with a manageable data file and usable codebook. Such work has little to no payoff, as the relevant data have already been published and one does not stand to gain much by having other researchers check one's results. Of course, the APA Ethical Principles are in place, and perhaps most researchers did feel obliged to act in accordance with them. However, like everybody else, researchers are busy people who generally have better things to do than prepare data files and codebooks for a few nosy psychologists from Amsterdam. Such is the human condition.

**Figure 1.**
*Percentages of Empirical Articles' Corresponding Authors in Different Response Categories*




## A Proposed Solution

Fortunately, there is a rather simple solution to this problem. That solution lies at the editorial end of the publishing process. It seems to us that in the electronic age, it would be easy to modify the manuscript submission process to avoid this problem. Only two modifications are necessary, namely: (a) that upon the acceptance of a manuscript for publication, authors be required to submit an ASCII file with anonymized data, as well as a codebook in standardized format and (b) that the journal publish the data and codebook on the World Wide Web as an electronic appendix to the article.

Naturally, authors would only be required to provide data that were used in the analyses reported in the article. They would not be required to provide all the data that were gathered in the study under consideration; such a requirement would be both troublesome and unreasonable. If there were proprietary or other issues that precluded the publication of data, authors could request dispensation (which would have the additional benefit of enhancing transparency with respect to the important question of who actually owns the data).

As an example of how this would work, suppose that an author reports a multivariate analysis of variance with $n$ = 1,000, three dependent variables, four independent variables, and a covariate. In this case, the corresponding data file is simply a text file of eight columns and 1,000 rows. The codebook simply states which column corresponds to which variable in the article (in addition to the coding of missing data, if any). Note that we are not proposing the publication of the sort of full codebooks and scoring protocols that allow the reader to reconstruct the entire research process; hence, for truly detailed reanalyses, one would still have to contact the author. Rather, we are proposing a simple and easy method to ensure that the APA ethical guidelines are effective. In most cases, this would not cost the author or the journal much more work than is already being done in the editorial process. But the benefits of these modifications would be considerable.

First, it would become much easier to reanalyze data; one could, for instance, quickly and effectively confirm the robustness of results with respect to various statistical analyses (e.g., parametric or nonparametric techniques). Second, the feasibility of doing meta-analyses would be greatly enhanced, which would facilitate assessments of the magnitude of effects and the stability of research findings. Third,

it would be easy to address any question one might entertain with respect to data massaging or overly favorable descriptions of results; in addition, such instances of dubious science would, in all likelihood, become highly infrequent as soon as authors knew that other researchers would be looking over their shoulders. Fourth, the procedure would open up a new critical dimension in science because it would be not just the researcher's interpretation of the data that entered the public discourse of science, but the data themselves. Fifth, once data were available on the Web, they would remain available (to all, including the author himself or herself), so that the scientific process would no longer be hampered by the fact that, after a few years have passed, researchers are often unable to locate their data, may have moved to a different university, or may have discontinued scientific work altogether. Sixth, and perhaps most important, scientific evidence should be publicly accessible as a matter of principle; anybody who wants to play in the scientific arena will have to come in with open sight. A procedure like the one proposed would thus increase the openness of scientific research.

It seems to us that, considering the ratio of the benefits achieved in this manner to the costs involved in terms of extra work, this is a bargain. We therefore suggest that the APA journals incorporate the proposed procedure in the publication process.

## REFERENCES


American Psychological Association. (2001). Ethical standards for the reporting and publishing of scientific information. In *Publication manual of the American Psychological Association* (5th ed., pp. 387–396). Washington, DC: Author.

Craig, J. R., & Reese, S. C. (1973). Retention of raw data: A problem revisited. *American Psychologist, 28,* 723.

Wolins, L. (1962). Responsibility for raw data. *American Psychologist, 17,* 657–658.



Correspondence concerning this comment should be addressed to Jelte M. Wicherts, Department of Psychology, Psychological Methods, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. E-mail: j.m.wicherts@uva.nl