ELSEVIER

Technology in Society 27 (2005) 437–451

Technology In Society

www.elsevier.com/locate/techsoc

Assessing the science–society relation: The case of the US National Science Foundation's second merit review criterion

J. Britt Holbrook*

Department of Philosophy and Religion Studies, College of Arts and Sciences, University of North Texas, P.O. Box 310920, Denton, TX 76203-0920, USA

Abstract

The science–society relation exhibits a tension between scientific autonomy and societal control of the direction and scope of scientific research. With the 1997 formulation of two generic merit review criteria for the assessment of National Science Foundation proposals—one for intellectual merit, and a second for 'broader impacts'—this tension between science and society took on a unique institutional expression that has yet to work itself out into a well-accepted balance of complementary interests. This article examines some of the issues associated especially with the second 'broader impacts' criterion.

Keywords: National Science Foundation (NSF); Merit review; Broader impacts, pure, basic and applied research; Philosophy of science policy

1. Introduction

Since its early modern origins, modern natural science has struggled to develop appropriate standards for quality assessment. In particular, one of the main issues has been the extent to which science ought to be judged only on its own terms. The history of science provides extensive literature on the effort to establish science as an

* Tel.: +1 940 565 4048; fax: +1 940 565 4448.
E-mail address: jbrittholbrook@unt.edu.

0160-791X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.techsoc.2005.08.001

autonomous human activity independent especially of religious or political manipulation. The experience of Galileo Galilei (1564–1642) and efforts to break free of church criticism of the heliocentric theory is perhaps the most well-known case in the religious arena. The Communist effort to promote the genetic theories of T.D. Lysenko (1898–1976) is an oft-cited case in the political arena. Such historical lessons taught us to espouse the Enlightenment ideal of a science that exists independently of religious or political influence for the ultimate benefit of society as a whole. From the very beginning, however, questions have occasionally been raised about whether scientific autonomy might be carried so far as to create an imbalance in the science–society relation. The Romantic response to Enlightenment science, for instance, questioned whether science could indeed stand on its own. Are there not times when scientific knowledge distorts lived reality? Is technological power not only a boon but also a danger to human welfare? This debate may be seen in terms of a conflict between advocates of internal and external criteria for evaluating science, with internalists championing autonomy while externalists argue for more societal control over the direction and scope of scientific research. Nevertheless, there are some on the inside (i.e. actual scientists) who recognize the importance of external considerations. Indeed, in the 1960s Alvin Weinberg, the physicist administrator of Oak Ridge National Laboratory, argued that external criteria had a proper and significant role to play in any scheme for assessing science.

In post World War II America, the debate surrounding the formation of the US National Science Foundation (NSF) also reflected this ambivalence: the strong autonomy advocacy of Vannevar Bush's Science—The Endless Frontier (1945) was moderated by the more pragmatic arguments of the Steelman Report (1947), which advocated more limited scientific autonomy in the name of public benefit. Although by the time the NSF was actually created in 1950 many of Bush's specific proposals for its formation were abandoned, his notion of the strong autonomy necessary for basic scientific research was institutionalized with the creation of the NSF and the development of protocols for the internal peer review of proposals to be funded by the federal government. We can credit Bush's rhetorical genius for arguing that a large degree of scientific autonomy was in fact necessary for producing the kind of societal benefits desired; and we can blame either the wishful thinking or woeful logic of many societal decision makers for the assumption that scientific autonomy was, therefore, sufficient for producing societal benefits. Decision makers have often made this assumption, however, with the result that tension persists between advocates of internal criteria of scientific merit and advocates of broader external criteria for assessing science.

With the 1997 formulation of two basic merit review criteria for the assessment of NSF proposals, this tension between science and society, internal autonomy and external evaluation, took on a unique institutional expression that has yet to work itself out into a well-accepted balance of complementary interests. This article examines some of the issues associated especially with the second 'broader impacts' criterion in an effort to contribute to the further evolution of a discussion of a distinctive issue in what may be termed the philosophy of science policy.

2. Background

In 1997 the Naztional Science Board (NSB), the NSF's policy branch, approved two new generic merit review criteria to replace the four that had been in effect since 1981.[1] The two criteria approved in 1997 and currently used to evaluate all NSF proposals are: (1) What is the intellectual merit of the proposed activity? and (2) What are the broader impacts of the proposed activity? It is tempting to assign the 'internal' label to Criterion 1, which is chiefly concerned with scientific merit as judged by scientists, while assigning the 'external' label to Criterion 2, which is concerned with issues of education, infrastructure, diversity, and societal benefit. However, no such simple division of labels will do. For, insofar as both criteria are part of NSF's peer review process, i.e. insofar as both criteria are criteria for scientists to be judged by scientists, Criterion 2 is also an internal criterion. Nevertheless, NSF's emphasis of Criterion 2 introduces what many take to be considerations external to science into the (internal) peer review process.

