# The Controversy over Peer Review

## Recent studies of the peer review system show that its critics have yet to make their case.

Thane Gustafson

One of the most striking features of federal support of fundamental science is heavy reliance on scientists outside the government and on the decentralized decision-making processes of the scientific community. Much of the overall direction of federal support for fundamental science depends on the individual decisions of working scientists about which fields to enter and which lines of investigation to pursue. These decisions are communicated to government agencies through requests for funds for project support and through the opinions of nongovernment scientists who advise the government at many different levels. This advice is principally of two sorts: evaluations of requests for project support for their scientific merit and relevance to agency goals; and expressions of opinion as to which fields and lines of investigation are especially worthy of federal support.

The outstanding example of such decentralized mechanisms in federal science policy is "peer review," used by the National Institutes of Health and the National Science Foundation. In its purest form, peer review relies on the flow of unsolicited grant applications from individual researchers ("proposal pressure") (*1*) as the funding agency's principal guide for the planning of research support, and uses committees of nongovernment scientists as the main instrument of proposal evaluation and policy assessment (*2*).

In practice, peer review procedures and patterns of proposal vary widely from one government body to another and even within a single agency. In some agencies peer review is elaborate, proposals are largely unsolicited, and the influence of advisory committees is great; in others peer review is cursory, the influence of agency staff outweighs that of outside advisers, and research proposals originate in response to requests from the agency.

The author is an assistant professor of government at Harvard University, Cambridge, Massachusetts 02138, and a member of the seminar on Science, Technology, and Public Policy of the John F. Kennedy School of Government, Harvard University.

The most influential advisory committees and the most highly developed system of peer review are to be found in the NIH, which funds 75 percent of the biomedical research performed in American medical schools as well as more than 40 percent of all university research. The NIH peer review mechanism has two parts. Initial review of proposals for research grants is delegated to approximately 50 study sections composed entirely of nongovernment scientists. They are organized by scientific discipline and rank grant applications according to their scientific merit. The recommendations of the study sections are then forwarded to the pertinent NIH institute. By law, each institute must maintain an advisory council of nongovernment scientists and informed laymen. The advisory council, by working its way down the project ranking as established by the study sections, decides how far the institute's available funds can extend (thus eliminating many projects approved by the study sections). In addition, the councils may raise the priority of projects that are considered to be especially relevant to the institute's objectives. Without the prior approval of their council, institute officials cannot make research awards.

The result is a unique type of dual review which ranks projects at the first stage according to their scientific merit, and at the second stage according to their relevance to the institute's objectives as well as the nation's (the latter is reflected in the size of each institute's appropriations).

Thus, peer review, even in its purest form, is far from a device for giving a blank check to nongovernment researchers. Its interest as an administrative mechanism lies in the fact that it combines the intrinsic criteria of the scientists with the extrinsic criteria of government agencies, yet in such a way that the two types of criteria remain distinct.

In addition to two-stage review, which is unique to NIH, two further types of peer review that are widespread in the federal government are single-stage review and mail review. Single-stage review is used by certain divisions of the NSF (*3*). The initial evaluation of proposals is done by NSF program directors, who make a preliminary ranking. Although advisory panels of outside experts exist for each NSF program, their role is confined mainly to a review of cases considered borderline by the program director. The outside experts do not assign priority scores, and the final assignment of priorities is left to the program director. Outside scientists are not without influence, however; in some NSF programs the influence of the panels is greater than that of the program director (*4*).

Mail review, which is also used in NSF as well as in other federal agencies, is the form of peer review that gives the greatest role to the program director. Project proposals are sent by mail to outside reviewers around the country, but the opinions of the reviews are not binding on the program director. Even in this form of peer review, however, the influence of the program directors is limited by the fact that all large projects are subject to the approval of the National Science Board (*5*).

Peer review relies heavily on the discretion, integrity, and good judgment of working scientists and agency staff. The opinions of reviewers are confidential (*6*). Accountability is maintained (at least in principle) through the review of funding decisions by agency superiors and through periodic oversight by congressional committees.

Peer review has its critics, however. The system has undergone recurrent attacks from several sources and is currently under fire once again. A few conservative members of the House of Representatives have recently attacked the confidentiality of peer review in NSF and have questioned the integrity of its program officers. For example, Representative R. E. Bauman of Maryland denounced the peer review system in bitter terms on the floor of the House (*7*):

> I suggest that there is a need for revision of the basic system by which . . . research grants are made. They are handed out in an unregulated and secretive manner known as the "peer review system." This system allows cronies to get together and finance their pet projects, where grant application writing has become an art and where many people are not devoting themselves to basic research needs but rather to feathering their own nests.

Another congressman, Representative John Conlan of Arizona, has charged that in some instances the peer review system has been deliberately tampered with by program officers of NSF. Specifically, Conlan stated that in two instances the comments of outside reviewers had been misrepresented by NSF program directors. Representative Conlan further objected to the secrecy which has traditionally shrouded the identity of outside reviewers and criticized the refusal of NSF officials to make their reviews public (*8*).

However, these congressmen are not the only figures in government who are skeptical about peer review. The Office of Management and Budget in 1973 issued a sharply worded critique of the peer review system in NIH (*9*). Criticism also comes from outside the government. Representatives of small liberal arts colleges have complained that the peer review system is biased in favor of the major graduate universities (*10*). Feminist groups have charged that peer review and advisory groups are heavily biased against women scientists (*11*). And, there are protests that peer review groups, in choosing new members, discriminate against younger scientists (*12*).

