

Scientific Autonomy, Public Accountability, and the Rise of “Peer Review” in the Cold War United States

Melinda Baldwin, *American Institute of Physics*

Abstract: This essay traces the history of refereeing at specialist scientific journals and at funding bodies and shows that it was only in the late twentieth century that peer review came to be seen as a process central to scientific practice. Throughout the nineteenth century and into much of the twentieth, external referee reports were considered an optional part of journal editing or grant making. The idea that refereeing is a requirement for scientific legitimacy seems to have arisen first in the Cold War United States. In the 1970s, in the wake of a series of attacks on scientific funding, American scientists faced a dilemma: there was increasing pressure for science to be accountable to those who funded it, but scientists wanted to ensure their continuing influence over funding decisions. Scientists and their supporters cast expert refereeing—or “peer review,” as it was increasingly called—as the crucial process that ensured the credibility of science as a whole. Taking funding decisions out of expert hands, they argued, would be a corruption of science itself. This public elevation of peer review both reinforced and spread the belief that only peer-reviewed science was scientifically legitimate.

By the time this essay reaches print, it will have passed through a gauntlet familiar to scholars in every discipline. The editor of this journal will have sent it to two or three anonymous referees. If the referees deemed the contents potentially worthy of publication (and this article could have been rejected once, twice, or many times before finding a home), they will have offered comments on how it should be improved before publication. I will have addressed those critiques in subsequent revisions, strengthening the article’s evidence and arguments until the editor and referees agree that the most serious objections have been answered. The final

Melinda Baldwin is the Books Editor at *Physics Today*, a magazine published by the American Institute of Physics. Her research focuses on the history of scientific communication and scholarly practices. She is the author of *Making “Nature”: The History of a Scientific Journal* (Chicago, 2015). 1 Physics Ellipse, College Park, Maryland 20740, USA; mbaldwin@aip.org.

Acknowledgments. Research for this project was supported by the American Academy of Arts and Sciences, the American Council of Learned Societies, and Harvard University’s Department of the History of Science. I would like to thank Michael Barany, Evan Hepler-Smith, Aileen Fyfe, Michael Gordin, Erika Milam, the attendees at the November 2017 “Funding Bodies in Late Modern Science” workshop, the members of the Consortium for History of Science, Technology, and Medicine’s Physical Sciences Working Group, and the referees for productive questions, suggestions, and encouragement.

Isis, volume 109, number 3. © 2018 by The History of Science Society.

All rights reserved. 0021-1753/2018/0109-0005\$10.00.

538

result is a research article in a peer-reviewed journal—a piece of writing that will count for more professional credit, and hold more scholarly credibility, than work posted on my personal website or uploaded to a preprint server without external review.

Peer-reviewed writing and grants are given a special status in most scholarly communities, and this is perhaps nowhere more true than in the sciences, where findings that have not faced the test of peer review are usually seen as preliminary or unreliable. My favorite example of this attitude toward peer review has long been an evocative headline about a 2012 *Physics Letters B* paper: “CERN’s Higgs Boson Discovery Passes Peer Review, Becomes Actual Science.”¹

The most widely accepted story about peer review’s origin credits Henry Oldenburg with inventing it for the seventeenth-century *Philosophical Transactions of the Royal Society*, creating the impression that refereeing has been an unchanging part of science for over three hundred years. However, new historical work is beginning to shed more light on peer review’s development—and the real story is far more complicated than the neat tale of Oldenburg inventing refereeing out of whole cloth during the Scientific Revolution.

Most existing histories of peer review have focused on the emergence of the scientific referee during the nineteenth century or on the inner workings of referee systems at particular journals. Those studies have shown that refereeing was not initially thought of as a process that bestowed scientific credibility and that many high-profile journals and grant organizations had unsystematic (or nonexistent) refereeing processes well into the twentieth century. So how did we get to our current moment, in which peer review is thought to be so crucial to science that new findings aren’t “actual science” without it?

In this essay, I argue that the vision of peer review as a process central to science can be traced to the Cold War United States, where various stakeholders sought to navigate a growing tension between desires for scientific autonomy and public accountability in controversies over government science funding. I begin by tracing the development of refereeing systems at scientific journals, the institutions where scholars have done the most work on refereeing’s history. From there, I explore the development of refereeing at funding bodies, where comparatively little historical work has been done. Despite important differences between refereeing procedures at the two types of institutions, refereeing at both journals and grant organizations became more heavily emphasized in the late 1960s and 1970s, and I show that the Cold War United States is the crucial context for understanding that increased emphasis.

The massive expansion of government funding for scientific research in the postwar United States led to more scrutiny of science—and to suggestions that scientists should be more accountable to the public. Scientists balked at the suggestion that their methods or conclusions might be vetted by scientific laymen but did not want to surrender the public status or funding opportunities they had gained. A controversy over peer review procedures at the National Science Foundation (NSF) in the 1970s highlighted the conflict between scientists’ desire to keep scientific decisions in expert hands and the belief that public funds made scientists answerable to laymen and legislators. At a hearing about the NSF’s review procedures, various stakeholders argued that peer review was the only acceptable method of selecting which proposals to fund, and the outcome of the controversy placed increased emphasis on referee opinions at the NSF. The episode both reflected and helped cement the view that peer review was central to the proper practice of science.

¹ Sebastian Anthony, “CERN’s Higgs Boson Discovery Passes Peer Review, Becomes Actual Science,” *ExtremeTech*, Ziff Davis, Inc., 2012, <http://www.extremetech.com/extreme/135756-cerns-higgs-boson-discovery-passes-peer-review-becomes-actual-science> (accessed 29 July 2015).

REFEREING IN JOURNALS THROUGH WORLD WAR II

Because refereeing is so important to scientific and scholarly work, over the past several decades many scientists, physicians, journalists, and sociologists have written papers and books analyzing the contemporary state of peer review. Modern economists, sociologists, and philosophers of science have continued this investigation of how peer review functions within the scientific community, and there has been some excellent recent work on journal refereeing. The sociologist Joanne Gaudet has studied the methods and assumptions of peer review in journals, arguing that peer review systems are shaped by disciplinary conditions and economic contexts. The philosopher Carole Lee and her collaborators have undertaken a study of the functions of bias in peer review and have provocatively argued that it may not be possible or desirable to eliminate biases.²

Interestingly, peer review’s best-known historical origin story was penned not by a historian but by a pair of sociologists. In their 1971 paper “Patterns of Evaluation in Science: Institutionalization, Structure, and Functions of the Referee System,” Harriet Zuckerman and Robert Merton claimed that refereeing had its origins in the seventeenth-century Royal Society of London. Zuckerman and Merton wrote that when the Royal Society gave its Secretary, Henry Oldenburg, permission to compile the *Philosophical Transactions* in 1665, he immediately decided to gather expert opinions on the papers submitted for publication. This version of peer review’s history has been widely circulated and repeated in other scholarly papers, creating the pervasive impression that refereeing has been a part of science ever since the first scientific journal was created.³

