

Perspectives on Psychological Science

<http://pps.sagepub.com/>

Are Social Scientists Harder on Their Colleagues Than Physical Scientists Were on Theirs in the Past? Commentary on Trafimow & Rice (2009)

Raymond S. Nickerson

Perspectives on Psychological Science 2009 4: 79

DOI: 10.1111/j.1745-6924.2009.01108.x

The online version of this article can be found at:

<http://pps.sagepub.com/content/4/1/79>

Published by:



<http://www.sagepublications.com>

On behalf of:



[Association For Psychological Science](http://www.sagepub.com/content/4/1/79)

Additional services and information for *Perspectives on Psychological Science* can be found at:

Email Alerts: <http://pps.sagepub.com/cgi/alerts>

Subscriptions: <http://pps.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Issues in Publishing, Editing, and Reviewing

Are Social Scientists Harder on Their Colleagues Than Physical Scientists Were on Theirs in the Past?

Commentary on Trafimow & Rice (2009)

Raymond S. Nickerson

Tufts University

ABSTRACT—*Trafimow and Rice (2009; this issue) have written a thought-provoking article that addresses an important issue in a creative, informative, and engaging way. In a series of vignettes, the authors imagine how several of the better known developments of science might have fared if the manuscripts in which they were first described had been assessed according to the standards and predilections of current reviewers of manuscripts in the social sciences. In this commentary, I note points made by Trafimow and Rice with which I agree, mention some questions that the article raises that are important in my view, challenge the authors' assumption that contemporary social scientists generally treat the ideas of their colleagues more harshly than past physical scientists treated those of theirs, and express an opinion about the merits of the peer-review system as it currently functions in the social sciences. Although I acknowledge that the current system is far from perfect, I argue that it does a passably good job and question whether the reviews it produces are generally too harsh.*

David Trafimow and Stephen Rice have written a thought-provoking article (2009; this issue) that addresses an important issue in a creative, informative, and engaging way. In a series of vignettes, the authors imagine how several of the better known developments of science might have fared if the manuscripts in

which they were first described had been assessed according to the standards and predilections of current reviewers of manuscripts in the social sciences. The motivating assumption appears to be that current-day social scientists are harsher, or more persnickety, critics of colleagues' work than were their counterparts in the physical sciences in the past. The article is an interesting read, because of both the clear and concise account it provides of many of the more important ideas in the history of science and the questions it raises about the current manuscript review process in the social sciences.

My comments fall roughly into four categories. First, I note some points with which I heartily agree. Second, I mention some questions that the article raises that seem to me to be especially important. Third, I challenge the authors' assumption that contemporary social scientists generally treat the ideas of their colleagues more harshly than past physical scientists treated those of theirs. Finally, I express an opinion about the merits of the peer-review system as it currently functions in the social sciences.

POINTS OF AGREEMENT

A major goal of Trafimow and Rice, in their words, "is to dramatize that a reviewer who wishes to find fault is always able to do so" (p. 65), the implication being that the fact that a particular manuscript can be criticized is not sufficient reason to evaluate it negatively. This strikes me as an incontestable point, and the vignettes that are used to make the case heighten the reader's awareness of the numerous ways in which reviews of manuscripts can fail to serve their purposes, which I take to be en-

Address correspondence to Raymond S. Nickerson, 5 Gleason Road, Bedford, MA; e-mail: r.nickerson@tufts.edu.

sure that published research meets certain standards and providing useful feedback to authors.

Trafimow and Rice make numerous other points to which I find it easy to say “right on,” among them the following. “[S]ometimes it is more important to illuminate an important puzzle than to solve a minor one” (p. 75); “[I]t is often a good idea to break the rules, and reviewers should not penalize researchers for doing so without a good reason” (p. 76); and “Reviewers should not recommend extra measures or extra analyses just because they are fashionable in the field, they are routinely performed, or because the reviewers themselves do it in their own research” (p. 77).

