RT. A Journal on Research Policy & Evaluation **1** (2013)
Submitted on 17 March 2013, accepted on 17 April 2013, published on 21 July 2013

Doi: 10.13130/2282-5398/2914

# Blue skies, impacts, and peer review

**J.B. Holbrook[*], S. Hrotic[**]**

**Abstract** This paper describes the results of an on-line survey regarding the incorporation of societal impacts considerations into the peer review of grant proposals submitted to public science funding bodies. The survey investigated perceptions regarding the use of scientific peers to judge not only the intrinsic scientific value of proposed research, but also its instrumental value to society. Members of the scientific community have expressed – some more stridently than others – resistance to the use of such societal impact considerations. We sought to understand why. Results of the survey, based on a convenience sample of 428 participants (including individuals from four different funding agencies as well as non-academics), suggest that such resistance may be due to a lack of desire rather than a lack of confidence where judging impacts is concerned. In other words, it may be less that scientists feel unable to judge broader societal impacts and more that they are unwilling to do so.

**Keywords** peer review, impact, expertise, policy, autonomy, evaluation

## 1. Introduction

A recent feature in *Nature* [Bhattacharya 2012] describes issues surrounding the introduction of societal impact criteria to the process of grant proposal peer review as a "duel to the death" between researchers and research funders. Although hyperbolic, this description does capture some of the passion that researchers evince when they feel that peer review is threatened. Researchers in the UK and Canada recently staged protests against their own government funding agencies, both of which took the form of mock funerals. Their message: science is dead, and we have killed it. What assumptions underlie such strong reactions regarding peer review? It is this question we attempt to answer in this paper.

In the first section below, we review the current research on peer review. Most empirical studies of peer review treat peer review as a tool for evaluating the quality of research. These studies focus

[*] University of North Texas, Denton, Texas, britt.holbrook@unt.edu

[**]University of Vermont, stevenhrotic@yahoo.co.uk

on issues of reliability, bias, efficiency, validity, etc. We chose to focus on implicit foundational issues surrounding peer review: who counts as a peer, what is the value of research, and how should it be judged?

The next two sections of the paper describe one of the components of our own research: an on-line survey of peer reviewers, program managers, and grant applicants. We attempt especially to explore how respondents value research. We also explore respondents' levels of confidence with regard to judging two different aspects of research: (1) its intrinsic intellectual merit, and (2) its broader societal impact.

The final section discusses our results. Since our survey is based on convenience sampling, we cannot generalize our results to a defined population. We do not suggest, therefore, that we have conclusively answered any of our own questions; but we believe we have taken steps in the right direction. In attempting to draw out researchers' underlying assumptions regarding peer review, we hope at least to have opened up new questions for researchers on peer review, especially those researchers concerned to investigate the incorporation of societal impacts considerations into peer review processes. Our results and the questions they raise should also be of interest to research funding agencies that attempt to incorporate societal impacts concerns into their own peer review processes.

## 2. Background

Since the first half of the 20th century, public science funding agencies around the world have used input from members of the scientific community to inform their funding decisions, a process commonly known as peer review. Since the end of the 20th century, however, many agencies have begun to ask peer reviewers to assess grant proposals not only in terms of narrow questions of intellectual merit, but also in terms of broader matters of societal impact. Some agencies have incorporated such impact considerations directly into their peer review processes in the form of explicit societal impacts criteria. Asking scientific peers to judge the scientific merit of proposed research – the traditional purpose of peer review – makes intuitive sense: who better to judge scientific excellence than scientific experts? Asking scientific peers to judge the broader societal impacts of research, however, raises conceptual, even philosophical, issues.

In what sense are scientific experts specially qualified to judge broader societal impacts? Some have argued that they are not [Bozeman and Boardman 2009]. Should scientists consider the broader societal impacts of their research, or should they simply concentrate on the science? Others have argued not only that scientists have a responsibility to consider the broader impacts of their research, but also that asking them to do so takes scientists beyond the bounds of their disciplinary expertise [Frodeman and Holbrook 2007].

Are experts in societal impacts necessary to judge the broader impacts of proposed research? This would constitute the interdisciplining of the notion of a peer [Frodeman and Holbrook 2007; Frodeman and Briggle 2012; Nightingale and Scott 2007]. But why stop there? Should the members of society who will be most involved in its application – either as potential users of the research or simply as those most interested in or affected by it – be included in the (now more socially robust

and perhaps well-ordered, if no longer quite normal) peer review process? [Funtowicz and Ravetz 1993; Kitcher 2001; Nowotny et al. 2001].

This would constitute the transdisciplining of peer review.

Who, then, should count as a peer? What would extending the notion of a peer to non-traditional peers mean for the relationship between scientific knowledge production and the use of that knowledge by society? [Holbrook (2010a); Nowotny et al. 2001].

If everyone counts as a peer, then no one counts. Perhaps there are no experts anymore.

In addition to such abstract philosophical issues, the incorporation of societal impacts considerations into the peer review process raises practical concerns. Confronted by philosophical confusion regarding who should count as a peer, how can agencies identify appropriate peer reviewers? If they need to extend their search beyond the bounds of single disciplines, how should agency officials manage the sorts of problems that arise in cross-disciplinary communication? [Eigenbrode et al. 2007; Holbrook 2012a; O'Rourke and Crowley 2012].

Recent attempts to link the notions of trading zones and interactional expertise [Gorman 2010] appear promising. Perhaps the context of peer review itself offers some guidance for navigating interdisciplinary communication issues [Lamont 2009; Lamont and Mallard 2005].

How would extending the process of peer review to include reviewers from other academic disciplines (interdisciplining peer review), or even non-academic members of society (transdisciplining peer review), affect the peer review process and how it is managed? [Healy 1999; Scott 2007].

It is worth emphasizing that the philosophical issues and the practical concerns raised in the preceding paragraphs are not merely academic. Scott and Nightingale call on public science funding agencies to incorporate specific relevance criteria into peer review [Scott and Nightingale 2007].

In fact, many agencies have done so already. Moreover, agencies that have incorporated societal impacts considerations into the peer review process have encountered not only philosophic issues [Rip 2000], but also practical difficulties in doing so [Holbrook 2005; Holbrook and Frodeman 2011; Holbrook 2012b].

Given its importance both for regulating academic knowledge production and for negotiating the relationship between knowledge production and the use of that knowledge by society, there is still relatively little research on the process of peer review. Empirical studies of peer review are even rarer. There are a number of reasons for this relative inattention.

First, given the range of uses to which the process of peer review is put, there is no one thing that counts as *the* process of peer review [Hackett 1997].

Indeed, 'peer review' is a generic catch-all for a wide range of processes using expert judgment. The scope of the term 'peer review' varies according to context and scale. We can speak of peer review in the following contexts: (a) pre-publication review of a manuscript to determine its fitness for publication, (b) review of a faculty member's portfolio of work for the purposes of granting tenure or promotion, (c) review of grant proposals, (d) review of academic programs (for instance, a department, center, school, or college), (e) review of quasi-academic programs (for instance, funding agencies often use experts to evaluate *ex ante*, *ex nunc*, or *ex post* particular funding schemes), or (f) *ex post* review of scientific findings to be used in regulatory decision making.

It is also possible to distinguish between and compare peer review processes at various scales. For instance, one could investigate peer review of grant proposals at one agency compared to another, or within a particular agency at the program level, or from one proposal to another within an agency or program, or from one reviewer to another on a single proposal, or a single reviewer from one

proposal to another. Similarly, one could examine pre-publication peer review from field to field, from journal to journal, from editor to editor, from reviewer to reviewer, or following a single manuscript through successive submissions and iterations.

