Current Comments®

EUGENE GARFIELD
INSTITUTE FOR SCIENTIFIC INFORMATION®
3501 MARKET ST., PHILADELPHIA, PA 19104

Refereeing and Peer Review. Part 4. Research on the Peer Review of Grant Proposals and Suggestions for Improvement

Number 5 February 2, 1987

This is the conclusion to a four-part series on refereeing and peer review in science. The first two parts discussed the refereeing of scholarly articles prior to publication.[1,2] The third part focused on the mechanics of the peer-review system for the evaluation of research-grant proposals at the National Science Foundation (NSF) and the National Institutes of Health (NIH) and opinions about those systems.[3] This part examines the research on peer review and some proposed alternatives and improvements.

The COSPUP Study

One of the best-known and most thorough studies of peer review was conducted by sociologists Stephen Cole and Leonard Rubin, State University of New York (SUNY), Stony Brook, and Jonathan R. Cole, Columbia University, New York. At the request of the Committee on Science and Public Policy (COSPUP) of the US National Academy of Sciences (NAS), the Coles and Rubin examined the peer-review system of the NSF. Phase one of the study, a retrospective statistical analysis of "how peer review works in the day-to-day operation of the Foundation [NSF],"[4] (p. 17) was started in 1974 and completed in 1978. Phase two, coauthored by the Coles and COSPUP, reported the results of experiments designed to address the question of whether program officers influence the peer-review process through their selection of reviewers. It was started in 1978 and published in 1981.[5]

In phase one, the authors interviewed 70 scientists involved in all stages of the peer-review process, including current and former NSF program directors, advisory- and review-panel members, NSF section and division heads, and the director and associate director of the NSF.[4] (p. 18) To determine the most decisive factors in securing a grant, they collected data on 1,200 applicants, half of whom had been successful. In some cases, the authors examined not only the proposal but also the reviewers' comments, correspondence, and all paperwork connected with the funding decision.

Phase two was carried out in two stages. First, the Coles submitted 150 proposals previously reviewed by the NSF to a set of surrogate program directors. Half of the surrogates received proposals that had been edited in an attempt to conceal the authors' identities; the other half received copies that were exactly as they had been submitted to the NSF. The surrogate directors were asked to name a set of possible referees for the proposals, and the Coles once again attempted to conceal the identities of half the authors. None of the participants knew how the proposals had been rated by the NSF. The Coles asked them not only to evaluate the proposals but also, when applicable, to try to identify the authors. Reviewers of "blinded" proposals were also asked whether the removal of title pages, lists of references, budgets, and other identifying comments made the proposal more difficult to evaluate.[5] (p. 6-19)

The COSPUP Findings

The main conclusion of phase one is that peer review in the NSF functions fairly.[4] (p. viii-ix) The authors found a high correlation between high reviewer ratings and favorable funding decisions. They also found that an applicant's age and track record had little effect on the chances of getting a grant and that reviewers from major, "high-status" institutions treated proposals from researchers at prestigious institutions no differently than proposals from workers at less-prestigious institutions.

On the whole, the results of phase two corroborate the findings of phase one: the Coles found no evidence of bias on the part of program officers in their selection of reviewers and no evidence that external criteria such as gender, age, and race had any influence on reviewer decisions.[5] (p. 4) In the matter of blinded proposals, the Coles found it difficult to conceal authorship: "To omit all possible identifiers, in addition to the name(s) of the author(s) of the proposal, made the proposal almost unreadable," said Jonathan Cole.[6] This was reflected by the opinions of the COSPUP reviewers, who felt that the blinding process severely compromised the integrity of the proposals. Nevertheless, proposals that received high ratings by NSF reviewers generally received high ratings from COSPUP reviewers as well.

However, "there was a great deal of well-considered variance in opinions among equally qualified reviewers," in the words of Jonathan Cole.[6] "Thus, if we work with a small number of reviewers and a high variance in opinion, the outcome of an evaluation will depend greatly on the people selected to review the proposal.... This is not to imply that the process is 'unfair,' but that there is a substantial level of reviewer disagreement on rational grounds, e.g., quality of past work, priority given by a particular reviewer to the subject of the proposal, the assessment of the methods to be used, etc."[6]