However, more recent historical work has illustrated that Oldenburg did not employ any system resembling modern refereeing at the *Philosophical Transactions*. Other early scientific societies, such as the Académie des Sciences in France, did have procedures for evaluating the work of their members before it was circulated, but those systems of internal critique do not seem to map onto or lead into the systematic external refereeing we know today. Aileen Fyfe, Julie McDougall-Waters, and Noah Moxham have helpfully described this period as the “pre-

² Joanne Gaudet, “All That Glitters Is Not Gold: The Shaping of Contemporary Journal Peer Review at Scientific and Medical Journals,” uO Research, preprint (2014), pp. 1–23, https://www.ruor.uottawa.ca/bitstream/10393/31238/1/contemporary_shaping_peer_review.gaudet.pdf (accessed 29 July 2015); Carole Lee, “Commensuration Bias in Peer Review,” *Philosophy of Science*, 2015, 82:1272–1283; and Lee, Cassidy Sugimoto, Guo Zhang, and Blaise Cronin, “Bias in Peer Review,” *Journal of the American Society for Information Science and Technology*, 2013, 64:2–17. For other work analyzing the state of peer review see, e.g., Daryl Chubin and Edward Hackett, *Peerless Science: Peer Review and U.S. Science Policy* (Albany: State Univ. New York Press, 1990); Warren Hagstrom, *The Scientific Community* (New York: Basic, 1965); Frank Ingelfinger, “Peer Review in Biomedical Publication,” *American Journal of Medicine*, 1974, 56:686–692; Sheila Jasanoff, “Peer Review in the Regulatory Process,” *Science, Technology, and Human Values*, 1985, 10:20–32; Jasanoff, *The Fifth Branch: Science Advisers as Policymakers* (Cambridge, Mass.: Harvard Univ. Press, 1990), Ch. 4; Jasanoff, “Quality Control and Peer Review in Advisory Science,” in *The Politics of Scientific Advice: Institutional Design for Quality Assurance*, ed. Justus Lentsch and Peter Weingart (Cambridge: Cambridge Univ. Press, 2011), pp. 19–35; Bruno Latour and Steve Woolgar, *Laboratory Life: The Construction of Scientific Facts* (Princeton, N.J.: Princeton Univ. Press, 1986); Robert Merton, “Science and Technology in a Democratic Order,” *Journal of Legal and Political Sociology*, 1942, 1:115–126; David Pontille and Didier Torry, “From Manuscript Evaluation to Article Valuation: The Changing Technologies of Journal Peer Review,” *Human Studies*, 2015, 38:57–79; and Ann C. Weller, *Editorial Peer Review: Its Strengths and Weaknesses* (Medford, N.J.: Information Today, 2001).

³ Harriet Zuckerman and Robert Merton, “Patterns of Evaluation in Science: Institutionalization, Structure, and Functions of the Referee System,” *Minerva*, 1971, 9:66–100, esp. pp. 68–75. For repetitions of this origin story see, e.g., Von Bakanic, Clark McPhail, and Rita Simon, “The Manuscript Review and Decision-Making Process,” *American Sociological Review*, 1987, 52:631–642; Chubin and Hackett, *Peerless Science*; Robert Frodeman and Adam Briggie, “The Dedisciplining of Peer Review,” *Minerva*, 2012, 50:3–19; Ingelfinger, “Peer Review in Biomedical Publication”; Stephen Lock, “Does Editorial Peer Review Work?” *Annals of Internal Medicine*, 1994, 121:60–61; and Michael Mahoney, “Publication Prejudices: An Experimental Study in Confirmatory Bias in the Peer Review System,” *Cognitive Research and Therapy*, 1977, 1:161–175.

history” of peer review, highlighting the gaps between the general idea of having peers comment on each others’ work and the formal systems of “peer review” that would develop later.⁴

Far from springing fully grown from the Scientific Revolution, refereeing developed in the nineteenth century and spread slowly and haphazardly, encountering much skepticism and criticism along the way.⁵ While some scientific societies had procedures for orally vetting submissions to their journals as early as the eighteenth century, written reports by specialists in a paper’s field did not enter the picture until 1831, when William Whewell proposed that two Fellows of the Royal Society should write down their views on submissions to the *Philosophical Transactions* and have their comments published in the new journal *Proceedings of the Royal Society of London*.⁶

Whewell’s plan to publish referee reports was quickly abandoned, but the practice of sending papers submitted for the *Transactions* (and, later, the *Proceedings*) out for referee opinions endured and expanded. By the mid-nineteenth century, arranging refereeing was one of the chief responsibilities of the Secretaries of the Royal Society. Although Whewell had intended for his referees’ identities to be known to both the author and the journal’s readers, the Royal Society quickly decided that referees would give more candid advice if they remained anonymous. Indeed, their reports were not even sent to the paper’s author—the referee report was considered a confidential document for the internal use of the Royal Society. If one of the Royal Society Secretaries wished to communicate concerns to the author, he would use a personal letter to relay the referees’ comments without revealing their identities.⁷ Referee anonymity, therefore, was linked with the practice of refereeing very early in its history.

Over the course of the nineteenth and early twentieth centuries, an increasing number of scientific and learned societies in the United States and the United Kingdom adopted the practice of systematically consulting anonymous referees about submitted papers. Great Britain’s Geological Society and Royal Society of Chemistry, for instance, both adopted refereeing systems for their publications in the nineteenth century, while the American Physical Society and the American Sociological Association adopted them in the early twentieth century. During this period, concerns about the quality of the scientific literature led to the “referee” being reconceptualized as a gatekeeper—someone responsible for ensuring that a scientific paper merited publication. That vision of the referee as a guardian of the scientific literature led many

⁴ See, e.g., Aileen Fyfe, Julie McDougall-Waters, and Noah Moxham, “350 Years of Scientific Periodicals,” *Notes and Records: The Royal Society Journal of the History of Science*, 2015, 69:227–239; and Moxham and Fyfe, “The Royal Society and the Prehistory of Peer Review, 1665–1965,” *Historical Journal*, 2017, pp. 1–27, doi:10.1017/S0018246X17000334.