Ironically, NSB restructured the merit review criteria largely to respond to increased demand for an account of the societal benefits achieved by NSF funded projects. Congress passed the Government Performance Results Act (GPRA) in 1993. GPRA's purpose was to increase the focus of Federal agencies on improving and measuring 'results', which would provide congressional decision makers with the data they require to assess the 'relative effectiveness and efficiency of Federal programs and spending'. The message that 'results' are tied to funding has also been reinforced since President Bush took office by the President's Management Agenda (PMA), as well as the establishment of the Program Assessment Rating Tool (PART), designed specifically to tie GPRA to budget formation. As the 2001 NAPA Report on the newly restructured criteria notes, 'The immediate effect of this restructuring is to make the broader impact and societal objectives more visible—both to the scientific and engineering communities and to Congress' ([1], p. 17). There were, however, additional effects: (1) in 1997 Congress directed NSF to contract with NAPA to review the new criteria, and (2) questions arose within the scientific and engineering communities as to how to interpret and apply the new criteria, especially the second, 'broader impacts' criterion. NSF's new merit review criteria were being challenged on two fronts: on the external front, by members of Congress seeking immediate feedback on the 'results' of the new criteria, and on the internal front, by scientists and engineers who questioned the validity of the new criteria. These challenges focused in particular on Criterion 2.

3. The specter of philosophical issues

The NAPA Report notes that 'many reviewers either ignore Criterion 2 or in some cases regard it as irrelevant in the review of proposals', that many reviewers 'perceive Criterion 1 (scientific merit) and Criterion 2 (broader or societal impact) as in competition with each

[1] For a detailed comparison of the two current criteria to the four 1981 criteria, see [1], p. 6 and pp. 17–18. Hereafter, this report shall be referred to as the 'NAPA Report'.

other', and that many reviewers either 'disregard Criterion 2 altogether or simply merge social value into scientific merit' ([1], p. 13). Among the major recommendations of the NAPA panel is that 'there is a need to improve the conceptual clarity of the objectives of the new criteria as well as the language used in stating them…. This is true of the language of Criterion 2, in particular' ([1], p. 8).

Since the February 2001 NAPA Report, there have been repeated calls for clarification of the second criterion.[2] Such calls for clarification rest on the assumption that if those involved in NSF's merit review process exhibit difficulties with the interpretation and application of the second criterion, then the criterion itself must be unclear. However, this is a questionable premise: there are many other possible explanations for the difficulties surrounding the second criterion expressed by proposers and reviewers alike. As the NAPA Report states, 'Reviewers who tried to apply Criterion 2 as a matter of course in their own evaluation process, generally found its language reasonably clear' ([1], p. 71). This suggests that reviewers who already took Criterion 2 seriously had little difficulty understanding the language of the second criterion.

Yet this also suggests other possibilities: perhaps the language of Criterion 2 is not in and of itself conceptually unclear; perhaps there exist some reviewers who do not take Criterion 2 seriously. In fact, the NAPA Report supports this latter possibility: 'Some scientific communities have found Criterion 2 hard to accept. NSF received approximately 300–400 emails on the new criteria that showed a strong bifurcation of opinion. Approximately half saw NSF as having been too elitist and, therefore, welcomed the change to the new criteria. Half remained purists and didn't like the new criteria. Mathematicians, for example, were against the new criteria. Geophysicists have been for them' ([1], p. 83). Although the fact that approximately half of the reactions to the new merit review criteria were 'positive' while approximately half were 'negative' does indicate a 'bifurcation of opinion', that geophysicists are described as 'anti-elitist' while mathematicians are portrayed as 'purists' suggests that this 'bifurcation of opinion' may involve issues deeply rooted in disciplinary identity. In such a case, further clarification of the language of the second criterion may not be the best or only course of treatment: it may not be simply that some reviewers misunderstand the language of Criterion 2, but rather that different scientific communities (i.e. disciplines) interpret Criterion 2 differently. Such differing interpretations may rest on different disciplinary projects and perceptions. Some disciplines, for example, may see themselves as purely scientific and, therefore, necessarily unconcerned with the broader impacts of their research.

Indeed, as the NAPA Report states, 'the concept of broader social impact raises philosophical issues for many reviewers—in particular, reviewers who see their task as exclusively one of assessing the intellectual merit of proposals' ([1], p. 14, authors'

[2] Perhaps the most persistent calls for clarification of the language of Criterion 2 come from the reports of the Committees of Visitors (COV), outside experts who provide feedback to NSF on various aspects of program-level operations and outcomes of NSF-funded research. Among the program-level operations about which COVs provide feedback is a program's adherence to the merit review process, including its use of both merit review criteria, with special focus on the extent of each program's use of Criterion 2. However, such calls for clarification of Criterion 2 also appear in the reports of NSF's Advisory Committee for GPRA Performance Assessment (AC/GPA). AC/GPA provides advice to the Director regarding NSF's performance vis-à-vis GPRA.

emphasis). Although the NAPA Report raises the specter of 'philosophical issues' surrounding NSF's second criterion, it fails to pursue such issues in any detail.

