REFERENCES

1. Maxwell C: Hospital organizational response to the nuclear accident at Three Mile Island: implications for future-oriented disaster planning. Am J Public Health 1982; 72:275–279.
2. United Nations' Scientific Committee on the Effects of Atomic Radiation. United Nations, 1977.
3. Statement by HEW Secretary Joseph Califano before the Subcommittee on Energy, Nuclear Proliferation and Federal Services of the Senate Committee on Governmental Affairs, May 3, 1979, Washington, DC.
4. Report of the President's Commission on the Accident at Three Mile Island, 1979. Washington, DC: US Govt Printing Office, Document No. 052-003-00718-5 1979.
5. MacLeod GK: Some public health lessons from Three Mile Island: a case study in chaos. AMBIO (Royal Swedish Academy of Sciences), a Journal of the Human Environment 1981; 10:18–23.
6. Muller HA: Data on TMI disputed. Penna Med 1981;84:32–33.
7. New Castle Pennsylvania News: Administration Rapped for Delay in Thyroid Probe, United Press International: February 22, 1980.
8. Ford D: Three Mile Island, *New Yorker*, April 6, 1981, pp 49–120.
9. MacLeod GK: Medical ethics in the nuclear age. Penna Med 1981; 84:25–31.
10. Houts PS, Miller RW, Ham KS, Tokuhata GK: Extent and duration of psychological distress of persons in the vicinity of Three Mile Island. Proceedings of the Pennsylvania Academy of Sciences 1980; 54:22–28.
11. Dohrenwend BJ, Dohrenwend BS, Warheit GW, Bartett GS, Goldsteen RN, Goldsteen K, Martin JL: Stress in the community: a report to the President's Commission on the Accident at Three Mile Island. Ann NY Acad Sciences 1981; 365:159–174.
12. Gerusky TM: Oral Deposition to the President's Commission on the Accident at Three Mile Island, July 24, 1979.

Editor's Report—Peer Review Again

During 1981 the Journal received its usual number of unsolicited manuscripts, 703, more than four out of five of which we would be unable to publish because of space considerations, regardless of their merit.* Most of these submissions passed through the hands of one or more of the 487 reviewers listed on pages 297–298, or the Editorial Board members listed on page 237. The voluntary input of both groups is what makes any professional journal work. The credibility and integrity of such individuals has been challenged in the past, as I pointed out three years ago in these columns.[1] It has been challenged again, at least by inference, in a recent study of the National Science Foundation's (NSF) peer review processing of research grant applications.[2]

The authors of the NSF study are careful to point out that its results cannot be extrapolated to the grant processing of the National Institutes of Health which has also been criticized; the latter involves dialogue and consensus formation on the part of a panel which reviews both the proposal and any critiques of it by outside consultants. In the same way, the NSF study results could be said not to apply to the peer review process of professional journals; the latter is not consensus forcing, but the track record of the investigator is not one of the elements that influence decisions, as is the case with grant reviews.

Nevertheless, the NSF study is bound to cause comment and uneasiness among all those faced with making decisions of a comparable kind, not to speak of those who must, willy-nilly, submit to such decisions. The investigators reached the conclusion that getting a research grant depends to a significant extent on chance; they found reviewer disagreement to be such that, in about half the cases (they report), the values of the reviewers, rather than the merits of the project, determined the outcome.

This conclusion is reminiscent of the finding that referee agreement on papers submitted to professional journals is only slightly better than would be expected by chance.[3] It follows that if a paper is submitted to enough journals it stands a good chance of eventually being published; this corollary is sustained by the experience of others[4] as well as of this Journal.** Nevertheless, this is circumstantial evidence, subject to other interpretations. Before accepting it or the NSF conclusions at their superficial face value, it may be worth taking a closer look.

The NSF study took a series of grant requests that had been peer reviewed, selected another group of peer reviewers from a panel of "experts", and compared the outcome (accept or reject on the basis of numerical score) of the two sets. There is no escaping the fact that the granting agency or editor selects the reviewers, and that value judgments may enter into the selection. In the NSF study, however, the characteristics of the two sets of reviewers were similar, and it was not mathematically possible to demonstrate any systematic bias in reviewer selection so that this explanation was discarded. Procedural differences in the reviewing process seemed also to fail to explain the reversal in verdict that appeared (in one-fourth of the cases) when the scores of the two sets of peer reviewers were rank ordered. The size of the variance of the mean scores (i.e., discrepancies in the scores of individual reviewers reviewing the same grant) of each set of reviewers seemed to offer the best explanation of the reversals.

One problem with both the NSF study analyses and those of referee agreement in journal reviewing is the fact that reviewer recommendations on a check-off sheet must be translated into numerical values in order to be subjected to statistical treatment. Scaling is a common procedure, but interpretation of the terms (poor, fair, good, very good,

*In 1981 we published 83 articles, 56 public health briefs, one supplement, 30 editorials, 16 commentaries, nine "different views", 50 letters to the editor (usually with one or more responses), and four articles in special sections (Public Health Then and Now and Public Health and the Law).