⁵ On the history of refereeing see Melinda Baldwin, “Credibility, Peer Review, and *Nature*, 1945–1990,” *Notes Rec.*, 2015, 69:337–352; John Burnham, “The Evolution of Editorial Peer Review,” *Journal of the American Medical Association*, 1990, 263:1323–1329; Burnham, “How Journal Editors Came to Develop and Critique Peer Review Procedures,” in *Research Ethics, Manuscript Review, and Journal Quality: Proceedings of a Symposium on the Peer Review-Editing Process*, ed. H. F. Maryland and R. E. Sojka (Madison, Wis.: ACS Miscellaneous Publications, 1992), pp. 55–62; Imogen Clarke, “The Gatekeepers of Modern Physics: Periodicals and Peer Review in 1920s Britain,” *Isis*, 2016, 106:70–93; Alex Csiszar, “Peer Review: Troubled from the Start,” *Nature*, 21 Apr. 2016, 532:306–308; Csiszar, “Broken Pieces of Fact: The Scientific Periodical and the Politics of Search in Nineteenth-Century France and Britain” (Ph.D. diss., Harvard Univ., 2010); Drummond Rennie, “Editorial Peer Review: Its Development and Rationale,” in *Peer Review in Health Sciences*, ed. Fiona Godlee and Tom Jefferson (London: BMJ, 1999), pp. 1–13; and Weller, *Editorial Peer Review* (cit. n. 3), pp. 3–7.

⁶ Csiszar, “Peer Review.” See also Alex Csiszar, *The Scientific Journal: Authorship and the Politics of Knowledge in the Nineteenth Century* (Chicago: Univ. Chicago Press, 2018); Csiszar, “Broken Pieces of Fact”; and Moxham and Fyfe, “Royal Society and the Prehistory of Peer Review” (cit. n. 4).

⁷ Melinda Baldwin, “Tyndall and Stokes: Correspondence, Referee Reports, and the Physical Sciences in Victorian Britain,” in *The Age of Scientific Naturalism: John Tyndall and His Contemporaries*, ed. Bernard Lightman and Michael Reidy (London: Pickering & Chatto, 2014), pp. 171–186.

journals to place refereeing largely in the hands of a small number of elite scientists, who therefore had tremendous power to shape a journal’s contents in their field of expertise.⁸

In the interwar period, a broader push toward standardization during the Progressive Era had a tremendous impact on scientific practice, particularly in the United States. The quest for standardization seems to have been one impetus that influenced the development of increasingly formal refereeing procedures at British and American scientific societies. Refereeing procedures at the American physics journal *Physical Review*, for example, became much more standardized during the 1920s and 1930s. Referees who had once written free-form letters sharing their general impressions of *Physical Review* submissions were now asked to fill out forms assessing a paper’s suitability according to a predetermined list of criteria. However, most papers accepted for *Physical Review* never went out to referees at all; the editor accepted most papers on his own authority, consulting referees only when he thought he might want to reject a paper. It was not until the 1960s that all *Physical Review* papers were sent out for external referee opinions.⁹

Commercial scientific journals such as the *Philosophical Magazine* and *Nature* largely kept editorial decisions in-house until the mid-twentieth century.¹⁰ Many of these commercial journals were the personal projects of dynamic editors who felt qualified to evaluate any contributions that might come their way. Unlike the publications of scientific societies, which were published whenever the society had enough material to print an issue, commercial journals had to meet monthly or weekly deadlines and often could not afford to reject too many contributions—or wait for external referees to submit reports.

Refereeing also remained relatively uncommon outside the English-speaking scientific world in the nineteenth and early twentieth centuries. Journals affiliated with academic institutes in France and Germany, for example, generally did not employ refereeing procedures. In 1835, when the French Académie des Sciences founded the *Comptes Rendus Hebdomadaires des Séances de l’Académie des Sciences*, the Académie deliberately eschewed refereeing in order to publish the new periodical more quickly. Germany’s most prominent journals, such as the *Annalen der Physik und Chemie*, were generally under the control of a powerful editor who preferred to make publication decisions without relying on referee opinions. Scientists unaccustomed to outside refereeing did not necessarily see the practice as a superior system for scientific journals. Famously, in 1936 Albert Einstein was incensed to find that the editor of *Physical Review* had sent his submission to another physicist for commentary. In a terse note to the editor, John Tate, Einstein wrote that he and his coauthor “had not authorized you to show

⁸ On the referee as gatekeeper see Csiszar, “Peer Review” (cit. n. 5); and Clarke, “Gatekeepers of Modern Physics” (cit. n. 5). On various societies’ adoption of refereeing systems see Csiszar, “Broken Pieces of Fact” (cit. n. 5), pp. 70–76; Roberto Lalli, “‘Dirty Work,’ But Someone Has to Do It: Howard P. Robertson and the Refereeing Practices of *Physical Review* in the 1930s,” *Notes Rec.*, 2016, 70:151–174; and Dorris West Goodrich, “An Analysis of Manuscripts Received by Editors of the *American Sociological Review* from May 1, 1944 to September 1, 1945,” *Amer. Sociol. Rev.*, 1945, 10:716–725.

⁹ Lalli, “‘Dirty Work,’ But Someone Has to Do It.” On the move toward standardization during the Progressive Era see Deborah Coen, “Standardizing the Subject: Experimental Psychologists, Introspection, and the Quest for a Technoscientific Ideal,” *Technology and Culture*, 1993, 34:757–783; Alan Derickson, “Physiological Science and Scientific Management in the Progressive Era: Frederic S. Lee and the Committee on Industrial Fatigue,” *Business History Review*, 1994, 68:483–514; Nelson Kellogg, “Gauging the Nation: Samuel Wesley Stratton and the Invention of the National Bureau of Standards” (Ph.D. diss., Johns Hopkins Univ., 1992); and Perrin Selcer, “Standardizing Wounds: Alexis Carrel and the Scientific Management of Life in the First World War,” *British Journal for the History of Science*, 2008, 41:73–107.

¹⁰ On refereeing systems at these periodicals see Melinda Baldwin, *Making “Nature”: The History of a Scientific Journal* (Chicago: Univ. Chicago Press, 2015); Baldwin, “Credibility, Peer Review, and *Nature*, 1945–1990” (cit. n. 5); and Imogen Clarke and Jim Mussell, “Conservative Attitudes to Old-Established Organs: Oliver Lodge and *Philosophical Magazine*,” *Notes Rec.*, 2015, 69:321–336.