The Coles concluded that perhaps 25 to 30 percent of NSF funding decisions would be reversed if applications were evaluated by another, equally qualified group of reviewers. In both their phase-two monograph[5] and a paper they published in *Science* with statistician Gary A. Simon, SUNY, Stony Brook,[7] the Coles acknowledge that some scholars, taking note of this, will feel that the complicated system of peer review "does not buy you much."[5] (p. 43) Jonathan Cole points out, however, that "there is apt to be a great deal of disagreement on the contents of proposals...at the cutting edge of scientific inquiry,...and therefore, we should not be wholly surprised at the proportion of reversals."[6] Such reversals probably indicate that no "single, agreed-upon dogma"[7] is dominant in the fields studied, and, in fact, one of the most surprising results of the COSPUP study was that the level of consensus among reviewers was no higher in physics than in the social sciences.[4-7]

Phase one of the COSPUP study has been cited in over 74 papers since it appeared in 1978; phase two has been cited in 17. The *Science* article has been cited 49 times through 1986 and is one of four papers forming the core of a research front entitled "Alternatives to, arbitration in, and other aspects of peer review of scientific journals and research proposals" (#85-4243). The other three core papers include the classic study of patterns of evaluation in science by Harriet Zuckerman and Robert K. Merton, Columbia University;[8] a controversial study of bias in the journal refereeing process by Douglas P. Peters, University of North Dakota, Grand Forks, and Stephen J. Ceci, Cornell University, Ithaca, New York;[9] and a paper on the rate of agreement between reviewers by psychologist Grover J. Whitehurst, SUNY, Stony Brook.[10] All three papers were mentioned in Parts 1 and 2 of this essay.[1,2]

According to Jonathan Cole, the key policy implication of the COSPUP study was

that "the lower the number of reviewers used to evaluate the proposal, the greater the chance for...reversals."[6] As a result, the NSF now requires a certain minimum number of reviewers for every proposal it receives. Science journalist Tineke Boddé lists a number of other changes in the NSF system that have been made more recently.[11] For instance, the entire process has been streamlined, with a limit of 15 pages per proposal and a policy requiring a decision within nine months. Specific guidelines on conflicts of interest have been established, verbatim copies of all reviewer comments have been made available, and a system has been set up to reconsider declined proposals. Under certain circumstances, some proposals are now exempt from peer review, and program officers can extend existing grants without further review if they feel outstanding progress has been made.[11]

Peer Review in the NIH

Fourteen scientists and administrators from various agencies within the NIH were appointed to the NIH Grants Peer Review Study Team by then-acting director, Ronald W. Lamont-Havers. Chaired by Ruth L. Kirschstein, director, National Institute of General Medical Sciences, the team was charged with evaluating the NIH's peer-review system and with making, where applicable, recommendations for improvement.[12] In making its assessment, the team printed an open solicitation in the *Federal Register*[13] and mailed a memorandum to 30,000 individuals, asking for written comments on the peer-review system (1,500 replies were received). The team also held open public hearings for the scientific and lay communities. The team members considered everything they read and heard, according to William F. Raub, team member and deputy director, NIH, but the project was an informal survey and, ultimately, the recommendations the team made were based on a consensus of its members' best judgments.[14]

Virtually every recommendation made by the study team has been implemented.[14] Among these was the suggestion that guidelines on conflicts of interest and a formal appeals system for the reconsideration of rejected proposals be established. In addition, as part of the appeals procedure, the team suggested that specific criteria be established for reevaluating proposals and that an independent ombudsman be appointed to adjudicate disputes between the NIH and applicants. A change in NIH procedure that was recently instituted is the creation of two programs allowing the life of a grant to be extended for up to 10 years under certain very limited circumstances.[15]

In connection with the report by the NIH study team, Jonathan Cole suggests that a fruitful area for research would be a rigorous comparison of the NIH study-section approach to peer review with the individual approach used by the NSF. He says that "panels can evaluate the *relative* strengths of a set of proposals, but, in fact, each panel member, while voting on all, actually only reads a few. This leads potentially to an artificial consensus, where a couple of strong characters on the panel dominate the decision-making process."[6]

Studies of Scholars' Attitudes Toward Peer Review

Sociologist Gilbert W. Gillespie, Cornell University; Daryl E. Chubin, director, Technology and Science Policy Program, Georgia Institute of Technology, Atlanta; and physician George M. Kurzon studied the factors that help shape applicants' attitudes toward the system.[16] The authors expected to find that those who experienced success in obtaining funding would tend to be satisfied with the *status quo* and that those who failed to obtain funding would tend to blame the system.