Despite the dark suspicions of congressional critics, we should begin by noting that integrity is not the main issue; it is most unlikely that the integrity of external advisers and agency staff will be found wanting. In NIH study sections, for example, great care is taken to avoid conflicts of interest. Grant applications of study section members are never assessed by the section to which the applicant belongs. Study section members who belong to the same institution as the applicant whose proposal is under consideration are required to absent themselves during the debate and final vote (*13*). Recent statistical studies have shown that, in NSF, applications from top-ranked universities do not receive preferential treatment in the assignment of reviewers; no correlation exists between the rank of an applicant's department and that of the reviewer's (*14*).

What is at issue depends on the position of the critic. One can discern three principal preoccupations.

The first concerns the management of the peer review system, especially its integrity and its accountability to political authority. The second is directed at the principle which governs the choice of research proposals and of advisors—whether it should be that of merit, equity, or relevance to national goals. The third preoccupation focuses on the functions and limits of peer review. Under this category come questions as the extent to which new techniques of programming and evaluation, such as benefit-cost analysis, should be substituted for peer review and proposal pressure as policy-making instruments, and whether peer review should serve chiefly as a one-way signaling mechanism from the scientific community to the funding agencies, or as a two-way channel that includes a feedback link for the communication of agency priorities to the scientific community.

These three concerns—over management, principles, and functions—are present simultaneously in the three most prominent criticisms of the peer review mechanism. They are: (i) that there is favoritism and discrimination in the selection of advisory committee members and in their behavior; (ii) that agency staffs play excessive or improper roles in the selection of research projects; and (iii) that the peer review system is not the most reliable way to coordinate federal funding of basic research with the political goals of the government. My principal purpose is to review the evidence on these issues and to determine, where possible, their impact on the progress of fundamental research.

## Bias in Structure and Behavior of Science Advisory Committees

According to critics, members of external advisory committees are predominantly white, male, and more than 35 years of age, they are drawn disproportionately from prestige graduate universities, they discriminate against younger researchers and untried approaches, and they perpetuate these biases by their long terms of service and by their tendency to maintain an entrenched "old boy network" by nominating their personal acquaintances to succeed them.

Most of the critics share the premise that the selection of committee members and of research projects should be based primarily on professional achievement and scientific merit according to universalistic standards. Their principal concern is that biases in representation will lead to biases in evaluation of the "state of science," that is, of the intellectual opportunities facing the scientific community. Some critics, however, feel that to be representative in their advice, panel members must be representative in their socioeconomic characteristics. This is a popular viewpoint among reformers (*15*). However, it assumes that social characteristics or institutional affiliation influence the panel members' scientific judgment; and, as we shall see, this assumption is not easy to substantiate in the case of science advisory bodies for fundamental science.

Other critics are less concerned about scientific merit than about equity, defined as equality of funding among regions or as adequate representation for minority groups or less prestigious institutions. This is a more fundamental challenge, for it calls into question the concept of universalistic professional standards on which the peer review system rests.

## The Evidence on Representativeness

Membership on scientific advisory groups in the federal government is publicly available in most cases; the names, occupations, and addresses of nearly all the 20,000-odd individuals who advise the government in formal committees are published annually, as required by the Federal Advisory Committee Act of 1973 (*16*). In addition, the composition of NIH and NSF committees has long been made available in annual reports. Recently a major gap has been filled as a result of the decision of the National Science Board to make public the names of the approximately 35,000 reviewers throughout the country who advise NSF on project selection (*17*).

Unfortunately, this abundance of information has not yet been used for systematic studies of the characteristics of scientists serving on peer review committees. Published studies of other types of scientific advisers, if used with caution, can be used to gain an approximate idea of the characteristics of peer review groups. Since peer review groups tend to be more prestigious than most advisory committees but somewhat less so than policy-making advisory bodies, it seems reasonable to speculate that the characteristics of their members may fall somewhere between those of the two types of bodies.

The most thorough analysis now available of the social and professional characteristics of scientific advisers is contained in a study of the National Research Council (NRC) (*12*). The study reports that NRC advisers are an average of 10 years older than the population of scientific doctorate holders in the United States. While 50 percent of the advisers studied were more than 50 years old, only 3 percent were less than 35 (*12*, p. 53). The situation was similar in another agency investigated in the study; for the five senior advisory bodies of the Department of Defense the median age was 50 years (*12*, p. 54). The median age of NSF advisers in 1970 was also in the same range (47 years) (*18*).

National Research Council advisers were predominantly male. Whereas 7 percent of the holders of scientific doctorates in the United States in 1969 were women, 1 percent of the NRC advisers were female. Subsequent studies have shown that similar ratios prevail in NIH; in 1972 women accounted for 2 percent of the advisory body membership (*11*). This pattern is not universal, however; female scientists in NSF make up nearly 8 percent of the advisory group panels (*19*).

Certain regions were found to be better represented in the NRC than others, although the pattern varied with the importance of the advisory committee involved. For the NRC system as a whole, the South Atlantic region was the best represented (*12*). But on 35 policy-related committees the New England, Pacific, and Mid-Atlantic regions were dominant, while the South Atlantic trailed behind. The latter pattern is also characteristic of NSF panels; the New England, Pacific, and Mid-Atlantic states together, while employing about 50 percent of the nation's doctoral scientists and engineers (*20*), account for nearly two-thirds of NSF advisers (*19*). As for NSF reviewers, data for fiscal year 1974 show that California, Massachusetts, Maryland, and Washington, D.C., are represented considerably more than proportionately in relation to their populations of scientists (*14*).