[our manuscript] to specialists before it is printed. I see no reason to address the—in any case erroneous—comments of your anonymous expert. On the basis of this incident I prefer to publish the paper elsewhere.”¹¹

Well into the twentieth century, a reputable scientific journal could and did publish papers without consulting referees. Many journals that adopted refereeing in the mid-twentieth century did so not because of epistemic concerns but to lighten the editorial workload. At the American weekly *Science*, for instance, the editorial board performed almost all refereeing in-house for the first half of the twentieth century. The journal began relying more heavily on external referees in the 1950s after members of the editorial board complained that “the job of refereeing and suggesting revisions for hundreds of technical papers is neither the best use of their time nor pleasant, satisfying work.” Another example of a journal that adopted refereeing to ease editorial burdens is the *American Journal of Medicine* (AJM). When AJM was founded in 1946, its editor Alexander Gutman worked to ensure fast publication by handling the vast majority of acceptances or rejections himself. However, as the journal became more popular, Gutman was unable to keep up with the number of submissions. By the mid-1960s submissions to AJM were being sent out for external refereeing to ensure that the journal could maintain its reputation for fast publication.¹²

Other prominent English-language journals adopted systematic external refereeing even later. The *New England Journal of Medicine* began having two outside reviewers consider all potentially acceptable papers in the late 1960s. *Nature* began employing referees for every paper it published only in 1973.¹³ Well into the 1970s, the British medical journal the *Lancet* relied heavily on editorial judgment, with editors accepting or rejecting up to 90 percent of submissions themselves.¹⁴

Significantly, there seems to have been a growing perception that Americans placed more emphasis on systematic external refereeing than their counterparts in other nations. David Davies, the British geophysicist who edited *Nature* from 1973 to 1980, had been working at MIT when he accepted the position. He would later recall that he and his American colleagues were alarmed by *Nature*’s unsystematic attitude toward refereeing, and he felt that “getting the refereeing system beyond reproach” was his most important task as editor. By the late 1970s the editor of the *Lancet* worried that its articles might not be taken seriously in the United States unless the journal began employing peer review rather than allowing the editor to accept papers on his own authority. Despite such concerns, the *Lancet* decided to limit the influence its reviewers had over the editor. In a 1989 editorial, the journal complained that “in the United States far too

¹¹ Albert Einstein, quoted in Daniel Kennefick, “Controversies in the History of the Radiation Reaction Problem in General Relativity,” in *The Expanding Worlds of General Relativity*, ed. Hubert Goenner, Jürgen Renn, Jim Ritter, and Tilman Sauer (New York: Springer, 1999), pp. 208–234, on p. 208. Regarding practice at the *Comptes Rendus Hebdomadaires des Séances de l’Académie des Sciences* see Csizsar, “Broken Pieces of Fact” (cit. n. 5), pp. 86–95. For the German case see Lewis Pyenson, “Physical Sense in Relativity: Max Planck Edits the *Annalen der Physik*, 1906–1918,” in *The Young Einstein: The Advent of Relativity*, ed. Pyenson (Bristol: Hilger, 1985), pp. 194–214.

¹² “Minutes of the Joint Meeting of AAAS Editorial Board and Publications Committee, 10 June 1955,” Philip Abelson Papers, American Association for the Advancement of Science Archives, Washington, D.C. (hereafter cited as *Abelson Papers*), Box 1 (regarding submissions to *Science*); and Ingelfinger, “Peer Review in Biomedical Publication” (cit. n. 2) (regarding AJM).

¹³ Franz J. Ingelfinger, “The New England Journal of Medicine: Editor’s Report, 1967–1977,” *New England Journal of Medicine*, 30 June 1977, 296:1530–1535 (Ingelfinger noted that “about 20 per cent of submissions are weeded out by the associate editors and me as unacceptable for various reasons” [p. 1533]); and Baldwin, *Making “Nature”* (cit. n. 10), Ch. 6. On *Nature*’s early history see also Roy MacLeod’s rich account for *Nature*’s centenary issue, republished as “The Social Framework of *Nature*” in *The “Creed of Science” in Victorian England* (Aldershot: Ashgate, 2000), pp. 1–44.

¹⁴ Ingelfinger, “Peer Review in Biomedical Publication” (cit. n. 2), p. 687.

much is being asked of peer review” and proudly assured readers that at the *Lancet* “reviewers are advisers not decision makers.”¹⁵

REFEREEING AT FUNDING BODIES

Increased burdens on editorial boards provide one clue about why refereeing became more widely used at journals in the late twentieth century; however, that does not explain how the scientific community came to view refereeing as crucial for scientific legitimacy. To understand that shift, we must look not only at journals but at funding bodies. Sociologists have done a great deal of excellent work on refereeing systems at funding bodies, especially in the past decade. One notable project is the Comparative Assessment of Peer Review (CAPR), a four-year study of peer review procedures at six worldwide grant organizations that analyzed how peer review processes do or do not balance the scientific community’s desire for autonomy with funding bodies’ desire for relevant and accountable scientific research. Michèle Lamont’s 2009 book *How Professors Think* examined grant peer review panels in the humanities, illustrating interesting divides in what different fields consider “excellent” work, as well as illuminating the social functions peer review serves for members of the scholarly community.¹⁶

However, far less historical work has been done on the development of refereeing systems at funding bodies, perhaps because so few funding bodies used systematic refereeing procedures prior to the late twentieth century. Many funding bodies had unsystematic or internal review processes that placed heavy responsibility in the hands of organization employees. Private funding bodies such as the Rockefeller Foundation, for example, generally left funding decisions in the hands of trusted middle managers well into the postwar period, awarding money via what Robert Kohler has described as a “patronage system.”¹⁷ The same was true for many publicly funded grant organizations. The German Research Foundation, created in 1920 and initially called the Emergency Association for German Science, deliberately chose to rely on a small number of elite scientists for opinions on grant proposals, and much of the evaluation focused on the personal qualities of the applicants. Well into the twentieth century, a single three-man committee evaluated all applications for the Royal Society of London’s Government Grants; though all were invited to apply, the process awarded those grants almost exclusively to Fellows of the Royal Society.¹⁸

¹⁵ Baldwin, *Making “Nature”* (cit. n. 10), p. 180; Rennie, “Editorial Peer Review” (cit. n. 5), p. 4; and “Peers Reviewed,” *Lancet*, 1989, 33:1115–1116, on p. 1116.

¹⁶ Michèle Lamont, *How Professors Think: Inside the Curious World of Academic Judgment* (Cambridge, Mass.: Harvard Univ. Press, 2009). For work by sociologists on refereeing systems at funding bodies see Lutz Bornmann and Hans-Dieter Daniel, “Selection of Research Fellowship Recipients by Committee Peer Review: Reliability, Fairness, and Predictive Validity of Board of Trustees’ Decisions,” *Scientometrics*, 2005, 63:297–320; Bornmann and Daniel, “Selecting Scientific Excellence through Committee Peer Review: A Citation Analysis of Publications Previously Published to Approval or Rejection of Post-Doctoral Research Fellowship Applicants,” *ibid.*, 2006, 68:427–440; Liv Langfeldt, “The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome,” *Social Studies of Science*, 2001, 31:820–841; and Danielle Li and Leila Agha, “Big Names or Big Ideas: Do Peer-Review Panels Select the Best Science Proposals?” *Science*, 2015, 348:434–438. For some of the CAPR results see Frodeman and Briggie, “Dedisciplining of Peer Review” (cit. n. 3); J. Britt Holbrook, “The Use of Societal Impacts Considerations in Grant Proposal Peer Review: A Comparison of Five Models,” *Technology and Innovation*, 2010, 12:213–224; and Holbrook and Robert Frodeman, “Peer Review and the *ex ante* Assessment of Societal Impacts,” *Research Evaluation*, 2011, 20:239–246.