4. A brief historical outline of issues surrounding Criterion 2

One of the main reasons behind the 1997 restructuring of NSF's generic merit review criteria was the desire to link public investment in science with societal benefits, to demonstrate, in other words, that the people were getting a good return on their investment.[3] Congress had passed GPRA in 1993, and it was partly in response to such demands for demonstrable results that in 1995 NSF had adopted a new strategic plan, according to which, among the long-term goals of the Foundation was 'the promotion of the discovery, integration, dissemination, and employment of new knowledge in service to society' [3]. The goal of 'knowledge in service to society' was meant to link NSF's goal of world leadership in science and engineering with contributions to the national interest.

4.1. NSB–NSF Task Force on Merit Review

Also in 1995, NSB stated its desire to re-examine the merit review criteria that had been in effect since 1981, in light of NSF's new strategic plan. In 1996 the Board established the NSB–NSF Task Force on Merit Review to examine and evaluate the old criteria. In its Discussion Report [2] the Task Force recommended two generic criteria to replace the four 1981 criteria: (1) What is the intellectual merit and quality of the proposed activity? and (2) What are the broader impacts of the proposed activity? Among the perceived advantages of the proposed new criteria were that they would be helpful in connecting NSF investments to societal value while preserving NSF's ability to select proposals on the basis of scientific excellence, and that the new criteria were more clearly related to the goals and strategies of the new strategic plan. NSF published the recommendations of the Task Force on the Web, through press releases, and through direct contact with universities and professional associations and received around 300 responses from the scientific and engineering community.

In light of these responses, in 1997 the Task Force published its Final Recommendations [4]. The responses raised several concerns about the new criteria, including what the Task Force termed the issue of 'weighting' the criteria: Criterion 1 was perceived by respondents as more important than Criterion 2, or Criterion 2 was perceived as irrelevant, ambiguous, or poorly worded. Moreover, respondents expressed concern that for much of basic research it is impossible to make meaningful statements about the potential usefulness of the research. The Task Force noted that 'respondents may be interpreting this question too narrowly. While it may not be possible to predict specific potential applications for one's research, one should be able to discuss the value or

[3] For a brief description of the motivations behind the re-examination of NSF's merit review criteria, see [2], Section 1. Context of the Report. For a more detailed history of the development of NSF's new merit review criteria, including a 'Key Events and Decisions Timeline', see [1], pp. 23–31.

applicability of the line of inquiry or research area'. In response to the issue of 'weighting', the Task Force recommended stating that 'the criteria need not be weighted equally'. Ultimately, the Task Force judged the criteria to be flexible enough 'to be useful and relevant across NSF's many different programs', and recommended that the new criteria be adopted. Later in 1997, NSF issued Important Notice No. 121 [5], which announced NSB approval of the new merit review criteria, effective October 1.

With the approval of the new merit review criteria, NSB had effectively increased the profile of the importance of the societal benefits of NSF-funded projects. Yet in doing so, they had also laid the foundation for a continuing philosophical conflict. At stake for NSB was thoroughly integrating the merit review process with their new strategic plan, which had been designed to increase the profile of the societal benefit derived from NSF-funded research. This makes perfect sense, since NSB's purpose is to set policy for NSF. At stake for the Task Force on Merit Review was finding a way to carry out this process of integration by means of revising the merit review criteria. Again, this makes sense, since this is precisely what the Task Force was tasked to do. Yet scant attention was paid to what was at stake for the respondents to the Task Force's proposed new criteria: what was at stake for the scientific and engineering communities? While NSB approached the issue from a larger policy perspective, and while the Task Force focused on producing the most generic, flexible criteria that would integrate intellectual merit and societal benefit, members of the scientific and engineering communities expressed diverse reservations about the proposed new criteria: some did not understand Criterion 2, some did not find it very important, some claimed it was irrelevant, some claimed it was impossible. Is Criterion 2 unclear? Is it relatively unimportant? Is it irrelevant? Is it impossible to answer?

By recommending that 'the criteria need not be weighted equally', the Task Force was attempting to remain task-oriented: their rationale for this recommendation was that it would maintain the flexibility of the criteria. Yet in doing so, they also undermined the effectiveness of Criterion 2—whoever claimed that Criterion 2 was irrelevant was effectively given carte blanche to ignore it.[4] This freedom effectively allowed discussion of the fundamental differences surrounding Criterion 2 to be postponed.