Gillespie and colleagues sent a three-page, 19-item questionnaire to 719 researchers whose proposals had been approved or rejected by the National Cancer Institute of the NIH in 1980 and 1981. The questionnaire stated that those who did not return the survey would be assumed to be satisfied with peer review, so the authors find it noteworthy that 336 (47 percent) responded—although they do not presume that satisfaction with the system was the only reason for nonresponse. It is also interesting, the authors said, that 205 (61 percent) of the responses came from scholars whose proposals had been funded, since they expected a heavier response from scholars who had been denied funding.[16] Because the questionnaire was sent to researchers who had recently submitted proposals for review, it could not measure the attitudes of those whose discontent with the system had led them to give up submitting proposals.

As the authors expected, previous success in obtaining funding was found to be inversely proportional to a desire to change the system. Gillespie and colleagues also found that those who have been unsuccessful until very recently in obtaining funding tended to support the current process, while those who had been successful in the past but who had recently been denied funding tended to favor modifications to the system. The authors also concluded that several complaints about the peer-review system reflected a surprising ignorance of the procedures governing the operation of the system. For instance, those who believe that cronyism or old-boy networks control the process fail to take into account the limited time that an individual may serve in a review group and the NIH's strict requirements concerning the makeup of such groups, which ensure a balanced cross section of scientists that changes constantly.[16] Jonathan Cole points out, however, that the choice of a given individual reviewer from among a number of roughly comparable candidates "can be a function of social and intellectual ties with study-section members."[6]

Flaws in the System?

There may be instances in which peer review operates with unintended blind spots or unsuspected inefficiency. Alan L. Porter, School of Industrial and Systems Engineering, and Frederick A. Rossini, School of Social Sciences, Georgia Institute of Technology, studied the fate of proposals that "fall between the cracks" of the NSF's disciplinary programs.[17] After analyzing 257 reviews received by 38 approved, cross-disciplinary proposals in five different subject areas, they found that reviewer decisions were more favorable when the proposal fell within the reviewer's own discipline. In discussing this finding, the authors found it reasonable "for a reviewer of proposed research to favor that which is more familiar.... In such a case, one is apt to understand better what is planned; one may know the researchers personally or by reputation, and hence appreciate their expertise; and one can feel more secure in making strong recommendations."[17] Porter and Rossini conclude that interdisciplinary research proposals should not be reviewed in the same way as disciplinary projects.

A study by Anthony S. Russell, associate professor of medicine, and Michael Grace, both at the University of Alberta, Edmonton, and Bonnie D. Thorn, director of finance and administration, Arthritis Society, Toronto, Canada, supports the widespread belief that the peer-review process is unnecessarily long and complex.[18] Russell and colleagues examined 113 grant applications to the Arthritis Society, a national voluntary health organization, to determine whether there were any substantial differences between the initial assessment each proposal received in-house and the detailed, out-of-house review that followed. They found that in-depth reviews had little impact on the original rating, implying that review procedures that operate in a similar, two-tiered fashion could be greatly streamlined.[18] And in fact, in an analysis of nearly 1,400 reviews of about 200 NSF proposals, David Klahr, Carnegie-Mellon University,

Pittsburgh, Pennsylvania, found that independent mail reviewers had little impact on the final rating given to a proposal by panel reviewers.[19]

Suggestions for Further Change and Improvement

As I mentioned earlier, both the NSF and the NIH have instituted changes in their review procedures over the last few years. Nevertheless, there are plenty of suggestions for changing the system. Unfortunately, since so little empirical data exist, most of these suggestions are little more than remedies for *perceived* ills. It is hard to know which ones are worth implementing without further research.

One suggested change concerns the time and effort consumed by writing proposals and filling out forms. Typical of many scientists' feelings is a remark attributed to Nobel laureate biochemist Albert Szent-Györgyi (1893-1986). In an article published in *Chemical & Engineering News*, science journalist Howard J. Sanders reports that Szent-Györgyi once remarked that writing grant proposals filled his "scientific life with agony."[20] Rosalyn S. Yalow, Veterans Administration, New York, the 1977 Nobel laureate in physiology or medicine, suggests that researchers of demonstrated ability should not have to go through the process of making a formal application year after year for the renewal of funding.[21] Instead, they should receive a constant level of funding that is renewable every three years, subject to review of their progress.[22]

Rustum Roy, director, Science, Technology and Society Program, Pennsylvania State University, University Park, also wants research funding to be based on an investigator's performance.[23] But in a departure from other scholars' suggestions, he proposes a formula, "based on three kinds of post-hoc peer review,"[24] on which to base grants to individuals, university departments (or research units of a similar size), and institutions.[23,25,26] Roy claims his peer-review formula would eliminate the subjective elements of allocating grant money and does not tie funds to specific projects; instead, money would be administered at the departmental level and would be distributed based on a researcher's past performance, rather than on future promise (with allowances to be made for new or young investigators without track records).[26] Henry R. Hirsch, Department of Physiology and Biophysics, College of Medicine, University of Kentucky, Lexington, also proposed that all active faculty members ought to receive funding, varying to reflect the administration's judgment concerning "the costs and merits of different kinds of research."[27]