¹⁷ Robert E. Kohler, *Partners in Science: Foundations and Natural Scientists, 1900–1945* (Chicago: Univ. Chicago Press, 1991), pp. 68–70. See also Michael Barany, “A Postwar Guide to Winning a Science Grant,” *Physics Today Online*, 20 Mar. 2018, doi:10.1063/PT.6.4.20180320a.

¹⁸ Patrick Wagner, “‘Preserve of Full Professors’—On the History of the German Research Foundation between 1920 and 1970,” in *The German Research Foundation, 1920–1970: Funding Poised between Science and Politics*, ed. Mark Walker, Karin Orth, Ulrich Herbert, and Rüdiger vom Bruch (Berlin: Steiner, 2013), pp. 23–40; and R. M. MacLeod, “The Royal Society and the Government Grant: Notes on the Administration of Scientific Research, 1849–1914,” *Hist. J.*, 1971, 14:323–358.

Grant organizations affiliated with learned societies or with governments were most likely to utilize external refereeing, and some of them adopted it in response to Progressive Era pushes for standardization. For example, the American National Research Council, the private research arm of the National Academy of Sciences, developed increasingly formalized refereeing procedures during the 1920s and 1930s, in part to ensure fairness to researchers without established national reputations. But many government funding bodies, including ones formed in the mid-twentieth century, eschewed systematic external refereeing in favor of creating an in-house panel of researchers who would evaluate all grants. In Great Britain, the government created the Medical Research Council (MRC) in the aftermath of World War I to promote medical research. The members of the council were a mix of scientific experts and legislators, and funding decisions were left entirely up to the council members before World War II. In 1946, the MRC created an Appointments and Grants Committee to help deal with the administrative work of grant applications, but this committee was ultimately answerable to the council as well and appears to have made only occasional use of outside opinions.¹⁹

Similarly, when the U.S. government formed the National Institutes of Health (NIH) in 1948, its Division of Research Grants initially evaluated grant applications with little or no consultation from outside referees.²⁰ Instead, each application went first to a small “study section” composed of NIH-affiliated scientific experts in a particular field. From there, the study sections’ recommendations were forwarded to an NIH council of scientists and laymen, which added its own recommendations. Final decision-making power rested in the hands of the institute directors, heads of NIH member institutions such as the National Cancer Institute and the National Eye Institute. While the directors took the earlier evaluations into account, they were not obligated to follow the recommendations of the study sections or the council. Furthermore, NIH applicants would receive little information about why their grants had been accepted or rejected. Deliberations about the grants were considered confidential and internal to the NIH.

The National Science Foundation (NSF), established by the National Science Foundation Act of 1950, was more reliant on outside experts than the NIH. From the outset, its proposal review process was designed to involve a range of experts from across the basic science fields that the NSF was going to support. Some proposals were sent out for “ad hoc” mail review, in which copies of the proposal were mailed to scientists who submitted their comments by return mail. Other proposals were evaluated by special panels of experts assembled in Washington. The choice of the reviewers, as well as the choice of panel review versus mail review, was left up to NSF employees. By 1975, according to internal NSF statistics, 44 percent of proposals were evaluated by ad hoc mail review, 28 percent received panel review, and 28 percent received reports from a mix of the two systems.²¹

Like the NIH, however, the NSF ultimately left final decisions about which grants to fund in the hands of its leaders. Referee opinions were seen as only one piece in the decision—an

¹⁹ Richard Mandel, *A Half Century of Peer Review, 1946–1996* (Bethesda, Md.: National Institutes of Health, 1996), pp. 4–5; and A. Landsborough Thomson, *Half a Century of Medical Research* (London: Her Majesty’s Stationery Office, 1973), pp. 147–161.

²⁰ The roots of the NIH go back to the nineteenth century, when the Marine Hospital Service created a laboratory to study bacteriological theory. In 1930, a government act renamed this laboratory the National Institute of Health and gave the new NIH the right to award fellowships and grants for public health research. The grants program was initially modest, owing to Depression-era economic circumstances. However, in 1948 the NIH’s budget and grants programs underwent dramatic expansion, and the NIH was renamed the National Institutes of Health in recognition of the range of biomedical projects it now funded.

²¹ *National Science Foundation Peer Review Special Oversight Hearings: Hearings before the Subcommittee on Science, Research, and Technology* (Washington, D.C.: U.S. Government Printing Office, 1975) (hereafter cited as *Special Oversight Hearings*), p. 264. On the NSF review process more generally see George T. Mazuzan, “Good Science Gets Funded. . .: The Historical Evolution of Grant Making at the National Science Foundation,” *Science Communication*, Sept. 1992, 14:63–90.

important piece, but not the determining factor as to whether the NSF would award or decline funding. As one internal NSF document explained the system, “It is intended that the program official, taking into account all pertinent factors and information input, will make a final decision.”²² Furthermore, scientists who submitted proposals would not be given copies of their reviews, only a summary report prepared by an NSF employee stating the major reasons for the acceptance or rejection of the proposal.

Both the NSF and the NIH, however, would come to place more emphasis on external referee opinions following a 1975 controversy that I will discuss in more detail later in this essay. Today, just as it would be difficult to find a respectable scientific journal that does not use peer review, it is nearly impossible to find a major funding body that does not employ an external peer review process to evaluate its grants. Many major grant organizations now cite their use of external referees—that is, independent experts who are not employed by the grant organization—to illustrate the care they take with proposals. The Medical Research Council website has a page devoted to the details of its “external peer review process.” The German Research Foundation assures applicants that it consults over nine thousand external referees per year. The NIH has multiple pages on its website explaining the steps in their review process and the ways in which they ensure the “integrity and confidentiality” of grant refereeing.²³

