

Miller, T. I. *The effects of drug therapy in psychological disorders: A meta-analysis*. Unpublished doctoral dissertation, University of Colorado, 1978.

Rachman, S. *The effects of psychotherapy*. London: Pergamon Press, 1971.

Sloane, R. B., Staples, R. F., Cristol, A. H., Yorkston, N. J., & Whipple, K. *Psychotherapy versus behavior therapy*. Cambridge, Mass.: Harvard University Press, 1975.

Smith, M. L., & Glass, G. V. Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 1977, 32, 752-760.

GENE V GLASS

MARY LEE SMITH

Laboratory of Educational Research
University of Colorado

A Guide for Reviewers: Editorial Hardball in the '70s

The policies and practices of journal reviewers have undergone a marked shift recently. In the past, reviewers responded primarily to the judged importance, theory relevance, and originality of the idea under study. Such judgments are inherently subjective, however. In recognition of the scientific nature of the psychological enterprise, the new trend is toward simply listing the number of methodological flaws and conceptual alternatives that can be imputed to a given study. A recommendation as to whether to accept the article is then based simply on the sum of the flaws and alternatives. Some reviewers are not fully aware of this shift in practices, and the following rules are offered for those who wish to understand this new trend and join ranks with the forefront.

1. The best way to handle the question of the importance of the idea under investigation is to operate under the assumption that there is a perfect inverse correlation between its importance and the number of methodological flaws and alternative interpretations of the study.

2. The reviewer may similarly finesse the question of novelty of the idea under investigation. Be vigilant, however, for novel methods

of investigation. Novel methods are suspect because untried.

3. Add points for "real world" investigations because these are still in vogue. Subtract points for ecological validity in the laboratory setting, however, because this is not so much in vogue and, other things being equal, more ecological validity means messier stimulus materials. Messy stimulus materials mean many artifactual explanations.

4. Bear in mind that *every* study has artifactual possibilities and that your prowess as a reviewer consists in adducing as many of them as possible. Keep the list below handy, since most of the examples can be applied to any study.

a. *Experimenter demand characteristics*. The nice thing about this one is that you don't have to think about exactly *how* demand characteristics could have produced the results. Simply assert it. Editors know that all studies are prone to them and God knows the authors realize they are guilty. They know they didn't take all those Rosenthal precautions.

b. *Social desirability*. This is another one that doesn't actually have to be thought through. The lovely thing is that obviously the subjects wouldn't have said or done what they said or did if they didn't think it was desirable. There's really no way out of this one.

c. *Motivational relevance*. If the study engages the subjects' interest and deals with matters that are important to them, then assert that the findings were obtained only because of the motivations or defenses that were aroused by the procedures.

d. *Limitation to trivial situations*. If the study does *not* engage the subjects' interest, then so much the better. It may be claimed that the phenomenon under study would not hold up under any but the barren laboratory situations studied.

e. *Subjects' comprehension of instructions*. If the instructions to subjects were long and complicated, assert that the subjects probably

didn't understand them. The same criticism may be applied if the instructions were brief.

f. *Confounded manipulations*. Except for the very smallest and most circumscribed of manipulations, almost all independent variable manipulations involve changing more than one stimulus component. This provides a rich source of artifactual explanations. It is important to note that the artifactual explanations need not be at all plausible. The important thing is that they exist, and the author should be forced to rule them out methodologically. This may often mean that the manipulation will become so small and inconsequential that the infant phenomenon is thrown out with the alternative bathwater. But that's the author's problem, not yours.

g. *Verbal versus behavioral measures*. Criticize any article employing only verbal measures on the grounds that verbal measures are only verbal measures. Criticize any article having only behavioral measures on the grounds that the behavior might have been produced by any of a number of a cognitive states that were not tapped verbally. Criticize any article in which verbal measures follow the behavioral measures on the grounds that the verbal results were obtained only because the subject was trying to be consistent with his or her previous behavior. Criticize any article in which verbal measures precede the behavioral measures on the grounds that the behavioral measures were obtained only because subjects were trying to be consistent with their verbal reports. If the order of verbal and behavioral measures is counterbalanced, employ both of the latter criticisms.

5. Occasionally it may be difficult to find trivializing explanations, or the author may have spoiled things by arguing cogently against them. The enterprising reviewer is not daunted by this state of affairs. It is nearly always possible to find some interesting, nontrivial, alternative explanations for the results. In

that event, insist that more research be done to choose between the author's explanation and your own. If it would be extremely difficult to design research to test among the possibilities, that's the author's problem. After all, who chose this line of research?