Another charge is that advisory panels are dominated by representatives from prestige universities. Statistics for NSF and NIH do confirm that the top graduate institutions are heavily represented. An American Chemical Society (ACS) study of 11 NIH study sections in biochemistry showed that of 113 institutions represented, 32 contributed more than two-thirds of the advisers, and the eight institutions most heavily represented supplied nearly one-quarter of the advisers (*21*). In NSF the picture is similar: the 18 top institutions, which employed 9 percent of the nation's academic doctorate-holding scientists in 1971, contributed more than a third of the advisers (*22*).

In contrast, there is little evidence that the science advisory apparatus consists of a small number of individuals who spend long years on one committee or rotate from one committee to another. The ACS study of NIH biochemistry study sections found that only seven out of 366 advisers from 1964 to 1973 served twice. This appears to be in part the result of deliberate policy, for NIH's practice is to make no immediate reappointments to its study sections. Similarly, a study of science advisers to the U.S. Public Health Service (which included NIH study sections at the time of the study) showed that over the period from 1957 to 1964 fewer than 15 percent of the advisers served two full terms and fewer than 20 percent served more than one full term. The median seniority of Public Health Service advisers was a little more than 1 year during the period covered by the study (*23*).

If periods of service are so short, then recruitment must be very large. The study of the Public Health Service concluded that the intake of its panels would eventually include one out of every two active biomedical researchers in the United States. The ACS study of NIH biochemistry panels, which used a different approach, came to a similar conclusion.

In contrast to the rank and file of science advisers, there is evidence that the consistently influential advisers that serve the most important committees come from a smaller pool of individuals and rise in the science advisory hierarchy over a period of several years (*23*). Nevertheless, the dominant picture is one of great mobility among science advisers, which is consistent both with the extraordinary fluidity of the active research population in the American scientific community and with the high turnover rates in the professional staffs of science-funding agencies.

To sum up the evidence I have reviewed so far, the available studies confirm that the membership of science advisory committees does not strictly reflect the composition of the doctorate-holding scientific community, at least in some overall social and institutional features. The next question is what to make of this point.

First, there may be more appropriate criteria by which to judge representativeness. For example, although the prestige universities may appear to be overrepresented in comparison to the number of Ph.D.'s they employ, they seem underrepresented when one looks at the number of Ph.D.'s they train since the top 20 doctorate-granting institutions account for more than two-thirds of the output of doctorate-holding scientists.

Defenders of the peer review system assert that the selection of peer reviewers mirrors scholarly achievement. There is some evidence for this; according to a recent statement by Deputy Director of NSF Richard Atkinson, the distribution of NSF reviewers by state approximates the geographical distribution of scholarly publications. By that measure, the only overrepresented states are New York and the District of Columbia, and in the latter case the explanation is that many NSF proposals are reviewed by officials of other federal agencies, so as to ensure coordination (*14*, p. 11, figure 3).

A better indicator of merit would be the impact of research publications on the scientific community. Sociologist Hagstrom of the University of Wisconsin, by counting the frequency with which the publications of scientists from prestige institutions are cited, has concluded that the impact of those institutions is roughly in line with their representation on science advisory committees in Washington (*24*).

Most important, the representativeness of committee membership is significant only if we accept the premise of the advocates of representative reviewship that skewed membership leads to a skewed pattern of awards. I now turn to the evidence on distribution of research awards among institutions and among individuals in order to determine whether this is true.

## Patterns of Awards

Raw data on patterns of funding indicate a strong advantage for prestige institutions. In NIH, according to a recent memorandum by the Office of Management and Budget, applicants from ten institutions accounted for 46 percent of all grant funds in 1971, and their share then was greater than it had been 5 years before (*9*, p. 9). Similar patterns prevail in NSF. In 1974, research grant awards to the top 20 institutions represented one-third of total NSF obligations for all programs (*24*, p. 15). In the field of chemistry alone, according to an in-house study, the top 20 departments, employing 18 percent of the country's graduate chemistry faculty, received 49 percent of the NSF chemistry awards in 1973, an increase of 5 percent over 1972 figures (*25*). Moreover, the success rate of the top 20 departments was 53 percent, compared to 30 percent for all other departments (*26*).

However, the issue of concentration in funding is more complicated than it appears at first glance. For example, concentration does not always work to the advantage of the prestige institutions. A study of the allocation of NSF project grants in metallurgy and materials research shows that they are clustered among the middle-ranked institutions rather than the top-ranked ones (*27, 28*). In addition, the more significant pattern of concentration may not be among institutions, but within them. Several studies have found that when one takes into account the size of each institution, and particularly the number of faculty members capable of performing or supervising graduate-level research, the funding distribution among institutions becomes more nearly uniform. Thus, on a per capita basis, most of the inequality is among faculty within a single department. Whether the department is rated high or low, approximately 70 percent of the outside funding is received by the more productive half of the faculty, the half which receives more than 90 percent of the citations in the scientific literature (*29*).