Interesting contrasts emerge when we consider grant refereeing and journal refereeing side by side. Funding bodies’ frequent practice of assembling a panel of experts, often meeting in person, stands in contrast to the system of sending a paper out for two or three opinions at journals. Furthermore, while journal referees were in theory only judging the paper at hand, grant applicants had to submit CVs and evidence of previous work. Grant referees explicitly considered the applicant’s background as a factor in their decision, making the referee report a judgment about both this particular proposal and the scientist’s body of work—and, often, his or her personal qualities—as a whole. There is also an epistemological difference. Journal referees are evaluating science that has already been performed and judging whether the paper’s conclusions ought to be given the stamp of authenticity that publication in a specialist journal bestows. Grant referees are, in theory, judging science that has yet to be done, and there is often an implicit sense in discussions about grant refereeing that even the best proposals pose something of a risk. Perhaps the most notable historical contrast between grant refereeing and journal refereeing is how comparatively rare grant refereeing was up to the mid-twentieth century. Most funding bodies kept their decisions among a small number of experts, in a system akin to the editorial board model so many journals had employed before refereeing became too much of a burden for the board members. Furthermore, journal refereeing procedures seem to have had little influence on grant refereeing procedures, and vice versa; the two systems, as John Burnham argued, seem to have evolved largely independently of one another.²⁴

But that independence does not mean historians should continue to evaluate these two manifestations of refereeing entirely separately from one another. It cannot be a coincidence

²² Louis Levin, Office of Program Development and Analysis, memo, “Proposed Uniform Quality Rating System,” 8 Dec. 1965, National Science Foundation Historian’s Papers, National Archives and Records Administration, College Park, Maryland (hereafter cited as *NSF Historian’s Papers*), Box 25.

²³ “External Peer Review,” Medical Research Council, 2018, <https://www.mrc.ac.uk/funding/peer-review/external-peer-review/> (accessed 12 Jan. 2018); “The DFG Review Process,” *Deutsche Forschungsgemeinschaft*, 17 July 2013, http://www.dfg.de/download/pdf/foerderung/antragstellung/begutachtung/dfg_begutachtungsverfahren_130715_en.pdf (accessed 12 Jan. 2018); “Peer Review Policies and Practices,” National Institutes of Health, 2018, <https://grants.nih.gov/policy/peer/index.htm> (accessed 12 Jan. 2018); “Frequently Asked Questions: Confidentiality in Peer Review,” National Institutes of Health, 15 Jan. 2015, https://grants.nih.gov/grants/peer/confidentiality_faqs.htm (accessed 12 Jan. 2018); and “Integrity and Confidentiality in NIH Peer Review,” National Institutes of Health, 2018, https://grants.nih.gov/grants/peer/confidentiality_peer_review.htm (accessed 12 Jan. 2018).

²⁴ Burnham, “Evolution of Editorial Peer Review” (cit. n. 5), p. 1323.

that the late twentieth century—specifically, the late 1960s and early 1970s—was the moment when both journals and funding bodies found themselves under increasing pressure to employ external referees. Nor does coincidence seem likely when we consider that it was American journals and funding bodies that were the most likely to employ refereeing and Americans who were seen as more dependent on refereeing than their counterparts in other countries. Something else was happening in the United States as well: increasingly, Americans were calling refereeing by a new term, “peer review.”

THE RISE OF “PEER REVIEW”

The term “peer review” has an intriguing linguistic history. It appears to have originated not from journals, and not in reference to external referees, but from review committees at grant organizations and in the medical community. In the 1960s and early 1970s, for example, major American newspapers usually used “peer review” to refer to reviews of medical practices for compliance with Medicare and Medicaid.²⁵

To track the introduction of the term “peer review” into the English-speaking scientific community, it is useful to look at *Science* and *Nature*, two prominent scientific weeklies that contained news and commentary. *Science* first used the term “peer review” in a 1965 article about the NIH. Joseph D. Cooper explained that as part of an evaluation of the NIH’s review process, “project grantees . . . were asked to make assessments of the output of a system of project approval in which they were intimately involved both as grantees and as members of peer review groups.” However, *Science* used the term only infrequently until the 1970s—and then mostly in reference to review procedures specific to medical journals and the NIH.²⁶ Across the Atlantic Ocean, the British *Nature* did not use the term at all until 1971 and for most of the 1970s used it only to refer to American grant or institutional review in the biomedical sciences. One *Nature* article in 1975 described “the so-called peer review system” specifically as a “method for judging the relative merits of competing grant proposals.”²⁷

In the medical literature, the *New England Journal of Medicine* first used the term “peer review” in 1969 to refer to the procedures for ensuring that hospitals and physicians were abiding by the regulations of the new Medicaid funding programs: “Although state health authorities have been concerned with standard formulation and facility licensing and regulation for many years, few state programs can boast of the establishment of effective monitoring or ‘peer review’ methods to assure the adequacy of institutional and professional services under Medicaid.”²⁸

²⁵ See, e.g., Thomas O’Toole, “Nation Is Warned of Health Crisis,” *Washington Post-Times Herald*, 21 Nov. 1967, p. A1; Spencer Rich, “Social Security Bill Approved in Senate: Senate Drops Family Welfare Reforms in Bill,” *ibid.*, 6 Oct. 1972, p. A1; Eric Morgenthau, “Medical Checkup: Local Panels to Review Doctors’ Work and Bills Are Urged by Critics; Public, Private Health Plans Being Abused by Physicians, Backers of Panel Assert; Case of the Costly Earache,” *Wall Street Journal*, 16 Dec. 1970, pp. 1, 20; Jerry E. Bishop, “The Wrong Way to Fight Cancer,” *ibid.*, 16 July 1971, p. 6; “AMA Delegates Deride Aetna Life for Letter Impugning Doctor Fees; Firm’s Bid to Pay Cost of Legal Challenges to Bills Criticized; Consultation on Charges Urged,” *ibid.*, 23 June 1972, p. 6; “Medical Peer Review,” *New York Times*, 17 Dec. 1972, p. E10; and Nancy Hicks, “A.M.A. Will Back Peer Review Plan, But Seeks Changes,” *ibid.*, 27 July 1974, p. 1.

²⁶ Joseph Cooper, “Onward the Management of Science: The Wooldridge Report,” *Science*, 11 June 1965, 148:1433–1439, on p. 1434. For other appearances of the term in the journal see Philip M. Boffey, “Health Crisis: LBJ Panel Calls for Reshaping American Medicine,” *ibid.*, 1967, 158:1160–1162; Marti Mueller, “Oral Contraceptives: Government-Supported Programs Are Questioned,” *ibid.*, 1969, 163:553–555; and Barbara J. Culliton, “Dual Publication: ‘Ingelfinger Rule’ Debated by Scientists and Press,” *ibid.*, 1972, 176:1403–1405. See also “Medical Ethics Examined,” *Nature*, 1973, 242:7–8.