4.2. NAPA Report

Yet these issues would resurface in the 2001 NAPA Report. External pressure from Congress had not gone away: in 1998 the Senate directed NSF to contract with NAPA to review the effects of changes in the merit review criteria [8], a direction they reiterated in 1999 [9]. In 2000, NSF commissioned the NAPA study. Among the 'Major Conclusions and Recommendations' of the NAPA Report is that 'there is a need to improve the conceptual clarity of the objectives of the new criteria as well as the language used in stating them'. The report continues: 'Asking scientists to speculate about the possible

[4] NSF did not intend to give proposers and reviewers carte blanche, however, as [6,7] indicate: NSF requested that proposers and reviewers consider both intellectual merit and broader impacts in preparing and evaluating proposals for NSF.

future broader or societal impacts of a proposal raises a distinct level of discomfort for many reviewers. This discomfort is increased when precise definitions of some of the objectives of the new criteria remain ambiguous. The conceptual clarity of the new review criteria, therefore, needs to be improved so the criteria better reflect the intentions of NSF for instituting them. This is true of the language of Criterion 2, in particular' ([1], p. 8). It is interesting to notice that the NAPA Report does not claim that the discomfort caused by Criterion 2 is due to lack of clarity. Rather, it suggests that the discomfort caused by Criterion 2 is increased by a lack of conceptual clarity. As the Report goes on to suggest, *'Rewriting the language of the review criteria and restructuring their order is essentially treating only surface-level symptoms and not addressing underlying issues, about which there is considerable diversity of views within the scientific and academic communities. The ultimate differences about issues raised by Criterion 2 are not those of language but of belief*' ([1], p. 9, my emphasis). Ultimately, the NAPA Report asserts, *'the concept of broader social impact raises philosophical issues for many reviewers—in particular, reviewers who see their task as exclusively one of assessing the intellectual merit of proposals'* ([1], p. 14, authors' emphasis). In drawing attention to the fundamental issues surrounding Criterion 2, the NAPA Report provided yet another opportunity for discussion of those issues.

Yet in its FY2000 Report on its Merit Review System, NSF describes the recommendations of the 2001 NAPA Report as follows: 'The key finding was that it is too soon to make valid judgements [sic] about the impact and effectiveness of the new criteria. The NAPA report also highlighted the need to (1) improve the conceptual clarity of the criteria, (2) better communicate with proposers, reviewers and NSF staff about how the criteria are to be used, and (3) improve quantitative measures and performance indicators to track the objectives and implementation of the new criteria. NSF is implementing these suggestions beginning in FY 2001' ([10], p. 14). The NAPA Report did indeed conclude that it was too early to make a valid judgment about the effectiveness of the new criteria. However, to characterize this conclusion as 'the key finding' of the report is a bit misleading: it was one of five 'Major Conclusions and Recommendations', including also the need for quantitative measures to track the new criteria, the need for improving the conceptual clarity of the criteria, using targeted programs to address broader impact, and the need to move beyond simply modifying the language of the new criteria ([1], pp. 7–9). Moreover, the NAPA Report also offered four additional 'Recommendations to Expand NSF's Merit Review Process Improvement Initiatives', among which was included a recommendation to address the 'intellectual and philosophical issues' raised by the new criteria ([1], pp. 13–14). By emphasizing as 'the key finding' the NAPA Report's conclusion that it was too early to make a valid judgment about the effectiveness of the new criteria, and by agreeing to implement three so-called 'highlighted suggestions', the FY 2000 Report on the Merit Review System effectively gave the impression that, at least with regard to the NAPA Report, everything was under control. Moreover, by downplaying or even omitting other 'key findings' of the NAPA Report, the FY2000 Report on Merit Review effectively allowed discussion of the fundamental differences surrounding Criterion 2 to be postponed once again.

This is not to suggest that NSF had no interest in or intention of improving the merit review process. On the contrary, NSF has expended a great deal of time and resources on

improving merit review. One of the key areas on which NSF has focused in terms of improving the merit review process is increasing reviewer and program officer attention to both merit review criteria.[5] The FY2001 Report on Merit Review details the actions undertaken to insure that both criteria are addressed, including, but not limited to, (1) developing and disseminating a draft set of examples of activities that address the broader impacts criterion—in order to 'improve the conceptual clarity' of the criteria, (2) drafting revisions to the Grant Proposal Guide that instruct proposers that they must clearly address broader impacts in their proposals—in order better to 'communicate with proposers, reviewers and NSF staff about how the criteria are to be used', and (3) designing activities to increase program officer attention to the broader impacts criterion through training of new program officers and through electronic tracking of program officer use of both criteria in making recommendations to fund or decline proposals—in order to 'improve quantitative measures and performance indicators to track the objectives and implementation of the new criteria' [11]. In other words, NSF attempted to follow the suggestions of the NAPA Report as described in NSF's FY2000 Report on Merit Review.

Later in 2002, NSF issued Important Notice No. 127 [12], which informed proposers that, effective October 1, 2002, NSF would return without review proposals that did not separately address both merit review criteria within the Project Summary. Important Notice No. 127, therefore, rescinds the notion that proposers and reviewers have carte blanche to ignore Criterion 2. In sum, NSF was taking great pains to insure that Criterion 2 was being addressed throughout the merit review process.