**About 25 per cent of the papers we publish have been previously submitted to another journal. At least half of the papers we reject (tracked via MEDLINE three years after our rejection) have been published later in another journal.

excellent—or reject, accept if revised, accept) will obviously differ among reviewers. This is where values enter the picture, but such values express a reviewer's standards as applied to words or their understanding of the funding or publication process rather than their views about the merits of a proposal or paper. The contrast between values and merit are reflected in my experience by comparisons of the comments and ratings of two reviewers with virtually identical comments but different recommendations for or against publication.

A corollary problem is that the priorities of a research program or a journal, as well as the merits of a proposal or paper, enter into the decision of an agency or editor. In the case of journals, this factor was described with some hyperbole by Sir Theodore Fox, former editor of the Lancet.

> "The successful are selected in a competition which is often decided on factors quite outside the author's control—for instance, whether the journal already has three other papers on the same subject, whether the editor thinks he has published enough articles from Tierra de Fuego for the time being, or enough articles on Reiter's disease; or on the other hand that for the next few months he will take almost anything on this disease because someone he met in Seattle described it as one of the key conditions in contemporary medicine."[5]***

One must also point out that an unknown proportion of rejected proposals or papers undergo substantial rewriting as a result of the comments of reviewers, and eventually are accepted or published by the same authority which rejected the original submission. In the case of this Journal, I estimate this to be the case with about five to ten per cent of our published papers. Of those papers we publish that were previously submitted to other journals, three out of four have been substantially revised before we publish them, either before or after their submission to us.

Although not quite as high, a significant proportion of the papers first submitted to us and eventually accepted have also been revised substantially. Revisions of ths extent could be taken as evidence of the value of the peer review process. They could also be interpreted as indications that authors will do anything needed to get their paper published whether or not they agree with the criticisms—circumstantial evidence can often lead to the antipodes.

For all the criticism of the peer review process, no one has come up with another way to select a few out of many. In journal reviewing, modifications have been suggested such as the signing of reviews or (as in our case) blocking out the authors' identity where feasible. There is no evidence that either modification has removed any biased or chance factors that influence the decision. Furthermore, in the end it is the program administrator or the editor who makes the decision, not the reviewers—whatever their scores or opinions.‡

Another way of looking at the NSF study findings is to be grateful for the fact that another set of reviewers confirmed the fact that half of the grant request decisions was based on the proposal's merit or lack of merit. Actually, for reasons already stated, I think the proportion of decisions based (in the investigators' terms) on merit was probably higher than 50 per cent. If, in the case of journal reviewing (where even more decisions do not reflect on the intrinsic merits of a paper) we can asssume that as many as three-fourths of the decisions are made for good and sufficient reason while one-fourth of them are subject to chance, it would mean that a reversal, in the decision of the chance-influenced 25 per cent, would occur only in one out of eight papers that passed through the review process. Given the human condition, it would be foolhardy to think we can ever eliminate chance. It probably plays a larger role than this in most of our lives. Furthermore, in the case of journals, there are always second or third or more chances.

We should always question ourselves and strive to do better—but perhaps the NSF study findings do not negate the value of peer review in journal editorial decision making.

ALFRED YANKAUER, MD, MPH

Postscript

Since this editorial was written, as might have been predicted, letters to the editor of *Science* have burgeoned.[6,7] The letters raise most of the points in the preceding editorial as well as a few others for the more mathematically minded.

***In my subjective analysis of my own decisions, I found that factors other than quality or significance entered into the decision in one-third of the peer reviewed manuscripts rejected and were the primary reason for rejection in one-fifth of the cases.

Address reprint requests to Dr. Alfred Yankauer, Editor, *American Journal of Public Health,* Department of Family & Community Medicine, University of Massachusetts Medical Center, Worcester, MA 01605.

REFERENCES

1. Yankauer A: Editor's report: peer review. Am J Public Health 1979; 69:222–223.
2. Cole S, Cole JR, Simon GA: Chance and consensus in peer review. Science 1981; 214:881–886.
3. Ingelfinger FJ: Peer review in biomedical publication. Am J Med 1974; 56:686–692.
4. Relman AS: Are journals really filters? *In:* Coping with the Biomedical Literature Explosion: A Qualitative Approach. Rockefeller Foundation Papers 54–78, December 1978.
5. Fox T: Crisis in Communication. London: Athelone Press, University of London, 1965.
6. Atkinson RC, Singer IM, Cronbach LJ: The Peer Review Question (letters). Science 1981; 214: 1292–1293.
7. Clark AH. Luck, merit and peer review, Science 1982; 215:11.

‡In my experience, 18 per cent of the papers recommended for rejection by one reviewer are, in fact, accepted, while 22 per cent of the papers with recommendations by two reviewers to accept or accept if revised are, in fact, rejected. Most, but not all, of these decisions are made without the help of a third reviewer's recommendation.