Second, there are structural reasons that tend to discourage research on peer review, especially peer review of grant proposals. Public science funding agencies are generally committed to confidentiality regarding reviews and proposals. Such agencies are also budget-sensitive, meaning that they do not want to open themselves up to studies that may endanger their budgets. Chubin and Hackett provide a helpful taxonomy of studies of grant proposal peer review, dividing them into agency-sponsored, agency-funded, and independent studies [Chubin and Hackett 1990].

Agency-sponsored studies, such as the research on the US National Science Foundation (NSF) conducted by the Coles [Cole and Cole 1981; Cole, Rubin, and Cole 1978] or on the US National Institutes of Health (NIH) by Johnson [Johnson 2008], may have almost unlimited access to data; but they "share the problems of the 'court historian' who must please the king" [Chubin and Hackett 1990: 52]. Agencies also conduct ongoing smaller-scale studies that range from including independent observers in peer review panels (in the case of the European Commission [EC]) to having external visiting experts evaluate the peer review process on the program level (in the case of NSF) to conducting experiments on alternatives to peer review [Lightbourne 2010]. Some of these studies are made public, while some are not.

Studies funded by agencies but conducted by outsiders have more freedom of inquiry, but at the price of limited access. Such studies are typically limited to publicly available information, such as the names of grant winners and their proposal abstracts, and lack access to the reviews of proposals or the proposals themselves, much less access to failed proposals and their reviews. Having received a grant from a public funding agency to conduct research (as in our own case) may sometimes lead to some degree of cooperation from funding agencies. For instance, Langfeldt gained impressive access to the Research Councils of Norway [Langfeldt 2001, 2004].

We ourselves experienced different responses from the agencies we examined; some agencies see such research as helpful, and are thus cooperative, while others seem concerned about whether our research might hurt the agency.

Independent studies have unlimited freedom of inquiry but almost no resources. Indeed, Wennerås and Wold relate the fact that they had to go to court to gain access to the reviews conducted by the Swedish Medical Research Council [Wennerås and Wold 1997].

It is understandable that some agencies seek to avoid opening themselves up to criticism; it also makes sense that not all researchers can take an agency to court, or are willing to employ subterfuge [Peters & Ceci 1982].

Most of the literature on grant proposal peer review can be divided into the following categories. First, there are comparative studies, such as the Boden Report [Boden et al. 1991], or the report prepared by Lamont and Mallard on peer review in the US, UK, and France [Lamont and Mallard 2005]. Second, there are studies of the reliability, validity, efficiency, effectiveness, bias, or fairness of peer review [Bornman et al. 2007; Burggren 2009; Johnson 2008; Marsh et al. 2008; Viner et al. 2004; Wennerås and Wold 1997]. This second category is by far the most populated. A third category might include research on interdisciplinary peer review [Laudel 2006; Lamont 2009]. Finally, there are studies that attempt to put peer review in its larger social context [Chubin and Hackett 1990; Hackett 1997].

Despite the fact that agencies have included societal impacts considerations as part of peer review for over a decade, a significant gap in the literature is empirical research on the inclusion of societal

impacts criteria in peer review [Bornman 2012, 2013]. In fact, we know of only two such studies, both of which [Roberts 2009; Kamenetzky 2013] examine the US National Science Foundation's (NSF) Broader Impacts Merit Review Criterion. It is notable that Roberts concludes that including societal impacts as a peer review criterion may not in fact lead to an increase in the societal benefits of funded research.

Our own research is comparative, but our concern is to put peer review of grant proposals in a larger societal context. We therefore ask different questions than other researchers. Hackett, for example, discusses the notion of accountability in terms of accountability to one's peers [Hackett 1997: 57]; we tend to focus on accountability to the larger society and on societal impacts considerations as an attempt to introduce such a larger perspective on accountability into the processes and debates of peer review [Frodeman and Holbrook 2011; Holbrook 2010].

Insofar as we relate societal impacts to calls for the democratization of knowledge, postnormal science, or Mode 2 knowledge production, our research also has ties to much of the literature from science and technology studies mentioned above.

What makes our work different from some of the theoretical literature in science and technology studies, however, is our desire to lend our work empirical support. The Comparative Assessment of Peer Review (CAPR) study was a five-year (2008-2012) examination of the process of grant proposal peer review at public science and technology funding bodies, funded by the NSF's Science of Science and Innovation Policy (SciSIP) program. CAPR research was performed by an interdisciplinary team made up of philosophers (including Holbrook), an anthropologist (Hrotic), information scientists, a biologist, and others. CAPR focused on the ways that societal impacts considerations are incorporated into the peer review process across three US federal agencies: NSF, NIH, and the National Oceanic and Atmospheric Administration (NOAA) and three non-US agencies: the EC's Framework Programmes (FPs), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Dutch Technology Foundation (STW).

The CAPR study had two overall objectives. First, we sought to advance our understanding of the grant proposal peer review process at public science funding bodies. Second, we hoped to improve the process of grant proposal peer review, with particular attention to identifying and helping to ameliorate difficulties that arise with the incorporation of societal impacts considerations into the peer review process. CAPR is both framed by, and directed toward, the needs of the funding bodies we studied, an approach we call 'field philosophy' [Frodeman 2010; Frodeman, Briggle and Holbrook 2012a]. Table 1 offers a summary of some of the most important similarities and differences between the different models of grant proposal peer review at the six agencies included in the CAPR study.

For a more detailed discussion of CAPR's overall rationale and methodology, see [Holbrook 2010b]. For a comparison of NIH and NSF, see [Frodeman and Briggle 2012]. For a comparison of NSF and the EC, see [Holbrook and Frodeman 2011].

The next section presents an interdisciplinary and comparative survey conducted as part of the CAPR study. In conducting our survey, we hoped to add in a small way to empirical investigations of societal impacts and to promote the question of who should count as a peer as important areas for further research. We worked outside the investigating agencies (as appropriate for a comparative study of this kind) and took as given that access to existing data and cooperation of participants in such circumstances is problematic. However, we elected to proceed, in hopes that even qualified results would further encourage others – including agency insiders – to consider some of the issues described above.

**Table 1.** Comparison of 6 CAPR agencies in terms of 5 key aspects of peer review models

| Agency | Are Societal Impacts Criterion Explicit? | Do Reviewers Consider Impacts beyond Education and Outreach? | Does Agency Predetermine Expected Scope of Impacts? | Does Agency Predetermine Weighting of Impacts? | Do Reviewers Make Actual Funding Decision? |
|---|---|---|---|---|---|
| **NSF** | Yes | Yes | No | No | No |
| **NOAA** | Yes[a] | No[a] | Yes | Yes | No |
| **NIH** | No | Yes | No | No | No |
| **NSERC** | No | No | No | No | No |
| **STW** | Yes | Yes | Yes[b] | No | No |
| **EC-FP7** | Yes | Yes | Yes | Yes | Yes |

*Notes*: [a] NOAA has an explicit impacts criterion that focuses only on education and outreach. Reviewers may consider societal impacts beyond education and outreach – for example, the fit with NOAA's mission or strategic plan – but not under the rubric of NOAA's education and outreach impacts criterion.
[b] STW focuses explicitly on the 'utilization' of proposed research. For a discussion of the scope of 'utilization', see the CAPR Midterm Workshop Report.

## 3. The Comparative Assessment of Peer Review (CAPR)

This section describes the construction of a survey designed to uncover the values underlying the idea of a peer, of research, and of the relation between research and society. It also outlines the solicitation of participation in the survey – an area of particular methodological stickiness in the case of peer review.