6. The perspicacious reviewer leaves no theoretical statement unchallenged. Nearly always, theoretical statements are too broadly worded, and of course the results can never establish that the theoretical statement is generally true. Any theoretical statement left standing must be hedged with at least three qualifying phrases. In addition, any reviewer worth his or her salt can think up as good a theoretical position as the author's in a few minutes' time. The author may then be criticized for failing to take this position into account in the manuscript.

7. Final rule for editors: In the old days, articles were accepted or they were rejected outright with a brief explanation of the reason or reasons. The modern editor never makes such a simple-minded decision. Instead, all authors are sent a six- to eight-page letter supplementing and sometimes contradicting the reviewer's four- to six-page critical commentary. This letter begins with the following sentence: "The reviewer and I are in agreement that your manuscript is not acceptable in its present form." The author is then required to deal with the 10-14 pages of methodological, statistical, and conceptual criticism and to cut the length of the article by one third. Finally, the author is told that such a revision might be acceptable but that of course no commitments can be made in advance. This effectively discourages the authors of inferior manuscripts but leaves them feeling much better than if the manuscript had actually been rejected. The author of a superior manuscript will revise and resubmit, along with a 10- to 15-page letter attempting to justify his or her failure to comply completely

with the editor's instructions for revision. This initiates what is usually the first of several challenging exchanges between editor and author. Along the way, authors of less superior manuscripts tend progressively to drop out, leaving only the finest manuscripts for publication. These manuscripts, of course, have been greatly improved by dealing at length with all possible methodological and conceptual criticisms and by largely abandoning the theoretical considerations that gave rise to the study.

RICHARD E. NISBETT
University of Michigan

An Evaluation of 80 Psychology Journals Based on the Science Citation Index

Much recent effort has been devoted to determining a relative ranking of psychology journals in terms of excellence. Mace and Warner (1973) sampled opinions of departmental chairpersons to determine such ratings, while Koulack and Keselman (1975) obtained more extensive evaluations by sampling opinions of members of the American Psychological Association. These ratings have been criticized on various grounds (e.g., Boor, 1973; Buss & McDermott, 1976; Gynther, 1973; Hohn & Fine, 1973; Levin & Kratochwill, 1976; Porter, 1976). One major drawback to these analyses is that they simply survey subjective opinions that may be unrelated to more objective assessments of journal esteem and impact. There have been three recent attempts to obtain a more objective measurement of journal worth by using some form of citation analysis (Buss & McDermott, 1976; Porter, 1976; White & White, 1977).

The most comprehensive attempt to obtain an objective assessment of psychology journals was that of White and White (1977). Briefly, they sampled every 10th page of the 1974 *Social Science Citation Index* (SSCI) and counted the number of references to articles appearing in

57 psychology journals in 1972 and 1973. By multiplying the number of citations found with this sampling procedure by 10, they were able to estimate the total number of times that articles appearing in the 1972 and 1973 issues of the 57 journals were cited in 1974. In addition, by dividing this total by the number of articles appearing in the journals in those years they were able to calculate the average number of citations to an article in each of the journals. This number, the average number of citations per article, is referred to as the *impact factor* (Garfield, 1972). White and White (1977) ranked the 57 psychology journals by this impact factor.

There are several difficulties with the White and White (1977) attempt at assessment of psychology journals. First, since they sampled only every 10th page, there is the problem of the reliability of the figures they obtained. They correlated the rankings obtained from the first half of the sample with those from the second half, and their Pearson r of .72 does not inspire confidence that this sampling procedure produced reliable results. If the sampling procedure is not reliable, then obviously the derived estimates of total citations, the impact factors, and the ratings are also open to serious question. Obviously what is needed is an evaluation of psychology journals based on a much more exhaustive sample of citations.

Another difficulty of the White and White (1977) study is that a number of important psychology journals were not included in their sample. Some notable omissions from their study were *Cognitive Psychology*; *Psychophysiology*; *Learning & Motivation*; *Perception & Psychophysics*; *Memory & Cognition*; and *Physiological Psychology*.

The present report is another attempt to evaluate the relative impact of psychology journals using citation analysis, but correcting to a large extent the shortcomings of