IMPORTANT QUESTIONS

One of the many thought-provoking questions that Trafimow and Rice raise touches on the prickly issue of the role of political correctness as a determinant (among others) of what research gets done and what gets published. “How many behavioral science theories that deal with currently hot issues such as prejudice or political action are accepted because they are consistent with what academics want to believe rather than because they are actually good theories?” (p. 75). We will probably never know the answer, but it is a good question to ponder nevertheless.

Another relates to the issue of what constitutes a valuable theoretical contribution. In response to an imaginary reviewer’s objection to Galileo’s relativity principle on the grounds that it lacked empirical support, Trafimow and Rice ask: “But should Galileo really be required to have data to back up his theory? Is it not sufficient to propose a brilliant theory with the possibility that others can find the relevant data?” Perhaps the most famous instance in the physical sciences of a theory being put forth well in advance of any supporting data is Einstein’s theory of special relativity. Is it enough that a theory provide the basis for predictions that can later be put to an empirical test? It is a good question.

Trafimow and Rice speculate that perhaps there is an overemphasis on data in the social sciences, even in journals that specialize in theoretical articles, and that this may discourage the publication of ideas that do not yet have an empirical base. Such an overemphasis risks peremptorily dismissing great ideas. One can wholeheartedly agree with this in principle, while also realizing that great ideas typically are not recognized as such until long after the fact; one might make a case for defining great ideas as those that have been able to survive initial criticism and even ridicule. Einstein’s hypothesis that light exchanges energy with matter in a quantum fashion was not generally accepted among physicists until 20 years after he proposed it in 1905 (Trafimow and Rice, allude to this fact in Footnote 3, p. 72). The problem with putting into practice the principle that data should not be required for the initial airing of great ideas is that of determining which

of the many ideas that are vying for journal space are truly great. Regarding this problem, Trafimow and Rice offer no solution.

CHALLENGES

Trafimow and Rice appear to believe that social scientists today treat the ideas of their colleagues more harshly than did physical scientists in the past. I think this belief should be challenged. As a general rule, scientists who have been responsible for truly revolutionary theories have been treated with much greater respect by their successors, especially remote successors, than by their contemporaries or near contemporaries. Francis Bacon referred to Copernicus’s heliocentric theory as a fiction. Galileo rejected Kepler’s hypothesis that the moon is responsible for the tidal motions of the earth’s oceans as well as his contention that the planetary orbits are not perfectly circular. Huygens and Leibniz ridiculed Newton’s concept of universal gravity because they could not accept the idea of a force extending throughout space that was not reducible to matter and motion. Humphrey Davy dismissed Dalton’s ideas about the atomic structure of matter as more ingenious than important. Lord Kelvin (William Thomson), who died in 1907, several years after the work of J.J. Thomson on the composition of the atom, never accepted the idea that the atom was decomposable into simpler components.

Such examples could be multiplied many times over. Quite possibly some of the arguments made by Trafimow and Rice’s imaginary reviewers and editors in this article against the ideas considered were actually made by contemporaries when they were originally proposed. It is not hard to find cases in which the treatment of new ideas by their originators’ contemporaries yielded discouragement and even embitterment. Newton was so disheartened by the negative reaction to the publication of his *Opticks*, in which he reported his work on the nature of white light, that he refused to publish anything else for several years. Max Planck was sufficiently frustrated and embittered by his inability to get established scientists to pay attention to what he had to say about the second law of thermodynamics in his doctoral dissertation that he claimed to have learned the following remarkable fact: “A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it” (Planck, 1949). Arthur Eddington’s rejection in the mid-1930s of S. Chandrasekhar’s prediction that cold stars with more than a specified mass would collapse to a point sufficiently discouraged Chandrasekhar that he discontinued for the better part of his professional career the line of thinking that eventually led to the theory of black holes.