### 3.1. Survey design

Based on a review of administrative documents at various agencies (for example, NSF's Reports of Committees of Visitors), as well as discussions and interviews with agency officials and members of the scientific community who submit and review research grant proposals, we concluded that significant challenges arise with the inclusion of societal impacts considerations as part of peer review. Most importantly, proposers and reviewers have indicated resistance to addressing societal impacts in their proposals and reviews – a fact that creates obvious difficulties for agency officials who must manage the proposal review process [Holbrook 2005; Holbrook and Frodeman 2011; Holbrook 2012b].

Such resistance is also widely attested to in the literature; proposers and reviewers have often expressed the opinion that being asked to address societal impacts in proposals and reviews is burdensome, unreasonable, or even punitive [APS News 2007; Lok 2010; Bhattacharya 2012]. Such resistance to addressing societal impacts criteria may also help explain the fact that, as Roberts [2009] suggests, such criteria may be ineffective at increasing the societal benefits of funded research. Whether such impact criteria are effective will depend on their use by proposers, reviewers, and agency personnel. We thus chose to delve deeper into the attitudes of those groups toward incorporating considerations of the societal impacts of research into the peer review process.

We also recognized that members of society at large had a stake in peer review processes that are meant to help determine the distribution of public funds. We therefore attempted to target a broad range of stakeholders, including academics (many of whom had served as proposers or reviewers, and some of whom had not), agency officials, and non-academic members of the larger society. We designed an on-line survey intended to investigate three questions regarding attitudes toward the use of societal impacts considerations as part of the peer review process, namely:

Who counts as a 'peer' in peer review?
How does society benefit from research?
What is the relation between the intrinsic and instrumental value of research?

Participants were asked to rate their agreement with statements regarding these issues using 6- and 7-point Likert scales.[1]

*Who counts as a 'peer' in peer review?*

Recent theories in science studies (see above) have suggested extending peer review to involve members of different academic disciplines (interdisciplining the notion of a peer) or members of the larger society (transdisciplining the notion of a peer). Moreover, some agencies do practice such extended peer review. We therefore wanted to determine the range of attitudes of various stakeholders in the peer review process regarding who counts as a peer.

Participants were asked via a 6-pt. Likert scale to rate how involved in funding decisions each of nine groups of potential review 'peers' (including traditional, interdisciplinary, and transdisciplinary peers) should be, in their personal opinion. Labels on responses ranged from (1) "Should not be involved at all" to (6) "Should make funding decision." The nine groups were:

[1] The on-line survey program utilized prevented the survey from being completed twice from the same IP address. The following differences between the description here and the instrument used should be noted, especially: (1) some scales have been reversed for clarity. (2) The survey includes questions not discussed here, including demographic data and independent variables not discussed here. (3) In addition to asking participants to describe their *personal* agreement with a statement, in several cases participants were also asked to assess how well the statement agreed with their *experience* and with their agency's *policy* (if applicable). (We discuss only their personal agreement in this paper, but note that asking participants to assess all three could have the effect of cuing participants to calibrate their own opinions with reference to their agency's and/or their history.) Note also that the on-line version used skip-logic; therefore the written form appears to have redundant sections. The complete survey can be viewed at www.csid.unt.edu/files/CSID_CAPRsurvey_2010.pdf.

Academics from the same or similar field as the applicant, Academics from a dissimilar field, International academics, Researchers employed by the private sector, Funding agency officials, Non-funding agency government officials, Prospective 'users' of the research, Members of the lay public. Members of the international lay public.

*How does society benefit from research?*

Public funding of scientific, technological, and academic research exists because people believe that society somehow benefits from the research it funds. Precisely how those benefits accrue to society, precisely what those benefits may be, and precisely which members of society benefit, however, remain open questions. Policy makers, scientists, and science policy scholars have proposed various narratives to describe the science-society relation. These include, for instance, the so-called Linear Model, Vannevar Bush's (1945) "Endless Frontier," Polanyi's (1962) "Republic of Science," Pasteur's Quadrant (Stokes, 1997), the Haldane Principle, Barbara Mikulski's call for "strategic" research, the notion of "Blue Skies" research, and the so-called "Impact Agenda." Such narratives have also often been exposed by other science policy scholars as so many myths([Sarewitz 1996; Edgerton 2009]).

Regardless of their truth, however, such narratives can and do guide science policy. We also suspected that such narratives can and do affect or reflect stakeholder attitudes toward the inclusion of societal impacts considerations in peer review. For instance, someone who subscribes to a Bush-like narrative would believe that public funding for science is necessary to ensure societal benefits, but that consideration of those eventual benefits plays no part in motivating basic research. Someone subscribing to such a narrative would likely be either neutral or hostile to the idea of asking proposers or reviewers to address societal impacts. Someone who subscribes to the use-inspired basic research narrative associated with Pasteur's Quadrant, however, may be more sympathetic to impact criteria.

Based on a survey of the science policy literature, as well as preliminary interviews with members of the scientific community, we generated a list of nine different narratives as part of the CAPR survey. These included traditional narratives directly related to the linear model, such as:

> By spending money on research, we create a 'BODY OF KNOWLEDGE' that can be used for further research. This is necessary for the application of science, and the development of new technologies, which improve human lives.

We also created contrasting narratives, purposefully including some with negative valences:

> Technological development is INEVITABLE; society must adapt.

All narratives are included in Table 3 below.

The narratives were not meant to be mutually exclusive. Indeed, we intentionally presented overlapping narratives with slightly different emphases. We included more than one narrative to represent a strong sense of autonomy for scientific research – what one could call Blue-Skies-compatible narratives. We also presented more than one narrative to represent a more directive, societal needs-driven notion of the science-society relation.

*Intrinsic and instrumental value of research.*

We considered our questions regarding attitudes toward different narratives of the science-society relation to be important, but also to be perhaps too conceptually distant from the issue of peer review to make sense to everyone, including the survey participants. As a result, we decided to add a section of questions regarding the intrinsic and instrumental value of research. We described the distinction between intrinsic and instrumental value as follows:

> One could value research in two distinct ways: (1) for its INTRINSIC value (that is, one might pursue knowledge simply for its own sake); or (2) for its INSTRUMENTAL VALUE (that is, one might pursue knowledge because it enables us to accomplish some other goal).

Thirteen questions in the survey addressed this distinction. The first eight were arranged in pairs, in which participants were asked to indicate agreement with two broad, descriptive statements, identical except that one included the word "INSTRUMENTAL," the other "INTRINSIC." For example, participants were asked how strongly they agreed with the statement, "I would feel comfortable assessing the INSTRUMENTAL value of research" and "I would feel comfortable assessing the INTRINSIC value of research." Three more questions asked participants how much they agreed with statements regarding the flexibility of weighting instrumental and intrinsic values. Two others addressed specific indicators of instrumental value.

The questions regarding intrinsic and instrumental value were directly related to the peer review process. We wanted to know whether, for instance, participants thought it was as possible to measure intrinsic and instrumental value, whether peer review could be used to judge intrinsic and instrumental value, and whether they themselves felt qualified to judge intrinsic and instrumental value. We wondered whether participants would tend to favor intrinsic value and profess an inability to judge instrumental as compared to intrinsic value (a possible explanation for resistance to societal impact criteria). We also wondered whether participants would favor improving procedures for assessing instrumental value (perhaps so, if they thought instrumental value was important; perhaps not if they were only interested in the value of research for its own sake).