4.3. Quantity and quality in the application of Criterion 2

Given the attention now being paid by NSF to the use and abuse of Criterion 2, one would expect an improvement in its application; and to some extent this is true. The most recent (2004) Report of the Advisory Committee for GPRA Performance Assessment (AC/GPA) notes that NSF's merit review process is, on the whole, 'impressive' [13]. The AC/GPA Report also notes some improvement in the application of Criterion 2: 'One of NSF's original GPRA goals was to increase reviewer and program officer (PO) attention to both of the merit review criteria. It was noted in the two previous AC/GPA reports that consideration of the broader impact of the research continued to be somewhat inadequate. In 2003, 90% of the reviewers commented on both merit review criteria, up from 84% in 2002 and 69% in 2001. Thus, there has been considerable progress on addressing the two criteria' ([13], p. 46). That is, there has been considerable progress in the quantity of reviewers who address Criterion 2. 'However', the Report continues, 'the quality of response to the broader impacts criterion is still an issue. Several COV reports as well as comments from the AC/GPA indicate that the discussions of this criterion frequently lack substance and appear to be cursory at best, even though NSF now requires a one page discussion of both criteria in the project summary of the proposal. In 2003, 276 proposals were returned because this discussion was missing completely. The AC/GPA finds that the review of the broader impacts criterion remains a challenge for most reviewers. We noted

[5] In fact, this has been incorporated as one of the GPRA performance goals for the foundation since FY 1999.

some inconsistency in the completeness and quality of this part of the review and we recommend that NSF continues to focus on this issue' ([13], pp. 46–47). NSF's attention to Criterion 2 has produced improvement in terms of the *quantity* of proposers and reviewers who address Criterion 2; yet the *quality* of the responses to Criterion 2 remains a persistent problem.

This points to a limitation in NSF's use of quantitative analyses of the application of Criterion 2: even if 100% of proposers and reviewers were to address Criterion 2 in their proposals and reviews, if they do so in a manner that lacks substance, the question as to the broader impacts of the proposal will remain unanswered.[6] This also points to an area of vulnerability as regards NSF's Organizational Excellence goal of operating a credible, efficient merit review system.[7] Unless the quality of responses to Criterion 2 improves, the credibility of the merit review system will suffer. Moreover, a lack of substance in reviewer responses to Criterion 2 decreases the efficiency of the merit review process: determinations of the broader impacts of proposals are essentially left to the program officer alone.

4.4. Persistent problems

That problems persist with the quality of responses to Criterion 2 indicates that NSF's efforts to clarify the meaning of 'broader impacts' have not been entirely successful. This leaves open the question of whether Criterion 2 is in need of conceptual clarification. However, it also opens up the possibility that other factors are involved in the lack of quality responses to Criterion 2, namely the philosophical issues alluded to in the 2001 NAPA Report: is Criterion 2 inconsistent with Criterion 1? Is Criterion 2 irrelevant to basic scientific research? Is Criterion 2 completely unanswerable? Is the lack of quality response to Criterion 2 related to scientific disciplinarity? If we recall the initial impetus behind NSB's restructuring of the merit review criteria, i.e. to link scientific research to societal benefit, it is possible to make the point even more starkly. Is intellectual merit inconsistent with societal benefit? Is societal benefit irrelevant to basic scientific research? Is the question of the societal benefit of scientific research completely unanswerable? If so, then what implications does this have for NSB's desire to make the societal benefits of NSF-funded research more obvious? Is it at all possible to link scientific research to societal benefit? If so, how? If not, then why should the public continue to fund such research?

---

[6] Prior to 2004, NSF's GPRA Performance Plans addressed merit review in terms of a 'Management Goal', as opposed to a 'Strategic Outcome Goal', and set the specific goal on the use Criterion 2 at 70% usage. That is, NSF would be 'successful' in its Management Goal relating to the use of Criterion 2 if at least 70% of reviewers commented on Criterion 2 in their reviews. NSF was 'not successful' in FY2001, as only 69% of reviewers addressed both criteria; in FY2002 and FY2003, NSF was 'successful', since 84% and 90% (respectively) of reviewers addressed both criteria ([14], p. II-47). Since 2004, 'Management Goals' have become 'Organizational Excellence', a fourth 'Strategic Outcome Goal'. NSF's new 'Organizational Excellence' goal vis-à-vis merit review is to 'operate a credible, efficient merit review system' ([15], see also [13], p. 45).

[7] See [13], p. 45.

5. Science policy and Criterion 2

The issues of whether, to what degree, how, and under what constraints scientific research should be publicly funded were raised long before Vannevar Bush wrote *Science—the Endless Frontier*.[8] Nevertheless, it was Bush's answers to these questions that eventually led to the establishment of the National Science Foundation.[9] Even more important, however, is the continuing influence of Bush's ideas regarding the nature of scientific research and the relationship between scientific progress and societal benefit. Moreover, these ideas and the issues surrounding them are still relevant to NSF today and have important implications for NSF in the future.