A related issue is the extent to which concentration increases in times of budgetary crisis. Critics complain that in hard times the representatives of prestige institutions act to protect their own. Data from NIH provide some evidence on this point; from 1967 to 1970, when the purchasing power of NIH funding declined sharply, the number of institutions funded declined from 330 to 277 (*30*). However, during a similar period (1966 to 1972) the total pool of NIH grant recipients declined much less—from 10,250 to 8,150 (*31*). The implication is that the concentration of NIH funding in the prestige institutions, even when reckoned on a per capita basis, increased during this period, a point which receives confirmation from Office of Management and Budget data (*9*). Since the early 1970's the number of institutions funded by NIH has expanded again, although much of the increase is due to the new prosperity of the National Cancer Institute (NCI). In 1974 alone the NCI allocated its funds among 367 institutions (*32*).

In times of reduced funding, how new proposals and younger researchers fare in comparison with scientists applying for renewals is of interest. At first glance, hard times seem harder for the researcher seeking his first grant. From 1967 to 1970 the share of NIH research grants awarded to new projects declined from 23 percent to 16 percent (*33*). Since that time the downward trend in the percentage of NIH grants awarded to new applicants has continued for all NIH institutes except the NCI and the National Heart and Lung Institute (*34*). Particularly in the NCI the increased flow of money since 1971 has had a dramatically different effect. Despite the enormous influx of new-project applications (proposals to NCI now account for roughly half of all applications to NIH), the success rate of new principal investigators (that is, the ratio of awards to applications) has risen from 30 percent in 1970 to 54 percent in 1974 (*34*, p. 479).

Even in bad times, however, younger researchers among the new applicants have not been neglected. For NIH as a whole in 1971, applications of young scientists (age 36 and less) for traditional project grants were approved (that is, recommended by the study sections) at a higher rate than those of older scientists (*34*, p. 48). Funding data for 1972 and 1973 for NCI give the same picture (*31*, pp. 240 and 241).

> Of those investigators 35 years of age and under applying for research support, 52% were funded ... in fiscal year 1972, and 35% were funded ... in fiscal year 1973. Of those applicants older than 35, only 32% were funded ... in 1972 and 25% ... in 1973.

A similar pattern has been observed in internal studies of NSF funding of younger investigators (*35*).

Once again, we must question the meaning of these data and then ask whether the observed pattern of awards is unfair or inefficient. For those who value equity over scientific merit, the existing pattern is discriminatory. However, for those whose principal concern is scientific merit, the key question is whether the existing distribution of awards is out of line with the distribution of scientific quality in the country's research institutions. In order to find the answer some reasonable indicator of the quality of research is required (*36*).

One possibility is the response of the scientific community itself to the present pattern of awards, for presumably researchers are good judges of whether NSF and NIH are funding the best science. Until recently most researchers gave the appearance of being satisfied. But this was primarily because, until the late 1960's, most worthy proposals could find support from one federal agency or another. Unsuccessful applicants usually recast their proposals in response to reviewers' comments and tried again. In NSF, for example, the number of rejections was never so great as to warrant a separate review and appeals procedure. In recent years, however, funding for fundamental research has dropped and the number of funding sources has declined. More and more front-rank scientists find themselves the unhappy recipients of "declination" notices from NIH and NSF. The two agencies are now faced with mounting protests from these disgruntled applicants. Consequently, in order to justify the pattern of funding produced by the peer review system, more objective indicators are necessary.

One measure that is attracting increasing attention is the rate at which scientific publications stemming from NSF or NIH grants are cited in the scholarly journals. Although subject to certain distortions, the citation rate does provide a rough indicator of the impact of a scientific work within the scientific community. This makes it possible to cross-check the judgment of peer review groups by comparing it to the reaction of the scientific community as a whole. A study by the Rand Corporation is enlightening (*37*). By surveying the 15-percentile fraction of NIH grants that produced the most heavily cited publications, the study showed that their greater merit was recognized from the beginning by NIH study sections, which awarded that fraction a significantly better priority score than they did to the remaining group. Figures on applications for renewal were also revealing: the most frequently cited researchers received a much higher priority score in applications for renewal than they had in their initial applications. These findings appear to indicate that peer review groups have a good capacity for recognizing potential in a research proposal, and for recognizing outstanding achievement when they see it (*38*).

The Office of Management and Budget's charge that NIH study sections deliberately approve more applications than can be funded in order to make a case for higher appropriations for the agency (*9*) can also be illuminated by these points. Their claim appears to be weakened by the fact that the average priority score awarded to new grant proposals by NIH study sections has remained much the same since 1967 (*37*). This suggests that the study sections are doing much as they have always done; what has changed is the amount of funding available to cover the proposals they approve.

The link between citation rates and peer review scores also provides a means for judging the claim that peer review groups award too much money to prestige institutions. For example, an internal study of NSF awards in chemistry shows a striking correlation between the distribution of citation rates and the allocation of NSF funds. Of NSF awards in chemistry, 80 to 85 percent go to departments whose faculty averaged more than 60 citations per author during the 5-year period ending in 1972; the top four or five departments, which received the lion's share of NSF grants, averaged around 400 citations per author (*39*). If one accepts the premise that citation rates reflect the quality and impact of research results rather than simply the visibility of their authors, then the charges of bias leveled at peer review groups lose much of their force.

On balance, however, these conclusions are still tentative. There has been too little research into the impact of the peer review system on funding patterns and on the quality of the resulting research. At this writing, several major studies are under way. It seems certain that, in view of the growing popularity of citation analysis and the controversy that surrounds the peer review system, the questions addressed in this article will soon be much better understood (*24*).