Were physical scientists of the past more charitable to their colleagues than are social scientists today? It is far from clear to me that they were. Cohen (1985) describes the resistance that revolutionary ideas have typically been met in science this way: “[T]he profundity of a revolution in science can be gauged as

much by the virulence of conservative attacks as by the radical changes in scientific thought it produces” (p. 414).

THE PRESENT PEER-REVIEW SYSTEM

Surely every journal editor and grant review panel member has been impressed with how variable the reviews of a given manuscript or proposal can be. This impression is borne out by empirical studies of the peer-review process that reveal a relatively low level of agreement among reviewers of the same manuscript or grant proposal (Cicchetti, 1991; Marsh & Ball, 1989; Marsh, Jayasinghe, & Bond, 2008). Of particular interest in the present context is the finding of slightly more inter-reviewer agreement among social scientists than among physical scientists (Jayasinghe, Marsh, & Bond, 2003). The finding of low agreement among reviewers represents a serious challenge to the assumption, if anyone makes it, that the peer-review process ensures that only the best manuscripts get published or that only the most deserving proposals get funded.

So the process is flawed and has limitations. Could it be improved? Probably. Could efforts to improve it do more harm than good? Perhaps. Does it work acceptably well as is? I want to argue that it does.

In my experience, reviews of my own manuscripts have been extremely helpful, with very few exceptions. While reading Trafimow and Rice’s vignettes, I found myself sometimes wondering how accurately the examples reflect the way the vast majority of reviewers and editors actually express themselves. Many of the turns of phrase of the imagined reviewers and editors are, thankfully, foreign to me. I have never seen a letter from a reviewer or editor that included such comments as “everything you say, despite the fact that it could be correct, is so unlikely that only irrational people would believe it” (p. 67); “Both reviewers felt that your dependence on the exactness of your measures, and what they found to be byzantine reasoning, . . .” (p. 69); “you cannot prove a negative, which you were foolish enough to attempt” (p. 69); “. . . found your arguments about time dilation to be simply insane” (p. 71). I do not mean to claim that such comments are never made, but I strongly suspect that they are rare and not characteristic of the comments authors typically receive from either reviewers or editors. Perhaps Trafimow and Rice were sometimes caricaturizing for effect, but that was not clear.

Trafimow and Rice raise the question of “whether the conservative nature of science is always a good thing” (p. 75). This question takes a prevailing conservative bias in science as a given, which, for present purposes, I do not want to challenge. I find it easy to accept the notion that it is not always a good thing, but that it is generally a good thing. Perhaps this is the position that Trafimow and Rice had in mind when noting that the degree of conservatism that should be applied in evaluating new ideas is a judgment call that should be made on a case-by-case basis.

But what principle(s) should guide the judgment call? When all the reviews have been written and absorbed, the editor has to make a binary decision: accept or reject. A conscientious editor wants to accept only those manuscripts that should be published and to reject only those that should not be. But the basics of statistical decision theory make it clear that it is generally not possible to set a quality criterion so as to accept all the manuscripts that should be published without also accepting some that should not be. Relaxing the criterion so as to accept more of the deserving manuscripts ensures also acceptance of more manuscripts that should not be published. The challenge is to find a criterion that establishes an acceptable trade-off.

The question of what the requirements for acceptance (or rejection) of a manuscript should be undoubtedly deserves much thought and discussion. In their Table 1, Trafimow and Rice identify 18 reasons for rejection and associate them with the scientists against whose great works they might have been used. In most of the examples of how these reasons could have been applied, they illustrate how they could have been misused. The authors say very little about which of the listed (or other) reasons they believe to be legitimate elements of a critique of a manuscript or the conditions under which they would be valid. Nor do they give much space to discussion of what they consider the characteristics of an ideal, or at least adequate, critique to be. Their emphasis is on identifying characteristics that a review should not have, and the objective appears to be to minimize the likelihood of rejecting manuscripts that have potential; comparatively little is said about the characteristics of a good review or about the importance of precluding the publication of manuscripts that add mainly noise to the literature.