Bush coined the term 'basic research', even if the denotation of the term is difficult to distinguish from H. A. Rowland's 'pure science' [20]. One of the salient features of basic research, according to Bush, is its lack of concern with 'practical ends'—ends that are the proper province of 'applied research'. According to the Bush conception, applied research depends on basic research. In fact, Bush argues that ultimately technological, medical, and military advancements (along with their associated economic benefits) all fundamentally depend on basic research. Although the uses of basic research and the eventual benefits that will accrue are difficult to predict, to eschew basic research would result in grinding progress to a halt. Bush's conception of the dependence of societal progress on basic scientific research ultimately led to what has become known as the linear model.[10]

Bush's conception of basic research and its relationship to applied research and societal benefit were and are highly influential on NSF's self-perception and public image. A comparison of the language contained in documents removed in time by almost 50 years, *The Second Annual Report of the National Science Foundation from FY1952*, and the introduction to *The National Science Foundation at 50: Where Discoveries Begin from 2000*, reveals a remarkable consistency of views.[11] Moreover, in a recent talk delivered at The 30th Annual American Association for the Advancement of Science (AAAS) Forum on Science and Technology Policy (held April 21–22 in Washington, D.C.), new NSF Director Arden Bement continued to

---

[8] See the Report Overview of [16].

[9] This is obviously an over-simplification: NSF was not created until 1950, and its eventual form was quite different from the plan Bush laid out for a National Research Foundation [17]. For a brief account of the response to Bush's plan, see [18], pp. 50–57. For a detailed account of the relationship between Bush's plan and the Steelman Report, see [19].

[10] For more detailed discussions of the linear model and its relationship to V. Bush, see [21,18], especially chapter 1; and [22].

[11] Examples abound. From 1952: 'Basic research is the pacemaker for applied work. Basic research aimed at producing more adequate data and at times new fundamental scientific discoveries hastens the progress of applied research. It serves to clarify the practical problems to be solved and enables the applied research scientist to lay out the course of his work in the most direct and economical manner' ([23], p. 8). From 2000: 'At the National Science Foundation, we invest in America's future. Our support of creative people, innovative ideas, and cutting-edge technologies has led to thousands of discoveries vital to our nation's health and prosperity…. The point to remember is that these and other advances came only after long years of publicly funded basic research' ([24], p. 1). Indeed, as Stokes points out, the 1952 Annual Report contains a virtual restatement of the linear model [18], p. 54; and the 2000 introduction to Where Discoveries Begin invokes Bush by name.

appeal to the rhetoric of science on and as the 'frontier' and to refer to Bush by name. This consistency can be traced to the continued adherence of NSF to many of the ideas Bush laid out in Science—The Endless Frontier, in particular the 'basic' versus 'applied' distinction and the linear model.

Yet in recent years, many of Bush's central tenets (including the opposition between basic and applied research as well as the linear model) have faced increasing criticism from the science policy community. A 1997 letter written by then Speaker of the House Newt Gingrich calls for the abandonment of the Bush model in favor of 'a new, sensible, coherent long-range science and technology policy' ([16], 'The Speaker's Charge'). Daniel Sarewitz offers a critique of what he terms the 'myths' of postwar science policy, most of which he traces back to Science—The Endless Frontier [25]. Sarewitz suggests moving toward a new mythology that would foster scientific research that serves the public interest. He writes: 'The general idea is to graft mechanisms onto the system that create a stronger motivation for pursuing, and better tools for recognizing and measuring, direct contributions of science to societal goals' ([25], p. 172). Roger A. Pielke, Jr. and Radford Byerly, Jr. characterize the Bush model as a paradoxical social contract that 'would exclude societal concerns from setting research paths and priorities. Indeed, science is accountable through the paradox that research done to advance science—without any consideration of practical benefits—is justified by the practical benefits that ultimately result' ([22], pp. 42–43). Pielke and Byerly suggest that scientists need to renegotiate their social contract. Donald Stokes offers a detailed account of various paradigms of scientific research and advocates replacing the Bush paradigm of 'pure basic research' with a paradigm of 'use-inspired basic research' in which scientific research would be inspired by both a quest for fundamental understanding and considerations of use ([18], especially chapter 3). Although these approaches vary in their specific recommendations, a conspicuous point of agreement is their shared conclusion that the Bush model of scientific research must be abandoned.

Implicit in the claim that the Bush model of scientific research must be abandoned is the idea that the opposition of pure basic research to considerations of societal impact represents a fatal flaw in the scientific community's quest to justify continued public investment in scientific research. In other words, Gingrich, Sarewitz, Pielke, Byerly, and Stokes all agree that in order to justify continued public investment in scientific research, the scientific community must adopt a new model of scientific inquiry that incorporates intellectual considerations of the nature of scientific research with considerations of societal benefit. Among these thinkers, Stokes is the only one to point to the pivotal role NSF can play in the adoption of this new model ([18], pp. 151–152).