Up to this point we have concentrated exclusively on the use of outside scientists in the peer review system. However, in many programs and agencies professional staff officers play an important or even dominant role. Let us now examine their influence in peer review.

### Agency Staffs as a Force in Peer Review

Program officers of funding agencies such as NIH and NSF play an important role in the evaluation of proposals. In several important respects the program officers are themselves the peers of the applicants for research funds: they are professional scientists with a background in research; they are often not career employees of their agencies; and their length of service is usually short (*40*).

However, program officers in many cases (the executive secretaries of NIH are an exception) are expected to further the program objectives of their agencies, and these do not necessarily coincide with those of the scientific community. Consequently, the program director may feel torn by competing criteria and he may find himself at odds with the external advisory committee attached to his program.

However, the potential for conflict varies, because the role of the program manager differs from agency to agency and from program to program. In comparison with that in NIH, the peer review process in most other agencies is less formal and sharp, and the influence of agency staff is generally greater.

The question of how much influence agency staff members have in the process of peer review is also relevant. Even where the influence of outside scientists is greater, as in NIH, agency staffers such as executive secretaries of study sections play an important, if discreet, role. In NIH, the staff of the Division of Research Grants has the power to rearrange study sections (*2*, p. 26). The executive secretary appoints the study section chairman and together they recommend replacements for panel members (*41*). When applications first reach the NIH Division of Research Grants, it is an executive secretary, temporarily acting as referral officer, who decides to which institute to assign the application (*42*)—a crucial decision since institutes differ widely in the number of approved projects they can fund. Executive secretaries in NIH decide which study section members shall have primary responsibility for reviewing a given proposal; in NSF the program director selects the reviewers for each proposal. During the processing of applications, the program officers of both agencies are the principal contacts between reviewers and applicants (*43*). In NIH the executive secretary acts as the link between the study section and the next stage of review by interpreting the decisions of the study sections to the advisory council of the institute. In NIH advisory councils it is the professional program manager who has the responsibility for recommending departures from the priority ranking established by the study sections whenever he deems a proposal especially relevant to the institute's mission (*2*, p. 28). Finally, the program directors of both NIH and NSF have final authority in funding decisions, although they will only very exceptionally make awards to proposals that have been disapproved by reviewers or panels.

Thus, even in the programs in which external peer review panels have the greatest sway, there appears to be the opportunity for the agency staff to influence the process of proposal evaluation by shaping the agenda, channeling the flow of information to and from the outside advisers, or actually altering or overriding their decisions. In NIH this influence is discreet and informal while in NSF it is usually more important than that of the external advisers. In both agencies the importance of the professional staff appears to be growing.

Of late, program officers find themselves increasingly involved in controversy. There are several reasons for this: funds are short and researchers are anxious; agency leaders are under pressure to pursue a wide array of social objectives; reformers call for openness in administrative procedures and for direct accountability to the public; and in mistrustful times the good faith of all administrators is under suspicion. As the chief points of contact between scientists and agency leaders, the program officers are caught in the middle.

Recent congressional attacks on program officers have focused mainly on their integrity. For example, Representative Conlan has charged (*8*, p. 4):

> It is common knowledge in the scientific community that NSF program managers can get whatever answer they want out of the peer review system to justify their decision to reject or fund a particular proposal. Since program managers soon learn, like college students, which professor is good for an easy "A" and which can be counted on for an almost certain "C" or "D," it's no trick to rig the system.

Such charges are extreme. The real issue, just as in the case of the external advisors, is not the program officers' integrity, but rather their accountability. According to the congressional critics, NSF section heads and division chiefs do not have time to give more than a cursory look at program officers' recommendation (*8*, p. 7), and reviewers' reports are not made public; therefore, in this view there is no effective mechanism to check up on the decisions of program officers. The issue of accountability arises also in NIH, since several institutes have shifted their emphasis from grants (which require the traditional two-stage peer review) to contracts, in which the major decisions are made by staff officers (*44*).

Critics in Congress propose to remedy the problem of accountability both by making reviewers' names and comments public and by making information available to Congress about individual research projects in advance of final approval (*45*).

Such remedies would make a major departure from the accepted administrative doctrine, which has evolved out of long experience in dealing with technical programs in government. Three points are central to the doctrine: (i) in the assessment of scientific merit, a certain degree of confidentiality is required in order to encourage candor on the part of reviewers and to guard against the danger of plagiarism; (ii) congressional oversight is best conducted after the fact, not before, and should be devoted to assessment of overall policy rather than detailed review of individual projects; and (iii) accountability is best achieved by focusing responsibility on visible and answerable superiors, not on their subordinates. The measures recently proposed in Congress would discourage candor or even participation by outside reviewers; they would involve congressmen directly in "legislative service" in support of scientists from their districts; and they would place the political spotlight directly on the program officers rather than on their policy-making superiors.

These dangers were clearly present in the minds of several members of the House Subcommittee on Science, Research, and Technology at congressional hearings on peer review this summer.

If the Conlan-Bauman proposals are not adopted—and it seems unlikely that they will be—the question is raised as to what reforms of the peer review system are needed. To answer this question it is necessary to look carefully at the purposes of the peer review system and its inherent limitations.

### Purposes and Limitations of Peer Review

Peer review and proposal pressure are signaling mechanisms for providing funding agencies with information about the possibilities of science. The decentralized systems of rewards and communications of the scientific community supply external advisory panels with the necessary knowledge and cues for appraising proposals and recommending new policies. Most important, the peer review system is based on the assumption that the flow of proposals to funding agencies will provide a reasonably accurate sampling of the universe of research opportunities arising in all fields of science. The objective of any reform should be to safeguard the accuracy of this signaling device.