## 3.2. Solicitation of participants

Locating appropriate participants was a non-trivial challenge. We were particularly interested in soliciting the participation of officials and peer reviewers in several key funding agencies (NSF, NIH, NOAA, NSERC, STW, and the EC). These are relatively high-status group with high demands on their time and little motivation to participate in external research that could in effect criticize their agencies in a public forum—indeed, it was reported to us anecdotally that participation in *our* research could conflict with internal responsibilities and ongoing agency-sponsored research. We anticipated response rates to be low (potentially biasing our results in the direction of those with a personal interest in critiquing peer review systems); but we elected to proceed, because we were familiar with the kinds of topics internal surveys of peer review processes addressed and believed important questions (in particular, challenging the assumptions surrounding expertise and who should count as a 'peer') were being missed.

Participants were invited in four ways throughout 2010. (1) On-line lists of agency officials (e.g., program directors) at each of these agencies were used to send personalized invitations. (2)

Instructors at various universities were also personally invited. Thirty universities in the US, Canada, UK, and Europe were randomly selected, to get as broad a regional representation as possible; this required we compromise with the number of disciplines approached. We targeted Biology, Physics, Psychology, and Anthropology departments to represent a range of physical and social sciences. (3) An invitation to participate was posted on several 'science studies' listserves and webpages, including our own. (4) A Facebook ad was placed, targeting users who listed relevant topics of interest ("NSF," "Peer review," "Science funding," etc.). Although the Results section below describes the responses of "non-academics" (i.e., those who had never worked professionally within academia), these are not representative of the general public: over 80% of "non-academics" held a graduate degree and demonstrated (through Facebook profile self-description) interest in peer review processes.

657 total responses were collected. 167 did not complete the survey and were excluded, leaving 490 usable responses.[2] As part of the survey, participants were asked for their "primary agency affiliation" if any; their level of experience as agency officials, peer reviewers, or applicants; and if they had ever worked as professional academics. For the purposes of this report, we chose to describe the four best represented agencies (NSF, NIH, EC, NSERC) and participants without a professional academic background. This excluded a further 62 academics with other agency affiliations.

Thus, the results described below reflect 428 participants. Participants included 141 from NSF (i.e., who listed NSF as their primary agency affiliation), 85 from NIH, 77 from the EC, 54 from NSERC, and 71 who had never worked as a professional academic and who had no agency affiliation. Participants included 78 agency Officials, 171 Peer Reviewers, and 48 Applicants.[3] 42% were female; 57% were from the US (by citizenship), 25% from Europe or the UK, and 13% from Canada.[4]

We must emphasize that our data do not describe agencies' policies. Indeed, the survey was designed to elicit *personal* opinions of individuals involved in peer review.[5] Further, our sample is not demonstrably representative of the population of individuals working within these agencies: given our unofficial role, our lack of access and agency authority limited our methods to a convenience sample, and (therefore) only descriptive data is reported here.

However, we hoped even preliminary data along unexplored lines would spur further research (including agency internal studies, which could ensure high participation rates) with broadened research questions. Though we cannot state with confidence that our results reveal what relevant populations actually believe, we believe they clearly demonstrate there are possible (and possibly common) schema for the mechanisms and aims of peer review and of research funding that have not been considered.

---

[2] Participants who skipped a few questions, but then continued, were not excluded.

[3] We assumed that Officials would have had experience as Peer Reviewers, and Peer Reviewers as Applicants. We gave precedence to upper levels of the hierarchy. Thus "Officials" described here were all participants who reported having average or above experience as Officials; "Peer Reviewers" had average or above experience as peer reviewers but were not Officials, etc. Note that 70 of the 357 academic participants had below average experience in all three categories.

[4] 5% held a dual citizenship, were a citizen of another country, or did not report their citizenship.

[5] Academic participants were asked to describe their agency's "current practice" and their "past experience" within their agency before describing their own opinions (e.g., what classes of individuals should be considered "peers" in peer review: academics from similar fields, academics from other fields, private sector researchers, prospective "users," etc.).

Put differently, we decided the methodological risk was worth the intellectual (and potentially societal) reward: our aim was to have a broader impact rather than simply adhering to standards of intrinsic worth.

## 4. Results

This section describes the average responses of all 428 participants (from NSF, NIH, NSERC, EC, and non-academics) to the central survey questions related to three main topics investigated by this research: who counts as a peer, narratives for how society benefits from research, and instrumental and intrinsic value. Note that inferential tests have not been made in acknowledgment of the irregularity of our sample; our goal was to conduct a pilot sufficient to demonstrate the relevance of these questions (e.g., who qualifies as a review "peer"), not definitively answer them.

*Who counts as a peer?*

Participants were asked to rate (on a 6-point Likert scale) how involved nine different groups of potential 'peers' should be in the grant decision process. Averages of all participants' responses were made. Based on interviews and other preliminary research, nine groups of potential peers were considered (Table 2).

Unsurprisingly, academics with similar disciplinary background – traditional peers – were perceived to be most important in peer review, with participants indicating that they should be most involved in funding decisions. Interestingly, participants felt that traditional peers should be more influential in making the funding decisions than funding agency officials. In fact, however, only one agency involved in the CAPR project, the European Commission, actually allows reviewers to make funding decisions (within budgetary limits). At all other CAPR agencies, reviewers merely have input into funding decisions that are made by agency officials.

EC respondents (and NSERC) were markedly more open to the idea of international peers. This is hardly surprising, given the international character of the European Union. EC respondents were also more open to influence by private-sector researchers. NSF respondents were more inclined than other agency respondents to favor influence by funding agency officials. This, too, is not surprising, since NSF employs a strong program officer model (modeled after the US Office of Naval Research). Perhaps more surprising is that non-academic participants—that is, participants with no significant experience in the peer review process—respond quite similarly to the 'experts,' with the exception of a greater emphasis on prospective users of the research and less reliance on the expertise of private sector researchers.

In general, according to participants, the further one is from the research field or from the Academy, the less involved one ought to be in the funding decision. Academic researchers from a dissimilar field were ranked above private sector researchers; and funding agency officials, who are generally trained in the same or similar field to which academic researchers apply for funding, were ranked above both. Potential users of the research were considered on the border between a peer and a non-peer. Non-agency government officials and members of the lay public were essentially not

considered peers at all. To emphasize: these are normative claims, with participants responding to questions regarding who ought to count as a peer.

**Table 2.** Who should count as a peer? *Mean score; standard error in parentheses*

| | **NSF** | **NIH** | **NSERC** | **EC** | **Non-academics** |
|---|---|---|---|---|---|
| Academics from the same or similar field as the applicant | 4.8 (0.7) | 4.7 (0.9) | 5.0 (0.9) | 4.8 (0.8) | 4.6 (1.0) |
| International academics | 3.0 (1.0) | 3.1 (1.0) | 4.0 (0.8) | 4.5 (1.0) | 3.4 (1.1) |
| Funding agency officials | 4.2 (1.5) | 3.2 (1.5) | 2.6 (1.3) | 2.7 (1.3) | 2.8 (1.3) |
| Academics from a dissimilar field | 3.0 (1.0) | 3.1 (1.0) | 2.8 (1.3) | 3.3 (1.2) | 2.9 (1.0) |
| Researchers employed by the private sector | 2.7 (1.1) | 2.8 (1.0) | 3.0 (1.2) | 3.4 (1.2) | 2.6 (1.0) |
| Prospective 'users' of the research | 2.3 (1.1) | 2.1 (1.0) | 2.0 (1.2) | 2.6 (1.3) | 2.7 (1.2) |
| Non-funding-agency government officials | 1.8 (0.9) | 1.6 (0.7) | 1.6 (1.1) | 1.9 (0.8) | 1.8 (0.8) |
| Members of the lay public | 1.5 (0.8) | 1.7 (0.8) | 1.6 (0.9) | 1.5 (0.8) | 1.7 (0.8) |
| Members of the international lay public | 1.4 (0.7) | 1.4 (0.6) | 1.3 (0.6) | 1.5 (0.7) | 1.5 (0.7) |

*Note*: U.S. agencies (especially NSF) appear to favor a relatively strong input from Funding Agency Officials. However, note also higher variability in responses, even among NSF participants (standard error = 1.5).