Indeed, Stokes recommends a paradigm of use-inspired basic research that is highly compatible with NSF's current merit review criteria. Unfortunately, *Pasteur's Quadrant* was published in 1997, the same year in which NSF's new merit review criteria were put into effect. It was, therefore, impossible for Stokes to relate his paradigm to NSF's new merit review criteria. However, the consensus among policy analysts that a new model must replace the Bush model and the existence of a new model that closely corresponds to NSF's merit review criteria presents NSF with an

opportunity to take center stage in adopting a new vision of the nature of scientific inquiry that combines a quest for fundamental understanding with considerations of societal benefit. Why, then, would NSF continue to operate under the Bush model?

One plausible answer is that there are unresolved philosophical issues in the scientific community that hinder the adoption of a new model of scientific inquiry that incorporates fundamental understanding with societal benefit. These are, of course, the same unresolved philosophical issues that have hindered the incorporation of the broader impacts criterion into the merit review process. Another plausible answer is that insofar as NSF has avoided addressing the fundamental issues surrounding the scientific community's opposition to Criterion 2, NSF has failed to see the relationship between those issues and the larger questions of policy. What is necessary, then, is to begin to address those fundamental philosophical issues.

6. Conclusion

Scientists and engineers deal professionally with what we might term broadly 'matters of fact', questions that are, in principle, resolvable by empirical means. Whether the theory of evolution correctly infers a common ancestry for all living things is just such a matter of fact, as is the question of whether life exists elsewhere in the universe: both are susceptible, in principle, to empirical testing. On the other hand, we also often encounter the opposite sort of question, a sort not empirically resolvable even in principle, which we might term broadly 'matters of opinion'. Whether spinach actually tastes good is such a matter of pure opinion, and no amount of empirical testing will settle the issue.[12] However, we continue to adhere to a damaging prejudice, one Nietzsche diagnoses as the 'faith in opposite values', if we reduce all matters to one of these two opposites. It is simply not the case that anything and everything not susceptible to empirical testing is a matter of pure opinion.

Unlike scientists and engineers, philosophers are professionally accustomed to operating in areas that represent a middle ground between 'objective facts' and 'subjective opinions'.[13] Questions of ethics or aesthetics, for instance, cover this middle ground. We can refer to such middle-ground matters as 'philosophical issues'. If such philosophical questions are not susceptible to scientific proof, neither are they relegated to the realm of mere opinion. Rather, philosophical issues are subject to what one might term a reasonable discussion. Far from presupposing any sort of strict notion of rationality, such a reasonable discussion presupposes only that we are addressing an issue that, although not susceptible to empirical testing, is nevertheless too important to be relegated to the realm of pure opinion. To take part in such a reasonable philosophical discussion need not involve us in aimless and endless metaphysical meanderings as abstruse as they are abstract. Whether,

[12] Notice that this is a claim about the spinach, not a claim about whether you or I or most people like spinach, which would obviously be empirically testable matters of fact.

[13] Philosophers have not, however, tended toward questions of science policy. For a catalog of this lapse in philosophical attention, as well as a counter-example to the usual tendency, see [26]. Philip Kitcher is perhaps the most notable exception [27].

for example, murder is wrong is neither an empirically testable matter of fact nor a matter of pure opinion; nevertheless, a reasonable discussion of the matter has in fact led to a general consensus that murder is wrong. Some might argue that there are cases, such as self-defense, in which murder is justifiable. But now we are involved in a reasonable discussion.

I propose that we begin a reasonable public discussion of the fundamental 'philosophical issues' surrounding Criterion 2.[14]

Some, including some philosophers, may view philosophy from a narrow disciplinary perspective as concerned only with abstract issues. Some, including some philosophers, may view the following list of what I shall term 'philosophical issues' as including issues more properly termed sociological or political. I contend, however, that the following issues are philosophical, even if they represent a new field of philosophy—what has been termed elsewhere the philosophy of science policy [27]. Finally, I do not intend the following list to be exhaustive. On the contrary, I expect and hope that others will raise different issues with Criterion 2. Nevertheless, I put the following 'philosophical issues' up for debate: (1) whether Criterion 2 is in need of conceptual clarification, (2) whether Criterion 2 is inconsistent with Criterion 1, (3) whether, and to what extent, reactions (either 'positive' or 'negative') to Criterion 2 depend on one's disciplinary identity, (4) whether, and to what extent, reactions to Criterion 2 depend on the degree to which one has (consciously or unconsciously) incorporated the Bush model of the opposition between basic and applied research, (5) whether NSB's approval of the new criteria entails a tacit commitment to a new model of scientific inquiry, (6) whether NSB is pushing NSF to move away from supporting basic research, (7) whether we can find appropriate qualitative as well as quantitative measures for the application and interpretation of Criterion 2, (8) whether one's mostly scientific peers possess the necessary expertise to assess the 'broader impact' of one's proposal, (9) whether Criterion 2 should be modified or abandoned, and (10) whether NSF could or should use Criterion 2 as a sort of fulcrum in support of a leadership role in developing a new national science policy.