Roy's proposal met with considerable individual criticism. In a number of letters written in reply to his original editorial in *Science*,[23] various scientists expressed misgivings about jettisoning the "informed judgment"[28] and the concern with quality that they feel are intrinsic to peer review in its present form.[29,30] But in his reply, Roy says these objections assume that peer review "is in some mysterious way linked with the progress of science" and that the process can accurately predict the quality of research not yet performed. Roy states that both claims are totally unsupported.[31]

Another funding alternative to peer review, supported by a "small but vocal number of scientists," as Sanders puts it,[20] involves block grants, a system common throughout Europe,[3] in which funds are awarded to a research institution for allocation as it sees fit. The money would not go directly to an individual; instead, distribution would be determined by department heads or administrative officials. But Sanders notes that most US scientists strongly oppose a block-grant system, in the belief that a department head or administrative official or committee is less qualified to decide how to allocate research funds than an expert peer-review group.[20] Moreover, according to Joshua Lederberg, president, Rockefeller University, a block-grant system

would merely substitute "the politics of the institutions for the politics of the review committees."[32] And the people who make the funding decisions not only won't be anonymous to those in need of funds, they will have to live and work with them daily, and thus, as Sanders writes, "are less apt to make their choices impartially."[20]

Some scientists also question the underlying assumption of the present peer-review system: that only experts from an applicant's field or a closely allied discipline are qualified to judge that research proposal. David Apirion, Department of Microbiology and Immunology, Washington University School of Medicine, St. Louis, Missouri, suggests the creation of a class of professional, salaried science reviewers to replace peer review.[33] As Apirion puts it, "In all other branches of human creative enterprise [such] as literature, music, sculpture, etc., the producers of new works as well as the performers of new and old works are often judged by a special class of persons, reviewers and critics, who are seldom actively involved in the expansion of the particular discipline that they are entrusted to judge and evaluate."[33]

Pressures on the Peer-Review System

Several authors made observations concerning peer review that bear emphasizing. Yalow pointed out that there is a certain deadening effect—or dishonesty—inherent in trying to explain or justify research that has yet to be done; if your project is so low-risk that you already know what you expect to find, Yalow asks, then how original or important can it be?[21,22] Daniel H. Osmond, University of Toronto, notes that there may be a certain amount of pressure, once funding is approved, to "groom" research results to fit the expectations of the granting agency.[34] Perhaps the biggest problem with peer review, however, isn't really a problem with peer review at all, but rather with the amount of funding available. In the "golden years" of the 1950s and 1960s, money for research was relatively plentiful and granting agencies generous; now, with money tight and with so many applicants, even deserving projects are sometimes denied funding.[35] As Lederberg says, "When there's not enough [money] to go around, some people are inevitably hurt—sometimes arbitrarily and unfairly."[32] Frustration with such decisions carries over to the system by which the decisions are rendered.

Obviously, the process of peer review grinds on in spite of such troubling issues. There was a consensus of views expressed by scientists interviewed for Sanders's wide-ranging special report. In spite of all the complaints and all the faults hinted at, peer review is still considered the best method by which society places its bets on the most fruitful research.[20] Yet the credibility of peer review in the eyes of both the public and the scientific community is threatened by the activities of those who lobby Congress directly for funds. Richard C. Atkinson, former director, NSF, and currently chancellor, University of California, San Diego, and physicist William A. Blanpied, currently international studies specialist at the NSF, warn that the abandonment of peer review might reduce science to just another special-interest group, with funds being allocated based on political acumen rather than on a consensus of what best serves the advancement of scientific knowledge.[35] To prevent more institutions from joining those that have already abandoned the system, further changes in peer review may be necessary. But we should not confuse the forest with the trees. Without a strong peer-review system, albeit constantly reexamined, science might become tentative and inefficient.

* * * * *

My thanks to Stephen A. Bonaduce and Terri Freedman for their help in the preparation of this essay.