*Scale*: 6 Should make funding decision; 5 Should dominate funding decision; 4 Should influence funding decision; 3 Should probably participate; 2 Should witness but not participate; 1 Should not be involved at all.

*How does society benefit from research?*

Participants were asked to rate (on a 7-point Likert scale) their agreement with nine short narrative descriptions of the relationship between funded research and societal impacts.

The nine narratives, in order of decreasing agreement, are in Table 3.

**Table 3.** How does society benefit from research? *Mean score; standard error in parentheses*

| | NSF | NIH | NSE RC | EC | Non-academics |
|---|---|---|---|---|---|
| By spending money on research, we create a 'BODY OF KNOWLEDGE' that can be used for further research. This is necessary for the application of science, and the development of new technologies, which improve human lives. | 6.7 (0.6) | 6.7 (0.7) | 6.7 (0.9) | 6.6 (0.8) | 6.6 (0.9) |
| Scientific research leads to advanced technology, which impacts a country's ECONOMY; the financial success of the country benefits its citizens. | 6.2 (0.9) | 6.3 (1.0) | 6.0 (1.2) | 6.3 (0.8) | 6.2 (1.2) |
| Creating a 'body of research' is a good idea. However, the connections to benefits for society are rarely predictable. Research allows for the SERENDIPITOUS discovery of solutions to problems the researcher may not be aware of. | 5.9 (1.4) | 6.2 (1.2) | 6.1 (1.2) | 5.6 (1.4) | 5.8 (1.2) |
| Society benefits INDIRECTLY. By understanding more about the world, people's perspective broadens. This helps them make better decisions and be better citizens, to everyone's benefit. | 5.3 (1.4) | 5.4 (1.6) | 5.3 (1.4) | 5.4 (1.4) | 5.5 (1.4) |
| The greatest benefits of research occur because the immediate effects (e.g., education, trained professionals, new technology) make a country more COMPETITIVE in a GLOBALIZED world. | 5.2 (1.5) | 5.1 (1.5) | 5.1 (1.6) | 5.1 (1.6) | 5.0 (1.6) |
| Technological development is INEVITABLE; society must adapt. | 4.5 (1.8) | 4.6 (1.9) | 4.3 (1.8) | 4.1 (1.9) | 4.6 (1.8) |
| All research impacts society one way or the other; the real challenge is to LIMIT possible NEGATIVE impacts. | 3.8 (1.7) | 3.7 (1.7) | 3.3 (1.5) | 4.1 (1.8) | 4.2 (1.7) |
| In order to generate actual societal benefits, we must adopt a PROBLEM-CENTERED approach, instead of generating new information for its own sake. | 3.3 (1.6) | 3.3 (1.8) | 2.8 (1.5) | 4.4 (1.9) | 4.0 (1.8) |
| There ARE NO significant BENEFITS to society from public funding of scientific/technical research. | 1.2 (0.7) | 1.2 (0.9) | 1.3 (1.1) | 1.4 (1.1) | 1.4 (1.0) |

*Note*: As agreement with narratives decreases, standard deviations increase (except for the narrative stating scientific research has no social benefit).

*Scale*: 7 Strongly agree; 4 Neutral / No opinion; 1 Strongly disagree.

In short, Blue-Skies-compatible narratives dominate. CAPR participants continue to place their faith in serendipity and indirect benefits far above problem-centered approaches to funding research. Economic benefits and global competitiveness are also agreed-upon benefits, though the first is apparently more palatable than the second. Participants were essentially neutral on technological determinism and the proposition that negative impacts must be limited. The most surprising result was how low the problem-centered approach was rated by participants identified with NIH – as low as those associated with NSF, which is a basic research funding organization. Again, non-academics' normative opinions follow similar patterns, though they are on the high end of the spectrum both for problem-centered approaches and believing that *negative* impacts must be minimized.

**Table 4.** Assessing intrinsic and instrumental value. *Mean score; standard error in parentheses*

| | **NSF** | **NIH** | **NSERC** | **EC** | **Non-academics** |
|---|---|---|---|---|---|
| Peer review procedures for the consideration of the INTRINSIC value of research need to be improved. | 4.8 (1.6) | 5.3 (1.5) | 4.1 (1.9) | 4.8 (1.6) | 5.1 (1.6) |
| Peer review procedures for the consideration of the INSTRUMENTAL value of research need to be improved. | 4.2 (1.6) | 4.4 (1.7) | 3.8 (1.6) | 4.8 (1.7) | 4.8 (1.6) |
| I think the INTRINSIC value of research is extremely important, and should be considered when funding decisions are made. | 6.1 (0.9) | 5.9 (1.2) | 6.0 (1.2) | 5.7 (1.2) | 5.8 (1.1) |
| I think the INSTRUMENTAL value of research is extremely important, and should be considered when funding decisions are made. | 4.5 (1.6) | 4.7 (1.8) | 4.1 (1.7) | 4.8 (1.6) | 4.9 (1.6) |
| INTRINSIC value cannot be measured. | 3.2 (1.6) | 3.0 (1.7) | 3.4 (1.9) | 3.0 (1.7) | 3.6 (1.7) |
| INSTRUMENTAL value cannot be measured. | 3.2 (1.5) | 3.3 (1.7) | 3.0 (1.6) | 3.0 (1.6) | 3.2 (1.4) |
| I would feel comfortable assessing the INTRINSIC value of research. | 6.0 (1.2) | 6.0 (1.2) | 5.7 (1.8) | 5.8, (1.4) | 5.7 (1.3) |
| I would feel comfortable assessing the INSTRUMENTAL value of research. | 5.2 (1.5) | 5.4 (1.5) | 4.8 (1.8) | 5.3 (1.5) | 5.3 (1.5) |

*Note*: Unlike Tables 2 and 3, statements in Table 4 are not organized (i.e., listed in order of agreement.)

*Scale*: 7 Strongly agree; 4 Neutral / No opinion; 1 Strongly disagree.

*Intrinsic and instrumental value*

The survey included 13 questions that related to intrinsic and instrumental values as they relate to peer review. The first eight questions are pairs asking for comparisons between intrinsic and instrumental value (see Table 4). In essence, we asked participants to consider the appropriate roles of intrinsic value and instrumental value in funding decisions. We also asked how comfortable participants would feel with *assessing* intrinsic and instrumental merit.

We expected much more divergence between participants' attitudes toward intrinsic and instrumental value, skewed in favor of intrinsic value. Although participants do show some tendency in this direction, feeling that intrinsic value is more important than instrumental and feeling more comfortable assessing intrinsic than instrumental value, the differences were striking in how small they were. Participants feel that both intrinsic and instrumental value can be measured, and they feel comfortable assessing either (albeit somewhat more comfortable with assessing intrinsic value). They also think intrinsic value should be considered in funding decisions, but evince more ambivalence toward considerations of instrumental value. Interestingly, non-academics were slightly less confident that intrinsic value could be measured than were academics. This coincides with patterns above, for example emphases on the users of research as peers in the peer review process and a high emphasis on problem-centered research strategies, both of which suggest more concern with instrumental value.

In the second set of five questions (see Table 5), we asked about procedural issues concerning the consideration of intrinsic and instrumental value in the peer review process. We also asked about two specific instrumental justifications for research (economic benefit and social justice).