Trafimow and Rice argue that “if the theory seems unlikely to be true, it might be a crackpot idea but it also might be brilliant, as brilliant theories often seem unlikely to be true” (p. 77). Granted that one cannot always rule out the possibility that an idea is really brilliant, even one that appears to be untrue, but it does not follow that all theories that seem unlikely to be true deserve journal page space. Nor do Trafimow and Rice claim that it does. They argue that perhaps reviewers should ask themselves whether the theory, if true, would be important, and if the answer is positive, “perhaps the reviewer’s opinion that the theory is unlikely to be true should be deemphasized” (p. 78).

This proposal raises an interesting question. Imagine that importance can be quantified on a scale from, say, 0 to 10, and that one can independently assign a probability that the idea is true. Should the decision regarding publication be made solely on the basis of judged importance, say if the judged importance is 7 or greater? Should it be made on the basis of the product of judged importance and the estimated probability of it being right, say if the product is rated 5 or greater? The second criterion would allow rejection of even very important ideas if the estimated probability of their being true is sufficiently low, but

the first would not. Of course, with either criterion, the exact cutoff value for acceptance would be debatable. What I wish to argue here is that the importance of an idea, assuming it is true, should not be enough to justify acceptance without regard for the probability that it is indeed true. It is easy to generate numerous hypotheses that would be enormously important if true but that have a vanishingly small probability of being so.

Trafimow and Rice argue that “if the reviewer has a difficult time getting his or her mind around the concept, then perhaps the author has a truly innovative or revolutionary perspective” (p. 77). If the point is that reviewers should keep in mind their own fallibility as judges of the quality of other’s ideas, I strongly agree. But difficulty in getting one’s mind around a concept could also be a sign of a confused idea or lack of clarity in exposition; it is not, by itself, a justification for recommending publication.

Trafimow and Rice contend also that the ability to point to one or more explanations of a finding alternative to the one proposed is not, by itself, adequate grounds for rejection. This seems right to me. One needs to evaluate the relative merits of the alternative explanations in terms of such considerations as plausibility, parsimony, and consistency with published data.

Sometimes a reviewer’s objections to a manuscript can be discounted by virtue of being considered by an editor to be wrong. Trafimow and Rice illustrate this by reference to an imaginary reviewer’s contention that Einstein’s special theory of relativity was only incrementally different from Galileo’s theory of relativity. Clearly when a criticism is known by an editor to be factually wrong it should be ignored. The more difficult question is how to judge the merits of criticisms that are known to be correct, or at least not known to be wrong.

Trafimow and Rice anticipate the possibility that readers might interpret their assessment of the review process as it is currently done in the social sciences as an argument that reviewers should be less critical, and they contend that this would be a misinterpretation of their intentions. “On the contrary, our position is that researchers should be more critical. But part of critical thinking is to criticize the potential criticisms, and determine if they really are important, and be willing to consider that they might not be” (p. 77). This strikes me as an excellent point—casting a critical eye on one’s own criticisms before passing them on for others’ consumption is an eminently good idea.

This suggestion and others in Trafimow and Rice’s thought-provoking article invite reflection on the importance of training in reviewing. I am not aware of data regarding how much effort is made in psychology graduate programs to teach students to be good manuscript reviewers. For those preparing for a research career, reviewing is an important skill and worthy of some focused attention. What does it take to be a good reviewer? What are the characteristics of a good review? Long before becoming independent researchers, students should have assimilated a set

of standards for assessing the publishability of research reports: they should understand the difference between constructive and destructive criticism, they should realize the importance of attitudes and objectives in reviewing, such as the difference between trying to help an author improve a manuscript and exhibiting one’s own superior cleverness, and so on.