²⁷ Colin Norman, “Peer-Review under Review,” *Nature*, 1975, 258:382. See also “Academy Critic Shoots Wildly But Hits Sore Spot,” *ibid.*, 1971, 229:151–152; and “Has Science Never Had It So Good?” *ibid.*, 1971, 233:169–172.

²⁸ Samuel Levy, “A Perspective on Medicaid,” *New Engl. J. Med.*, 1969, 281:297–301.

Similarly, the term was first used in the *Journal of the American Medical Association (JAMA)* in 1970 to describe physicians’ assessments of each others’ practices under the Professional Standards Review Organization (PSRO) associated with Medicaid; this continued to be JAMA’s most frequent use of the phrase through the 1970s. “Peer review means different things to different people,” the physician Irvine H. Page explained in a 1973 editorial for JAMA. “To most American physicians it means PSRO, to the British House of Lords it means Peers examining other Peers for moral turpitude, and to the scientific community, it means Study Sections and Councils that determine a grantee’s financial and possibly research future.” Notably, journal refereeing was *not* one of the definitions Page offered, and JAMA would not use the term “peer review” to refer to journal refereeing until 1981.²⁹

Franz J. Ingelfinger, the influential editor of the *New England Journal of Medicine*, was one of the first people to refer to journal refereeing as “peer review” in print. In 1968, Ingelfinger was invited to write an opinion piece for the *Lancet* on the purpose of the medical journal. He wrote that “to be accepted the article must pass peer review, and ideally either must pass with high honours or must undergo appropriate revision.” The term would not appear again in the *Lancet* until the 1970s. Ingelfinger was also the first to use the term “peer review” for journal refereeing in *Science*, which he did in a 1970 article on the current state of medical literature. In a 1970 article on the *New England Journal of Medicine*’s review process, a member of the editorial staff (likely Ingelfinger) wrote that the journal tried to maintain quality and integrity through “a ‘peer review’ of submitted articles whose findings are weighed in final judgment by editors and their advisers.” The biologist Thomas Jukes (an Englishman at the University of California, Berkeley, and a regular *Nature* columnist) was the first *Nature* author to call journal refereeing “peer review,” in 1977.³⁰

The shift from “refereeing” to “peer review” is not merely a linguistic curiosity. Calling external refereeing at journals or funding bodies “peer review” established that the evaluation of a paper or grant proposal could only be done by experts—the peers of the person who submitted the work. The new term established a narrow range of acceptable reviewers and implicitly deemed those without a scientific background unqualified to evaluate the work in question. Moreover, it seems clear that this shift in terminology originated in the United States. One *Nature* contributor, Lord Zuckerman, indicated that many considered the phrase an Americanism, referring to the process of “what the Americans call ‘peer review’” in a 1972 letter to the Correspondence column.³¹

FUNDING BODIES IN THE POSTWAR UNITED STATES

The late twentieth-century United States seems to be a crucial context for understanding when and how “peer review” came to be seen as central to science. Changes to the way science was funded in the United States provide one potential clue as to why. Between 1948 and 1953,

²⁹ Irvine H. Page, “A Different Peer Review,” *J. Amer. Med. Assoc.*, 1973, 225:1240–1241; and Lois DeBakey and Selma DeBakey, “Journal Editors and the Press: Cooperation Not Conflict,” *ibid.*, 1981, 245:2296–2297 (first use to refer to journal refereeing). For uses of the term in the journal during the 1970s, largely in the context of PSRO assessments, see, e.g., Robert Katz, “Peer Review Involving Nonmembers of Medical Societies,” *ibid.*, 1970, 213:1686; A. Michael Pardue, “Peer Review,” *ibid.*, 1973, 225:67; and Calvin Kunin and Herman Efron, “Guidelines for Peer Review: Veterans Administration Ad Hoc Interdisciplinary Advisory Committee on Antimicrobial Drug Usage,” *ibid.*, 1977, 237:1001–1002.

³⁰ Franz J. Ingelfinger, “To Impart the Precepts and the Instruction,” *Lancet*, 5 Oct. 1968, 292(7571):766–767; Ingelfinger, “Medical Literature: The Campus without Tumult,” *Science*, 1970, pp. 831–837, on p. 836; “There Shall Be Weeping and Gnashing of Teeth,” *New Engl. J. Med.*, 1970, 282:393–394; and Thomas Jukes, “Peer Review,” *Nature*, 1977, 265:203.

³¹ Lord Zuckerman, “Letter: Cancer Research,” *Nature*, 1972, 240:497–498.

federal spending on scientific research in the United States increased by a factor of twenty-five, adjusted for inflation.³² In the 1950s, as Cold War tensions rose, few Americans questioned funding for scientific research; to many legislators and taxpayers, the idea that the Soviets might be ahead of the Americans in science and technology seemed sufficient justification for any spending. However, the most intense post-*Sputnik* anxiety soon abated, and as early as the mid-1960s legislators and analysts were questioning whether the government funding for science was yielding desirable results. In 1966, the Department of Defense (DOD) issued the Project Hindsight report, the result of an investigation into the outcomes of DOD funding for scientific research. The Hindsight report concluded that while applied research had yielded massive benefits in the form of new military technologies, the DOD's investment in basic research had not produced similar advances. The report led many policy makers to question military and government expenditures on basic research.³³

There were also questions about whether organizations like the NSF and the NIH were spending their funding wisely. In the early 1960s, Democratic Congressman Lawrence H. Fountain of North Carolina raised concerns that NIH grants were being mismanaged. Fountain had become alarmed after examining some of the NIH's appropriations requests and learning that scientists could adjust their grant budgets without NIH approval. He pushed for a closer examination of the NIH's finances, claiming that significant funds were being redirected to non-research expenses.

Fountain's investigation brought him into conflict with NIH director James Shannon, a confident, polished physiologist who first made his name as a researcher at the New York University School of Medicine. Initially Shannon seemed to agree that some grants might benefit from closer supervision and agreed to consider new supervisory rules. However, the NIH moved slowly to implement any changes. In March 1962, a frustrated Fountain assembled a subcommittee to evaluate the NIH's progress and called Shannon to testify.³⁴

At the hearing, Shannon came across as dismissive, even arrogant. He assured the committee that grant recipients "are selected on the basis of a rigorous screening by their scientific peers" and that "all subsequent administrative actions having to do with adjustments of budgets, and so forth, are essentially trivial in relation to this basic selection process."³⁵ He insisted that scientists themselves were the best judges of how grant money should be spent and that the inner workings of the scientific community naturally safeguarded against fraud or excessive requests. In Shannon's view, the NIH's internal grant review process chose the best scientific applicants, and that was sufficient to guarantee that funds were spent in a responsible manner.