Three of the statements that participants largely disagree with have something in common: "Research should be justified by social justice," "Research should be justified by greater economic benefit," and that funding agencies should use a standard weight between intrinsic and instrumental value all imply an *a priori* decision regarding *what kinds* of value should have precedence. In contrast, the two statements that suggest *ad hoc* procedures (instrumental value can only be determined "case by case" and varies by discipline) were received favorably. In short, this second set of responses seems to show that participants wish to retain autonomy in assessing value and are less well-disposed to top-down imposition of value judgments regarding intrinsic and instrumental value. One exception is the EC respondents' attitudes toward the question of predetermined weighting for intrinsic and instrumental value. Although still on the negative side of neutral, EC respondents were quite a bit less negative than respondents from the other agencies. EC does in fact predetermine weights for impact (see Table 1). Unlike other participants EC respondents did not disagree with the idea of justifying research on the grounds of social justice.

**Table 5.** Importance of intrinsic and instrumental value. *Mean score; standard error in parentheses*

| | NSF | NIH | NSERC | EC | Non-academics |
|---|---|---|---|---|---|
| The importance of INSTRUMENTAL value relative to the INTRINSIC value can only be determined case by case. | 5.6 (1.3) | 5.8 (1.4) | 5.8 (1.3) | 5.7 (1.3) | 5.9 (1.2) |
| The importance of INSTRUMENTAL value relative to the INTRINSIC value varies by academic discipline. | 5.8 (1.4) | 5.6 (1.5) | 5.8 (1.4) | 5.6 (1.7) | 5.6 (1.5) |
| Research should be justified by greater social justice. | 3.0 (1.6) | 3.1 (2.0) | 2.6 (1.6) | 3.9 (1.9) | 3.5 (1.6) |
| Research should be justified by greater economic benefit. | 2.5 (1.5) | 2.6 (1.6) | 2.2 (1.5) | 3.3 (1.9) | 2.8 (1.7) |
| Funding agencies should use a standard weight of INSTRUMENTAL to INTRINSIC value in peer review panels. | 2.2 (1.3) | 2.3 (1.5) | 2.1 (1.4) | 3.5 (1.9) | 2.4 (1.4) |

*Scale*: 7 Strongly agree; 4 Neutral / No opinion; 1 Strongly disagree.

## 5. Discussion

The science-society relation is the subject of many narratives that both express and shape our attitudes toward public funding for scientific research. What these narratives share is the fact that each offers an account of the way that science affects society and society affects science – an account that is almost always framed in terms of scientific autonomy and societal accountability. Some narratives emphasize scientific autonomy, suggesting that the best way for society to benefit from scientific research is not to interfere with it [Humboldt 1970]. In addition to advocating autonomy, such Blue-Skies-compatible narratives often incorporate the Aristotelian notion that what is good for its own sake (intrinsic value) is better than what is good for the sake of something else (instrumental value). These narratives also often speak in terms of benefits to society, as if scientists somehow bestow gifts upon society, rather than meeting societal needs. This latter, perhaps rather more servile, description of the science-society relation tends to be featured in narratives that emphasize accountability over autonomy.

Peer review is a sort of self-policing. As such, it is meant not only to ensure accountability within the scientific community, but it also to promote both the autonomy of the scientific community and

the accountability of the scientific community to society [Holbrook 2010a; cf. Hackett 1997]. For the past two centuries, it has been taken largely for granted that academic disciplines define peers. With the rise of interdisciplinarity, however, this assumption began to be questioned. The introduction of societal impacts considerations as part of the peer review of grant proposals raises the stakes even further. If 'who ought to count as a peer' is now a legitimate question, our answers reveal another important facet of our conception of the science-society relation.

As we discussed above, people involved in peer review are difficult to access through research such as ours; this is a busy group with little motivation to participate in external examinations of their agencies' activities. We are not confident that our sample is representative – in particular, we considered the possibility that our results would necessarily be biased against peer review. That is, we anticipated that, of those we approached, the ones most dissatisfied with peer review, most frustrated with their own experiences, would be more likely to participate. This makes the relatively conservative responses – those affirming the traditional notion of a peer – all the more striking.

According to CAPR survey participants, disciplinary peers should be considered the most important in deciding what research should be funded – above funding agency officials, far above potential users of the research, and even farther above members of the lay public. Given repeated accounts of resistance to the inclusion of impact criteria as part of peer review, this result is hardly surprising. Any agency considering extending peer review to include users or members of the lay public should have it firmly in mind.

It is worth noting, however, that some agencies (including STW) already use such extended peer review. Another agency that reports good experiences with extended peer review is the US Congressionally Directed Medical Research Program, which has involved over 1500 non-scientist 'consumers' as part of scientific peer review since 1992. These agencies deserve further research. Research on interdisciplinary peer review may also be relevant to the transdisciplining of peer review. It would also be interesting to connect empirical research on extended peer review with research on knowledge utilization [for instance, Landry et al. 2001] that suggests that a research focus primarily on user needs does not in fact facilitate knowledge utilization.

Our results suggest that the desire for scientific autonomy trumps the demand for societal accountability, according to CAPR participants (mostly scientists). Narratives compatible with a Blue Skies agenda were most popular with survey participants. Participants felt that the benefits, including economic benefits, of scientific research generally outweighed risks and justified public funding for science – even as they appealed to serendipity to connect research results to societal needs. Problem-centered approaches to funding research were considered to be almost as disagreeable as the idea that funding science did not benefit society at all. In general, then, participants seemed attracted to narratives that emphasized autonomy and resistant to narratives that emphasized direction by forces outside the scientific community. In addition, participants generally also disliked predetermined weightings for intrinsic and instrumental value, indicating resistance to the establishment of a procedural standardization that would restrict their freedom in assessing grant proposals.[6]

There is also some evidence that participants would prefer placing more emphasis on intrinsic than instrumental value. Not only do they think that intrinsic value should have more influence on funding decisions than instrumental value, but they also favor intrinsic-compatible narratives (for instance, that research should contribute to a 'body of research' and that benefits accrue 'indirectly')

[6] As noted above, EC respondents are an exception to this claim.

above instrumental ones (such as a 'problem-centered' approach to research funding). This result is consistent with the results of a study reported by Bozeman and Boardman [2009: 190] which suggested that societal benefit is relatively unimportant to academic researchers. The same study also suggested that most researchers felt the government played too great a role in setting research priorities, indicating a desire for autonomy. The connection between placing a higher value on intrinsic than instrumental value and the desire for autonomy deserves more thought, as does the connection between autonomy and disciplinary expertise.

Given reports of resistance to the inclusion of societal impacts considerations in peer review, as well as the fact that CAPR survey participants expressed a clear inclination toward intrinsic value, the most surprising result of the survey was that participants indicated that they feel comfortable assessing instrumental value – almost as comfortable as they do in assessing intrinsic value. This suggests that resistance to the incorporation of societal impacts considerations into the peer review process is not due to a lack of confidence regarding their ability to assess instrumental value.

This finding is especially interesting in the context of suggestions that scientific experts lack expertise relevant to assessing societal impacts. As Bozeman and Boardman [2009: 189] put the point, "Why is the scientist who does research on the genetics of grasses any more qualified to judge social good than the person who mows the grass?" Bozeman and Boardman take this point about expertise together with the fact that scientists do not seem to care about societal impacts to support the claim that agencies ought not to ask scientists to judge societal impacts. But even if it is true that such expertise is lacking among scientific peers, it does not seem to affect their confidence in their ability to assess societal impacts. Thus, the scientific community's resistance to including societal impact considerations as part of peer review seems not to be based on a perceived inability to judge societal impacts. Instead, it seems, scientists simply do not *want* to address societal impacts considerations. They would prefer that proposals be judged on their intrinsic merit by disciplinary peers, trusting the fact that funding good science will somehow lead to societal benefits.