Recent developments in the support of fundamental science may make this objective difficult to attain. The bulk of funding for basic research is now concentrated in four federal agencies, all of which are under growing pressure to promote "targeted" research programs. Such concentration creates new networks of communications and rewards alongside the traditional ones, altering the "market" mechanisms of the scientific community. There is the danger that proposal pressure, filtered through peer review, will no longer provide a reliable guide to the best science. An illustration of this problem is the remarkable flood of proposals to the NCI since 1971, or the recent surge of energy-related applications to NSF. In both cases researchers appear to be responding to the news of increased funding. The question is whether their response really matches the actual distribution of intellectual opportunities in science.

Distortions in the signal generated by the proposal pressure–peer review system can arise from four sources:

1) The structure and mechanism of the peer review system itself can distort the signals reaching the funding agency from the scientific community, thus producing a false impression of the intellectual opportunities available in each field (selective channeling).

2) The application of extrinsic criteria and funding constraints by agency staff, by encouraging some directions of research and neglecting others, may sharply affect the distribution and content of proposals submitted in the next time period (indirect feedback).

3) Agency staff may influence the generation of proposals through direct publicity, requests for proposals, seminars, and other. External advisors may carry back to their departments information on new funding priorities in Washington or informal lore on which proposals will be best received (direct feedback).

4) The practice of funding by project rather than by institution may cause a pattern of over response by researchers to short-term swings in agency priorities, because of the necessity of applying for new funding at frequent intervals.

Some observers of the peer review system, particularly in government, would say that these features of the peer review system are virtues, not defects. Items 2 through 4 above stimulate a fast response by the scientific community to national needs as they are perceived in Washington. Nevertheless, the signaling function of proposal pressure and peer review is vital. Without an accurate picture of the state of science, funding agencies run the risk of pouring public money into fields whose titles are politically appealing but which are not scientifically ripe for major advances, and of neglecting others in which gains might be made, but which appear esoteric or unintelligible to the layman.

Ways of limiting distortions in the signaling function must be found. One method is to move away from the traditional project grant and make greater use of funding by institution, thus leaving project selection to local institutional management as is currently done in national laboratories and coherent area programs. Funding takes the form of subvention for the institution as a whole, or a particular research center or broad field within an institution, and leaves allocation of resources as a responsibility of the local management, to be justified only after the fact. High scientific merit is maintained through the use of outside visiting committees, overseeing corporations, or various kinds of internal peer review. Institutional funding is already widely used in programs that require an interdisciplinary, well-coordinated approach.

An institutional approach helps to reduce several of the sources of distortion discussed above. It reduces the risks bearing on the individual researcher. When project grants are the sole source of research funds, the funding of a particular proposal nearly becomes a life and death decision for the career of a scientist or the survival of a department. This exacerbates the sensitivity of researchers to news of changing priorities in Washington (*46*). In addition, funding by institution clearly separates decisions on the political relevance of a field of research from those on the scientific merit of individual proposals.

However, many of the institutional funding mechanisms currently in use have been criticized by the scientific community. Many scientists argue that in national centers and area programs scientific merit is not as carefully scrutinized as in universities. An example is the National Center for Atmospheric Research (NCAR), which is funded as a unit by NSF. Until 1973 NCAR scientists were given wide latitude in their selection and execution of research proposals. Peer review took the form of general supervision by a University Consortium. According to a recent evaluation of NCAR, however, the results of this approach were disappointing (*47*). Scientists at the NCAR had few publications to their credit in proportion to the support provided to them; their research was not interdisciplinary as had been advertised in justifications of the NCAR approach; and the work and the quality of its personnel were not as highly regarded among atmospheric scientists as should have been expected, given the amount of funding and of logistic support. According to critics, one of the reasons for the undistinguished performance of the NCAR was a lack of close supervision by the managing consortium.

Another example of the mixed results of funding by institution is the center-grant program of NIH. Research funds in such centers are allocated at the local level. However, critics charge that this system produces "peer reviews at home rather than peer review at a distance," which carries the potential for conflicts of interest and personality (*13*). Similar problems plagued the NCI in the management of its Cancer Chemotherapy National Service Center in the early 1960's (*42*, p. 206).

Thus, while institutional funding may be the answer for certain large and closely coordinated programs, it will not necessarily enable funding agencies to obtain a reliable sampling of the best research opportunities. For most types of fundamental research the traditional project grant, selected by peer review, with overall priority among fields and subfields determined at least in part by proposal pressure, appears to provide the best available guarantee of scientific merit and accurate information. It is important, however, to extend existing safeguards: to choose advisors and agency staff who are representative of the best science; to limit their terms of service; to separate as much as possible the evaluation of scientific merit from that of funding so as to reduce the dependence of researchers on the priorities or biases of any one agency or congressional committee; and finally, to subject the entire system to periodic review and criticism. Given the present—admittedly tentative—state of our knowledge about the impact of the traditional peer review system, there is not a convincing case that the system's defects warrant the risk of sacrificing its virtues.