REFERENCES

1. **Garfield E.** Refereeing and peer review. Part 1. Opinion and conjecture on the effectiveness of refereeing. *Current Contents* (31):3-11, 4 August 1986.
2. ----------. Refereeing and peer review. Part 2. The research on refereeing and alternatives to the present system. *Current Contents* (32):3-12, 11 August 1986.
3. ----------. Refereeing and peer review. Part 3. How the peer review of research-grant proposals works and what scientists say about it. *Current Contents* (4):3-8, 26 January 1987.
4. **Cole S, Rubin L & Cole J R.** *Peer review in the National Science Foundation: phase one of a study.* Washington, DC: National Academy of Sciences, 1978. 193 p.
5. **Cole J R, Cole S & the Committee on Science and Public Policy, National Academy of Sciences.** *Peer review in the National Science Foundation: phase two of a study.* Washington, DC: National Academy Press, 1981. 106 p.
6. **Cole J R.** Personal communication. 13 November 1986.
7. **Cole S, Cole J R & Simon G A.** Chance and consensus in peer review. *Science* 214:881-6, 1981.
8. **Zuckerman H & Merton R K.** Patterns of evaluation in science: institutionalisation, structure and functions of the referee system. *Minerva* 9:66-100, 1971. [Reprinted as: Institutionalized patterns of evaluation in science. (Merton R K.) *The sociology of science.* Chicago, IL: University of Chicago Press, 1973. p. 460-96.]
9. **Peters D P & Ceci S J.** Peer-review practices of psychological journals: the fate of published articles, submitted again. *Behav. Brain Sci.* 5:187-95, 1982.
10. **Whitehurst G J.** Interrater agreement for journal manuscript reviews. *Amer. Psychol.* 39:22-8, 1984.
11. **Boddé T.** Evaluation of peer review draws mixed reactions. *BioScience* 32:10-2, 1982.
12. **Kirschstein R L, Akers R P, Brooks G T, Fretts C A, Gary N D, Goldwater W H, Green J G, Solowey M, Kaufman A A, Raub W F, Russell G F, Riseberg R J, Schiaffino S S & Wilson K S.** *Grants peer review: report to the director, NIH. Phase I.* Washington, DC: NIH, 1976. 226 p.
13. NIH grants peer review study team: establishment. (FR Doc. 75-23368). *Fed. Reg.* 40:40870, 1975.
14. **Raub W F.** Personal communication. 12 December 1986.
15. **Culliton B J.** NIH starts new grants program. *Science* 232:566, 1986.
16. **Gillespie G W, Chubin D E & Kurzon G M.** Experience with NIH peer review: researchers' cynicism and desire for change. *Sci. Technol. Hum. Val.* 10(3):44-54, 1985.
17. **Porter A L & Rossini F A.** Peer review of interdisciplinary research proposals. *Sci. Technol. Hum. Val.* 10(3):33-8, 1985.
18. **Russell A S, Thorn B D & Grace M.** Peer review: a simplified approach. *J. Rheumatol.* 10:479-81, 1983.
19. **Klahr D.** Insiders, outsiders, and efficiency in a National Science Foundation panel. *Amer. Psychol.* 40:148-54, 1985.
20. **Sanders H J.** Peer review. How well is it working? *Chem. Eng. News* 60(11):32-43, 1982.
21. **Yalow R S.** Peer review and scientific revolutions. *Biol. Psychiat.* 21:1-2, 1986.
22. ----------. Peer review: some suggestions. *Chem. Eng. News* 57(40):5, 1979.
23. **Roy R.** An alternative funding mechanism. *Science* 211:1377, 1981.
24. ----------. Personal communication. 29 November 1986.
25. ----------. Funding science: the *real* defects of peer review and an alternative to it. *Sci. Technol. Hum. Val.* 10(3):73-81, 1985.
26. ----------. Alternatives to review by peers: a contribution to the theory of scientific choice. *Minerva* 22:316-26, 1984.
27. **Hirsch H R.** *A proposal for per capita distribution of research funds with administrative flexibility.* 1983. 5 p. (Unpublished paper.)
28. **Liebman J C.** Letter to editor. (Alternative to peer review?) *Science* 212:1336, 1981.
29. **McCreery R L.** Letter to editor. (Alternative to peer review?) *Science* 212:1336, 1981.
30. **Kalt M R.** Letter to editor. (Alternative to peer review?) *Science* 212:1336-7, 1981.
31. **Roy R.** Letter to editor. (Alternative to peer review?) *Science* 212:1338-9, 1981.
32. **Lederberg J.** Personal communication. 17 August 1986.
33. **Apirion D.** Letter to editor. (Research funding and the peer-review system.) *Fed. Proc.* 38:2649-50, 1979.
34. **Osmond D H.** Malice's Wonderland: research funding and peer review. *J. Neurobiol.* 14:95-112, 1983.
35. **Atkinson R C & Blanpied W A.** Peer review and the public interest. *Issues Sci. Technol.* 1(4):101-14, 1985.