Bozeman and Boardman suggest that scientists lack any special expertise in societal impacts. There are at least two problems with this account. First, it unreasonably limits expertise to narrow academic realms, as if the boundary between research and the world were impermeable. The fact is, research on the genetics of grasses may well imbue the researcher with an appreciation of grasses outside the laboratory, including their potential uses in a variety of applications. *Prima facie*, scientific expertise about grasses *does* seem more relevant to considering the broader societal impacts of scientific research about grasses than expertise in mowing the grass. Second, their comparison with "the person who mows the grass" suggests that we all possess the same claim on being able to answer questions about the social good, about which there can be no question of expertise. This reduces answers to questions about the social good to mere expressions of subjective preference. But issues dealing with the social good are not mere matters of taste; asking whether it is good for society not to discriminate on the basis of race or gender is not the same as asking whether spinach tastes good. The key difference between questions about the social good and questions of taste is that we can present arguments about the social good. Arguments about matters of pure opinion, however, are illegitimate. It therefore seems premature to assume that scientific expertise is irrelevant to considering the societal impacts of science.

The claim that we cannot expect scientists to address the societal relevance of their research also gains credence based on the assumption that assessing broader societal impacts is somehow much more difficult than assessing intrinsic intellectual merit [Holbrook and Frodeman 2011].

To date, however, this appears to be an untested hypothesis: we have found no study that tests whether assessing societal impacts is in fact more difficult than assessing intrinsic merit. Moreover, CAPR survey participants indicate that they *would* feel comfortable assessing the instrumental value of research. Not only can we find no evidence that assessing impacts *is* in fact too difficult, but also the CAPR survey provides some evidence that assessing impacts may *not* be too difficult. Again, whether assessing potential societal impacts is more difficult than assessing potential intellectual merit is a question that deserves further research, not an assumption that everyone should accept at face value.

Agencies working to improve the incorporation of societal impacts considerations into their own peer review processes ought not to be overly distracted by ubiquitous claims that impact criteria themselves are unclear. Instead, they should focus not only on reviewers' explicit areas of scientific expertise, but also their implicit assumptions about the science-society relation. Agencies should consider new ways of making those implicit assumptions explicit. Agencies should also strongly consider making their own views about the science-society relation explicit. For example, the Research Councils UK (RCUK) has a one-page document titled "RCUK Expectations for Societal and Economic Impact" that details RCUK expectations in terms of accountability in return for the "considerable **flexibility** and **autonomy**" given to funding recipients" [RCUK (2010), emphasis in original]. This approach – not only articulating the agency's accountability expectations, but also tying those expectations to the agency's granting of autonomy to those it funds strikes us as a promising strategy.

Scientists who resist the inclusion of impact considerations as part of peer review ought to reconsider their approach – assuming, that is, they value their autonomy. Currently, scientists are being asked to propose and judge impacts. To offer a twist on RCUK's (and other agencies') granting of autonomy regarding the conduct of the research they fund, the fact that scientists are asked to propose and judge impacts grants them considerable flexibility and autonomy regarding the consideration of impacts in peer review. Scientists who resist including impacts as part of peer review ought to consider the possibility that if *scientists* do not determine impacts, *someone else* will determine impacts for them. Determination to resist impacts may well undermine scientific autonomy [Frodeman and Holbrook 2011; Holbrook and Frodeman 2012].

The CAPR survey suggests that incorporating societal impacts considerations into the peer review process really is more a philosophical issue or a matter of cultural adaptation than an issue of rewriting criteria to clarify or redefine 'impact'. Narratives help orient us in the world, guiding our interpretations of people, events, and relations between institutions. Perhaps the most promising area for further research would be to come up with better – and more socially robust – narratives of the science-society relation.

## Funding

This work was supported by the US National Science Foundation under Grant No. 0830387. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

## Acknowledgements

We thank Robert Frodeman for his helpful guidance as Principal Investigator of the CAPR project and for his comments on the paper. We thank Kelli Barr for help in conducting the literature review and putting together Table 1.

## References

APS News (2007) 'NSF's "broader impacts" criterion gets mixed reviews', *American Physical Society News* **16**. Available at http://www.aps.org/publications/apsnews/200706/nsf.cfm, last accessed 8 June 2011.

Bhattacharya, Ananyo (2012) 'Science Funding: Duel to the Death', *Nature* **488**. Available at http://www.nature.com/news/science-funding-duel-to-the-death-1.11073, last accessed 19 March 2013.

Boden, Margaret, Ash, Eric, Edge, David, Reece, Charles, Skehel, John and Williams, Peter (1991) 'Peer review: A report to the advisory board for the research councils from the working group on peer review', *Medical Research Council of England*. Available at http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003951, last accessed 8 June 2011.

Bornmann, Lutz, Mutz, Ruediger and Hans-Deiter, Daniel (2007) 'Gender differences in grant peer review: A meta-analysis', *Journal of Informetrics*, **1**: 226-238.

Bornmann, Lutz (2012) 'Measuring the societal impact of research', EMBO Reports, **13**(8): 673-676.

Bornmann, Lutz (2013) 'What is societal impact of research and how can it be assessed? A literature survey'. Journal of the American Society of Information Science and Technology, **64**(2), 217-233.
Bozeman, Barry and Boardman, Craig (2009) 'Broad impacts and narrow perspectives: Passing the buck on science and social impacts', *Social Epistemology*, **23**: 183-198.

Burggren, Warren W. (2009) 'Implementation of the National Science Foundation's "broader impacts": Efficiency considerations and alternative approaches', *Social Epistemology*, **23**: 221-237.

Bush, Vannevar (1945) 'Science: The Endless Frontier', US Government Printing Office, Washington, DC. Available at http://www.nsf.gov/od/lpa/nsf50/vbush1945.htm, last accessed 8 June 2011.

CAPR Midterm Workshop Report. Available at http://csidcapr.unt.edu/fedora/repository/capr:675/OBJ/, last accessed 8 June 2011.

Chubin, Daryl E. and Hackett, Edward J. (1990) *Peerless Science: Peer Review and U.S. Science Policy*. Albany: State University of New York Press.

Cole, Jonathan R. and Cole, Stephen (1981) *Peer Review in the National Science Foundation: Phase Two of a Study*. Washington, DC: National Academy Press.

Cole, Stephen, Rubin, Leonard and Cole, Jonathan R. (1978) *Peer Review in the National Science Foundation: Phase One of a Study*. Washington, DC: National Academy Press.

Edgerton, David (2009) 'The "Haldane Principle" and other invented traditions in science policy', *History and Policy*. Available at http://www.historyandpolicy.org/papers/policy-paper-88.html#S7, last accessed 8 June 2011.

Eigenbrode, Sanford D., O'Rourke, Michael, Wulfhorst, J. D., Althoff, David M., Goldberg, Caren S., Merrill, Kaylani, Morse, Wayde, Nielsen-Pincus, Max, Stephens, Jennifer, Winowiecki, Leigh and Bosque-Pérez, Nilsa A. (2007) 'Employing philosophical dialogue in collaborative science', *BioScience*, **57**: 55-64.

Frodeman, Robert and Holbrook, J. Britt (2007) 'Science's social effects', *Issues in Science and Technology*, **23**: 28-30.

Frodeman, Robert and Holbrook, J. Britt (2011) 'NSF's struggle to articulate relevance', *Science*, vol. 333, July 8, 2011: 157-58.