### References and Notes

1. Most project proposals come from an individual "principal investigator," who represents a small group of junior colleagues and collaborators, including graduate students. The advantage of this system is that it focuses scientific accountability: the "principal investigator" serves as the channel of responsibility between the project team and the sponsor.
2. J. G. Wirt, A. J. Lieberman, R. E. Levien, *R & D Management: Methods Used by Federal Agencies* (Rand Corporation R-1156-HEW, January 1974).
3. For background information on the various types of proposal review, see Wirt *et al* (*2*). Specific details on the roles of program directors, advisory panels, and outside reviewers in each NSF program may be found in *Peer Review and Proposal Evaluation: Staff Study* (National Science Foundation, Washington, D.C., June 1975).
4. An interesting question for future investigation is the mechanism of the distribution of power between the program director and his advisory panel. In some cases, it may be a matter of tradition; in other cases, power is wrested from the advisory panels by the section head. For example, before 1966 the advisory panel for chemistry is said to have played a very influential part in NSF chemistry awards. With the arrival of a new section head, however, the panel's powers were sharply reduced. It is interesting to speculate, incidentally, on the connection this fact may have with the remarkable subsequent stability of the chemistry section staff. Turnover among program directors could conceivably be inversely related to the power they enjoy.
5. Projects involving more than $500,000 during the first year or more than $1,000,000 over the lifetime of the project must be approved by the National Science Board (NSB). However, the NSB rarely rejects a basic research project after it has received a favorable recommendation from the NSF staff; such rejections, when they occur, are based on considerations of policy, not of scientific merit. By contrast, the NSB has never succeeded in reinstating any project rejected by the top level staff of NSF, even though the project may have been energetically recommended by the program officers at the disciplinary level.

6. Paraphrases of the opinions of reviewers may be made available to applicants upon request. The NSB is considering whether to release the names of reviewers under some circumstances.
7. *Congressional Record*, 24 June 1975, p. H6015.
8. Statement by Representative John B. Conlan (R-Ariz.) before the Subcommittee on Science, Research, and Technology of the U.S. House Committee on Science and Technology, 22 July 1975.
9. Office of Management and Budget, *Drug Research Reports*, **16** (No. 22), 4 (1973).
10. See, for example, the statement by J. L. Powell, Provost of Oberlin College, before the Subcommittee on Science, Research, and Technology of the House Committee on Science and Technology, 24 July 1975.
11. D. Shapley, *Science* **175**, 1090 (1972).
12. Charges such as these led to the formation by the National Academy of Sciences of a Committee on the Utilization of Young Scientists and Engineers in Advisory Services to Government. See the report of the Committee, *The Science Committee* (National Academy of Sciences, Washington, D.C., 1972).
13. N. Wade, *Science* **179**, 158 (1973).
14. Statement by R. C. Atkinson, Deputy Director, NSF, before the Subcommittee on Science, Research, and Technology of the House Committee on Science and Technology, 23 July 1975.
15. For a review of the issue of representative bureaucracy, see C. Mosher, *Democracy and the Public Service* (Oxford Univ. Press, New York, 1968); K. J. Meir, *Am. Political Sci. Rev.* **69**, 526 (1975).
16. Public Law 92-463. A crucial exception to the coverage of the act is the many advisory committees of the National Academy of Sciences–National Research Council, which account for nearly half of the total number of advisory committees in the federal government.
17. *Science & Government Report* **5** (No. 13), 8 (1975).
18. The median age of NSF advisers has undergone some interesting changes over the last 20 years. From 51 years in 1954, it dropped to 44 in 1966, and rose again to 47 years in 1970. (These data are drawn from a forthcoming study by N. Mullins *et al.* at Indiana University. I thank Professor Mullins for making some of his findings available to me.)
19. National Science Foundation, *NSF Management Statistics* (prepared by the NSF Administration Directorate, July 1975), p. 19. It should be observed that the apparent sex bias is no doubt a reflection of the age bias. That the percentage of Ph.D.'s awarded to women has been increasing in scientists would reflect the sex discrimination of graduate education 20 or 30 years ago.
20. NSF Administration Directorate, *An Analysis of the Geographical Distribution of NSF Awards as Compared with Other Selected Indicators* (National Science Foundation, Washington, D.C., 1975), p. 1.
21. Findings of the Public Policy Committee, Division of Biological Chemistry, American Chemical Society. I thank E. W. Westhead, professor of biochemistry at the University of Massachusetts (Amherst), for a summary of the committee's data.
22. Preliminary results of a study of the NSF advisory system by N. C. Mullins, *Social Stud. Sci.* (1 July 1975).
23. N. C. Mullins, *Sci. Stud.* **2**, 7 (1972).
24. Hagstrom's work is described in N. Wade, *Science* **188**, 429 (1975). It is conceivable, however, that this finding would not hold for peer review panels; since the latter are highly prestigious, they may draw more than proportionately from prestige institutions.
25. Data presented by NSF to the Committee on Science and Public Policy of the National Academy of Sciences, cited with permission.
26. Atkinson (*14*, pp. 13, 23). The success rate is defined as the ratio of successful applications to total applications.
27. D. Kuhlmann-Wilsdorf, University of Virginia, before the Subcommittee on Science, Research, and Technology of the House Committee on Science and Technology, Washington, D.C., 29 July 1975.
28. Kuhlmann-Wilsdorf's statement has aroused considerable controversy, which is discussed in D. Shapley, *Science* **189**, 622 (1975). The explanation may be that metallurgy-materials research in the prestigious institutions is generously supported through the Materials Research Laboratory program, which is block-funded.
29. Kuhlmann-Wilsdorf (*27*, p. 607). Similar conclusions were reached by the Westheimer Panel in the mid-1960's. See National Academy of Sciences–National Research Council, Committee for the Survey of Chemistry, *Chemistry: Opportunities and Needs* (National Academy of Sciences, Washington, D.C., 1965).
30. T. J. Kennedy, Jr., J. F. Sherman, R. W. Lamont-Havers.
31. C. D. Douglass and J. C. James, *Science* **181**, 241 (1973).
32. F. J. Rauscher, Jr., *ibid.* **189**, 115 (1975).
33. Kennedy *et al.* (*30*, p. 604). In NSF the percentage of funds awarded to new proposals has averaged about 10 percent in recent years. NSF memorandum, *Selection of Physics Research Proposals for NSF Support*.
34. Data drawn from J. T. Kalberer, Jr., *Cancer Res.* **35**, 473 (1975).
35. The NSF studies were performed under C. Falk at the behest of the NSB.
36. In practice, the various categories are not as neat as I have painted them here. The NSF, for example, applies several criteria such as scientific merit primarily and geographic and disciplinary distribution secondarily. As Deputy Atkinson explained in a recent statement, NSF program officers have no set formula for weighing the merit and equity criteria; rather, when choosing between two proposals of roughly equal merit, they decide to a large extent on the basis of geography. But as Atkinson stressed, it is done largely "by intuition" (*14*).
37. G. M. Carter, *Peer-Review, Citations, and Biomedical Research Policy: NIH Grants to Medical School Faculty* (Rand Corporation, R-1583-HEW, December 1974).
38. Evidence in the same direction comes from D. Kuhlmann-Wilsdorf's study. For NSF as a whole, she found a close relation between the size of the NSF grant and the rate at which the applicants' overall work was cited (*27*, p. 7).
39. NSF report to National Academy of Sciences, Committee on Science and Public Policy, 1972. Cited with permission of K. Wilson.
40. To document this point it is instructive to compare the rosters of program officers listed in NSF's Annual Report for 1969 and 1974. Although reorganizations and expansions of disciplinary categories occasionally make comparison difficult, the overall picture that emerges is one of remarkable movement. Sections such as physics, atmospheric sciences, and oceanography underwent almost complete turnover. In fact, the only major research section for which the program officers in 1974 were the same as in 1969 was chemistry.
41. Wirt *et al.* (*2*, p. 60). On the other hand, in NSF all external advisors are appointed by the director.
42. U.S. Office of Science and Technology, *Biomedical Science and its Administration: A Study of the National Institutes of Health* (Government Printing Office, Washington, D.C., 1965), p. 193.
43. G. N. Eaves, *Fed. Proc.* **31** (No. 1), 3 (1972).
44. The NCI, which was the leader in this trend, is beginning to reverse itself. Some programs that were funded by contracts will be funded via special cancer research emphasis grants.
45. Representative Conlan has introduced legislation (HR 9892) that would establish a "peer review office" in NSF, which would maintain a detailed record of applications, reviews, and final actions. See J. Walsh, *Science* **190**, 253 (1975).
46. For interesting illustrations, see G. N. Eaves, *Fed. Proc.* **32** (5), 1541 (1973).
47. N. Wade, *Science* **182**, 36 (1973).
48. This article is based on a survey of American policy on fundamental science conducted for the Joint Working Group on Science Policy under the U.S.-U.S.S.R. Agreement on Cooperation in Science and Technology. A similar survey of Soviet science policy is being conducted simultaneously by the U.S.S.R. Academy of Sciences. I thank H. Brooks, D. K. Price, and W. W. Lowrance for advice and comments.