Fundamentally, it seems to me, these issues surrounding Criterion 2 are ramifications of the larger issue of the relation between science and society, issues of scientific autonomy and responsibility. However, Criterion 2 is unique in that it presents an actual case of these abstract issues. Moreover, this particular case is no isolated instance, but rather an institutionalized fact—one with which a large number of scientists and engineers have had some experience, and one with which more than a few have had some difficulty. I believe that the case of the NSF's Second Criterion presents us (not only us philosophers, but also you members of the NSB, you members of Congress, you policy scientists and political scientists and all you scientists and engineers) with a unique opportunity to address the relation between science and society. I look forward to our discussion.

---

[14] Although the content of his concern is different from my own, it is interesting to note that Alan I. Leshner, CEO of the American Association for the Advancement of Science, has recently called for just the sort of reasonable discussion I am proposing: 'We should try to find common ground through open, rational discourse' [28].

Acknowledgements

An earlier version of this paper served as part of a proposal to NSF to examine their merit review criteria. The author would like to thank several anonymous reviewers for their helpful criticisms. The author would also like to thank Bob Frodeman for insightful comments on many iterations of that proposal, as well as for pointing the author toward a philosophical examination of Criterion 2. Thanks also to Roger Pielke for some excellent guidance on sources on the linear model. Finally, the author wishes to thank Carl Mitcham for his extensive and skillful editing of this final version of the paper.

References

[1] National academy of public administration (NAPA). A study of the national science foundation's criteria for project selection, a report by the national academy of public administration for the national science foundation; February 2001.
[2] NSF. Task force on merit review's discussion report (NSB/MR 96–15).
[3] NSF. NSF in a changing world (NSF 95–24).
[4] NSF. Task force on merit review's final recommendations (NSB/MR 97–05).
[5] NSF. Important notice no. 121.
[6] NSF. 'Dear colleagues' letter to PIs and reviewers (NSF 99–172).
[7] NSF. Important notice no. 125.
[8] Senate report 105–53.
[9] Senate report 105–216.
[10] NSF. FY2000 report on the NSF merit review system (NSB 01–36).
[11] NSF. FY2001 report on the NSF merit review system (NSB 02–21).
[12] NSF. Important notice no. 127.
[13] NSF. Report of the advisory committee for GPRA performance assessment (NSF 04–216).
[14] NSF. FY2003 performance and accountability report (NSF 04–10).
[15] NSF. FY2004 performance and accountability report (NSF 05–01).
[16] *Unlocking our future: toward a new national science policy*, A report to congress by the house committee on science, september 24, 1998, available online at: www.house.gov/science/science_policy_report.htm.
[17] Bush Vannevar. *Science - the endless frontier*. Washington, DC: United States Government Printing Office; 1945 [Available online at: http://www.nsf.gov/about/history/vbush1945.htm].
[18] Stokes DonaldE. *Pasteur's quadrant: basic science and technological research*. Washington, DC: Brookings Institution Press; 1997.
[19] Blanpied WilliamA. Science and public policy: the Steelman report and the politics of post-world war II science policy In AAAS science and technology policy yearbook 1999 p. 305–20.
[20] Rowland HA. A plea for pure science. Science 1883;2(29):242–50 [August 24].
[21] Brooks Harvey. The Evolution of U.S. Science Policy. In: Smith BruceLR, Barfield ClaudeE, editors. in: Technology, r&D, and the economy. Washington,DC: Brookings Institution and American Enterprise Institute; 1996. p. 15–48.
[22] Pielke, Roger Jr A, Radford Jr Byerly. Beyond basic and applied In physics today 1998 p. 42–6.
[23] NSF. Second annual report of the national science foundation, FY1952.
[24] NSF. The national science foundation at 50: where discoveries begin, 2000.
[25] Sarewitz Daniel. Frontiers of illusion: science, technology, and the politics of progress. Philadelphia: Temple University Press; 1996.
[26] Special issue of philosophy today. In: Frodeman Robert, Mitcham Carl, editors. Toward a philosophy of science policy: approaches and issues, vol. 48; 2004, 2004. p. 5.

[27] Kitcher Philip. Science, truth, and democracy. New York: Oxford University Press; 2003.
[28] Leshner Alan I. Editorial: where science meets society. Science 2005;307(11):815 [February].

Britt Holbrook received his PhD in philosophy from Emory University in August 2004. He has served as Research Assistant Professor within the Department of Philosophy and Religion Studies at the University of North Texas since January 2005. He has authored an entry on NSF's second merit review criterion for MacMillan Reference's *Encyclopedia of Science, Technology, and Ethics* as well as co-authoring (with Robert Frodeman) an article on the policy ramifications of Criterion 2 for *Ogmius* (forthcoming).