Frodeman, Robert (2010) 'Experiments in field philosophy', *New York Times*, 23 November. Available at http://opinionator.blogs.nytimes.com/2010/11/23/experiments-in-field-philosophy/, last accessed 8 June 2011.

Frodeman, Robert and Briggle, Adam (2012) 'The dedisciplining of peer review', *Minerva*. DOI 10.1007/s11024-012-9192-9198.

Frodeman, Robert, Briggle, Adam, and Holbrook, J. Britt (2012) 'Philosophy in the Age of Neoliberalism', *Social Epistemology: A Journal of Knowledge, Culture, and Policy*. 26 (2-4). DOI**:** 10.1080/02691728.2012.722701.

Funtowicz, Silvio O. and Ravetz, Jerome R. (1993) 'Science for the post-normal age', *Futures*, **25**: 739-755.

Gorman, Michael E. (Ed.) (2010) *Trading Zones and Interactional Expertise: Creating New Kinds of Collaboration*. Boston: Massachusetts Institute of Technology Press.

Hackett, Edward J. (1997) 'Peer review in science and science policy'. In: Frankel, M.S. and Cave, J. (eds.) *Evaluating Science and Scientists: An East-West Dialogue on Research Evaluation in Post-Communist Europe*, pp. 51-60. Budapest: Central European University Press.

Healy, Stephen (1999) 'Extended peer communities and the ascendance of post-normal politics', *Futures*, **31**: 655-669.

Holbrook, J. Britt (2005) 'Assessing the science – society relation: the case of the U.S. National Science Foundation's second merit review criterion', *Technology in Society*, **27**: 437-451.

Holbrook, J. Britt (2010a) 'Peer review'. In: Frodeman, R., Klein, J.T. and Mitcham, C. (eds.) *The Oxford Handbook of Interdisciplinarity*, pp. 321-332. Oxford: Oxford University Press.

Holbrook, J. Britt (2010b) 'The use of societal impacts considerations in grant proposal peer review: A comparison of five models', *Technology and Innovation – Proceedings of the National Academy of Inventors*, **12**: 213-224.

Holbrook, J. Britt and Frodeman, Robert (2011) 'Peer review and the *ex ante* assessment of societal impacts', *Research Evaluation*, **20**(3): 239–246. DOI: 10.3152/095820211X12941371876788.

Holbrook, J. Britt (2012a) 'What is interdisciplinary communication? Reflections on the very idea of disciplinary integration', *Synthese*. DOI: 10.1007/s11229-012-0179-7.

Holbrook, J. Britt (2012b) 'Re-assessing the science - society relation: The case of the US National Science Foundation's broader impacts merit review criterion (1997 - 2011)', in *Peer Review, Research Integrity, and the Governance of Science – Practice, Theory, and Current Discussions*. Robert Frodeman, J. Britt Holbrook, Carl Mitcham, and Hong Xiaonan. Beijing: People's Publishing House: 328 – 62. [Denton, Texas]. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc77119/. Accessed November 1, 2012.

Holbrook, J. Britt and Robert Frodeman (2012) 'Resistance to impact criteria can lead to a tightening of the accountability noose', London School of Economics *Impact of Social Sciences Blog*. March 15, 2012.

Von Humboldt, Wilhelm (1970) 'On the spirit and the organizational framework of intellectual institutions in Berlin', *Minerva*, **8**: 242-250.

Johnson ,Valen E. (2008) 'Statistical analysis of the National Institutes of Health peer review system', *Proceedings of the National Academy of Sciences of the United States of America*, **105**: 11076-11080.

Kamenetzky, Julia R (2013) 'Opportunities for impact: Statistical Analysis of the National Science Foundation's broader impacts criterion' *Science and Public Policy*, **40** (1): 72-84. DOI: 10.1093/scipol/scs059.

Kitcher, Philip (2001) *Science, Truth, and Democracy*. New York: Oxford University Press.

Lamont, Michèle and Mallard, Grégoire (2005) 'Peer review in international perspectives: US, UK and France. Report commissioned by the Social Sciences and Humanities Research Council of Canada.' Available at http://www.wjh.harvard.edu/~mlamont/SSHRC-peer.pdf, last accessed 8 June 2011.

Lamont, Michèle (2009) *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, MA: Harvard University Press.

Landry, Réjean, Amara, Nabil and Lamari, Moktar (2001) 'Climbing the ladder of research utilization: Evidence from social science research', *Science Communication*, **22**: 396–422.

Langfeldt, Liv (2001) 'The decision-making constraints and processes of grant peer review, and their effects on the review outcome', *Social Studies of Science*, **31**: 820–841.

Langfeldt, Liv (2004) 'Expert panels evaluating research: decision making and sources of bias', *Research Evaluation*, **13**: 51–62.

Laudel, Grit (2006) 'Conclave in the Tower of Babel: how peers review interdisciplinary research proposals', *Research Evaluation*, **15**: 57-68.

Lightbourne, Jim (2010) 'NSF merit review criteria and process'. Available at http://csid-capr.unt.edu/fedora/repository/capr:354/OBJ/, last accessed 8 June 2011.

Lok, Corie (2010) 'Science funding: Science for the masses', *Nature 465*, 416-418. Available at http://www.nature.com/news/2010/100526/full/465416a.html, last accessed 8 June 2011.

Mallard, Grégoire, Lamont, Michèle and Guetzkow (2009) 'Fairness as appropriateness: Negotiating epistemological differences in peer review', *Science, Technology and Human Values*, **34**(5), 573-606.
Marsh, Herbert W., Jayasinghe, Upali W. and Bond, Nigel W. (2008) 'Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability', *American Psychologist*, **63**: 160-68.

Nightingale, Paul and Scott, Alister (2007) 'Peer review and the relevance gap: ten suggestions for policymakers', *Science and Public Policy*, **24**: 543-53.

Nowotny, Helga, Scott, Peter and Gibbons, Michael (2001) *Re-thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, MA: Polity Press.

O'Rourke, Michael and Crowley, Stephen J. (2012) 'Philosophical intervention and cross-disciplinary science: the story of the Toolbox Project' *Synthese*. DOI: 10.1007/s11229-012-0175-y.

Peters, Douglas P. and Ceci, Stephen J. (1982) 'Peer-review practices of psychological journals: The fate of published articles, submitted again', *Behavioral and Brain Sciences*, **5**(2): 187-255.

Polanyi, Michael (1962) 'The Republic of Science: Its political and economic theory', *Minerva*, **1**: 54-73.

RCUK (2010) 'RCUK expectations for societal and economic impact'. Available at http://www.rcuk.ac.uk/documents/innovation/expectationssei.pdf, last accessed 8 June 2011.

Rip, Ari (2000) 'Higher forms of nonsense', *European Review*, **8**: 467-86.

Roberts, Melanie (2009) 'Realizing societal benefit from academic research: Analysis of the National Science Foundation's broader impacts criterion', *Social Epistemology*, **23**: 199-219.

Sarewitz Daniel (1996) *Frontiers of Illusion: Science, Technology, and the Politics of Progress*. Philadelphia: Temple University Press.

Scott, Alister (2007) 'Peer review and the relevance of science', *Futures*, **39**: 827-45.

Stokes, Donald E. (1997) *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington, D.C.: Brookings Institution Press.

Viner, Neil, Powell, Philip and Green, Rod (2004) 'Institutionalized biases in the award of research grants: A preliminary analysis revisiting the principle of accumulative advantage', *Research Policy*, **33**: 443-54.

Wennerås, Christine and Wold, Agnes (1997) 'Nepotism and sexism in peer review', *Nature*, **387**: 341-43.