A few years ago, I distributed a questionnaire to researchers who had submitted manuscripts to the *Journal of Experimental Psychology: Applied* during its fledgling period. The purpose was not only to obtain some feedback about how the review process, as conducted by that journal, was perceived by manuscript contributors, but also to get some sense of what contributors expected or wanted from reviewers and editors (Nickerson, 2005). Respondents rated reviews and action letters on a 7-point scale with respect to clarity, justification, and helpfulness. Both reviews and action letters received relatively high ratings with respect to all three criteria; the lowest of the six means was 5.27. Ratings by authors of manuscripts accepted for publication did not differ appreciably from those by authors of rejected manuscripts.

Respondents also rated the importance of comments from reviewers regarding each of the following issues on a 7-point scale: soundness of method, appropriateness of data analyses, evidential justification of conclusions, clarity of exposition, theoretical importance of findings, objectivity of interpretation of results, adequacy of coupling to related work, practical importance of findings, and uniqueness of contribution. The order in which the issues are mentioned here reflects the order of their importance (from greatest to least important) as rated by the participants, but perhaps more revealing is the fact that all were rated relatively highly and the spread between the mean ratings of most-important and least-important was small (6.10 to 5.23). The importance of criticism being tactfully phrased was given a mean rating, also on a 7-point scale, of 5.03, suggesting that tact is appreciated, but not more than substantive feedback regarding the strengths and weaknesses of a manuscript.

Ninety percent of the respondents expected reviewers to do substantially more than advise an editor regarding a manuscript’s publishability. A majority (77%) expressed preference for an editorial decision with detailed substantive feedback regarding problems and suggestions for improvement 10 to 14 weeks after submission over a minimal response (editorial decision and main reasons if rejected) 4 to 6 weeks after submission. Also, 74% considered it better for authors if reviewers err on the side of being too demanding (too critical) than on that of being too lenient (not critical enough), and 81% considered that to be better for the field.

These data are from one small study, collected from contributors to a single journal, and, in particular, from those contributors who elected to respond to the questionnaire, so I do not mean to treat them as representative of the views of all contributors of manuscripts to psychology or social science

journals. However, at least for this sample of contributors, the data suggest that the review process has been perceived as working reasonably well, and they provide little evidence of a strong desire among contributors for reviews to become less stringent or for publication criteria to be relaxed in any major way.

If the contention that social scientists are unduly or unjustifiably critical as manuscript reviewers is valid, rectification would seem to call for less critical reviews with the consequence that social scientists would find it easier to publish. Would this be good for social science (or social scientists)? Perhaps. Would it be good for science or society more generally? Perhaps. But I think the answer is far from clear in both cases.

REFERENCES

- Cicchetti, D.V. (1991). The reliability of peer review for manuscript and grant submissions: A cross-disciplinary investigation. *Behavioral and Brain Sciences, 14*, 119–135.
- Cohen, I.B. (1985). *Revolution in science*. Cambridge, MA: Harvard University Press.
- Jayasinghe, U.W., Marsh, H.W., & Bond, N. (2003). A multilevel cross-classified modeling approach to peer review of grant proposals: The effects of assessor and researcher attributes on assessor ratings. *Journal of the Royal Statistical Society, Series A, 166*, 279–300.
- Marsh, H.W., & Ball, S. (1989). The peer review process used to evaluate manuscripts submitted to academic journals: Interjudgmental reliability. *Journal of Experimental Education, 57*, 151–169.
- Marsh, H.W., Jayasinghe, U.W., & Bond, N.W. (2008). Improving the peer-review process for grant applications. *American Psychologist, 63*, 160–168.
- Nickerson, R.S. (2005). What authors want from journal reviewers and editors. *American Psychologist, 60*, 661–662.
- Planck, M. (1949). *Scientific autobiography and other papers* (F. Gaynor, Trans.). New York: Philosophical Library.
- Trafimow, D., & Rice, S. (2009). What if social scientists had reviewed great scientific works of the past? *Perspectives on Psychological Science, 4*, 65–78.