**NEWS AND COMMENT**

# Energy: Plan to Use Peat as Fuel Stirs Concern in Minnesota

*Minneapolis–St. Paul*—In the northern reaches of Minnesota, in an area carved out by ancient glaciers, lie some of the largest and most desolate peat bogs in the world. Flat, swampy, mosquito-ridden in summer and frigid in winter, the region attracts few, if any, hunters and fishermen, and only an occasional stray scientist or self-styled "swamp freak." Early in this century, an effort to drain the bogs on a massive scale to turn the area into farmland proved a disastrous failure, partly because the drainage proved ineffective and partly because the farmers were unprepared for the intricacies of peat. Later attempts to find some other use for the peat resource—in small-scale agriculture, horticulture, forestry, or whatever—have met with only minimal success.

But now, as a result of the energy crisis, eyes in Minnesota are once more turning northward—this time with visions of tapping a new source of energy for this energy-deficient state. The Minnesota Gas Company (Minnegasco), the state's largest gas-distributing utility, has applied for a long-term lease on some 491 square miles of state-owned land—containing an estimated 200,000 acres (312.5 square miles) of peat—with the announced hope of eventually building a plant that would convert the peat to synthetic natural gas (methane). The Minnesota Energy Agency has received a proposal—submitted by the Midwest Research Institute (MRI) and Rouse S. Farnham, professor of soil science at the University of Minnesota—to investigate the possibilities of burning peat directly as a fuel for municipal power or heating plants. And the Minnesota Department of Natural Resources (DNR) has commissioned a preliminary technical and environmental assessment of peat use for fuel and other purposes. The assessment will be carried out by MRI with the help of a $93,960 grant to the Minnesota DNR from the Upper Great Lakes Regional Commission, a group whose other two members—Wisconsin and Michigan—also have substantial peat resources.

The scale of the gas company's proposal is staggering. The 491-square-mile tract it