However, an investigation by Fountain's committee had revealed that at least one NIH grantee had, indeed, mismanaged funds. The NIH had made several grants to a private company called Public Service Research, Incorporated. In their required accounting documents, Public Service Research had included several expenses that were not strictly research related,

³² David Kaiser, "Shut Up and Calculate!" *Nature*, 2014, 505:153–155.

³³ On Project Hindsight see Glen Ross Asner, "The Cold War and American Industrial Research" (Ph.D. diss., Carnegie Mellon Univ., 2006); Ronald Kline, "Construing 'Technology' as 'Applied Science': Public Rhetoric of Scientists and Engineers in the United States, 1880–1945," *Isis*, 1995, 86:194–221; Karl Kreilkamp, "Hindsight and the Real World of Science Policy," *Science Studies*, 1971, 1:43–66; and Audra J. Wolfe, "The Military-Academic-Industrial Complex and the Path Not Taken," in *Origins of the National Security State and the Legacy of Harry S. Truman*, ed. Mary Ann Heiss and Michale J. Hogan (Kirkville, Mo.: Truman State Univ. Press, 2015), pp. 143–164.

³⁴ For more on these controversies and their fallout for the NIH see Laura Stark, *Behind Closed Doors: IRBs and the Making of Ethical Research* (Chicago: Univ. Chicago Press, 2011), pp. 144–158.

³⁵ *The Administration of Grants by the National Institutes of Health: Hearings before a Subcommittee of the Committee on Government Operations* (Washington, D.C.: U.S. Government Printing Office, 1962), p. 14. See also Daniel Greenberg, *The Politics of Pure Science* (1967; Chicago: Univ. Chicago Press, 1999), pp. 278–279.

such as money to recruit personnel and funds for moving and renovating the company's offices. NIH policy did not permit those kinds of expenses to be charged to research grants, but the breach had apparently gone unnoticed and had not prevented Public Service Research from being awarded further grants. While no one accused the company of intentional fraud, even Director Shannon was forced to admit that Public Service Research had misspent the awards and that the NIH should have noticed the problems before Fountain's audit brought them to light.³⁶

The committee questioned Shannon extensively about the NIH's refereeing procedures, asking him to walk them through exactly how the NIH chose which proposals to fund. At the end of that questioning, they seem to have concluded that the NIH's method of choosing grants via internal review was unremarkable and unproblematic. Instead, Fountain and his supporters blamed the NIH's accounting system for the problems. Fountain concluded that NIH grant spending should be much more closely supervised and that grantees should have to obtain approval before changing their budgets, and he rallied his fellow fiscal conservatives to pass legislation requiring NIH grantees to account for their spending more precisely. The result was new and much stricter accounting practices for NIH grant recipients and significant reorganization at the NIH. Postaward accounting was no longer the responsibility of the Division of Research Grants; instead, once grants were awarded, all spending would be overseen by a new and separate grants management division.³⁷

It is significant that although many at the hearings agreed that Public Service Research probably should not have received multiple grants, no one suggested that the NIH ought to rethink its grant reviewing method. They only argued that grant *accounting* should be more strictly monitored. Internal review by the NIH was still considered an acceptable method of choosing which grants to fund, and the word "referee" appears nowhere in the transcript of the hearings. Insufficient reliance on external referees was not yet a point of criticism. That, however, would soon change.

CONTROVERSIES AT THE NSF IN THE 1970s

The 1970s proved to be an even more tenuous time for American governmental science funding. The Cold War entered its period of détente, and tensions between the United States and the USSR temporarily abated. This blunted one of the most persuasive arguments in favor of increased science funding. Furthermore, as the Vietnam War became more and more controversial, ties between scientists and the military came under scrutiny. Antiwar activists on many college campuses protested departments and research institutes that held large military contracts and in some cases persuaded universities to cut formal ties with research institutes that specialized in military work. An increasing number of scientists also expressed concerns over the power of the "military-academic-industrial" complex; a group called "Science for the People" disrupted several meetings of the American Association for the Advancement of Science (AAAS) meetings in the early 1970s, shouting down speakers with military ties (such as Glenn Seaborg, president-elect of the AAAS and the discoverer of plutonium) and decrying the political overtones of certain new scientific theories, most notably sociobiology.³⁸

³⁶ *Administration of Grants by the National Institutes of Health*, pp. 51–87.

³⁷ Mandel, *Half Century of Peer Review, 1946–1996* (cit. n. 19), pp. 96–99.

³⁸ On Vietnam-era campus activism and science see Ellen Herman, "Project Camelot and the Career of Cold War Psychology," in *Universities and Empire: Money and Politics in the Social Sciences during the Cold War*, ed. Christopher Simpson (New York: New Press, 1998), pp. 97–133; Stuart Leslie, *The Cold War and American Science: The Military-Industrial-Academic Complex at MIT and Stanford* (New York: Columbia Univ. Press, 1993); and Matt Wisnioski, "Inside 'the System': Engineers, Scientists, and the Boundaries of Social Protest in the Long 1960s," *History and Technology*, 2003, 19:313–333. On Science for the People see

Meanwhile, the U.S. economy was facing major challenges. The combination of an oil crisis with a period of stagflation led to widespread unemployment, a decrease in real wages, and budget restrictions at a wide range of organizations. *Science*, for example, adopted strict page limits in response to declining AAAS membership dues and increasing paper prices.³⁹ The federal government was under particularly intense pressure to cut its budget in the face of a decreased tax base and widespread economic anxiety. All of this combined to create an environment where legislators on both sides of the aisle had reason to question government spending on scientific research. By 1975, two Republican Congressmen and a Democratic Senator were vocally criticizing the NSF.

Congressman John Conlan (a Republican from Arizona), Congressman Robert Bauman (a Republican from Maryland), and Senator William Proxmire (a Democrat from Wisconsin) each had distinct reasons for questioning the NSF's budget and grant-making procedures. Conlan was a first-wave member of the Christian Right who was gearing up to run for the Senate, billing himself as a social conservative who would bring Christian values to his government service.⁴⁰ Two NSF-funded school programs had long been controversial among Arizona's social conservatives: "Man, a Course of Study" (MACOS) and the "Individualized Science Instructional System" (ISIS). MACOS was a social sciences curriculum that covered the social habits of several different species of animals, as well as the Netsilik Eskimos; critics accused it of promoting moral relativism and focusing on violence, incest, and cannibalism among animals.⁴¹ ISIS, a program aimed at fourth graders, came under fire for allegedly including too much explicit discussion of human sexuality.

During the hearings over the NSF's 1976 appropriations request, Conlan stridently criticized both programs and questioned whether the NSF was spending its budget in the interests of the American public. Conlan soon found an ally in Bauman, a rising star in the Republican Party.⁴² Bauman, a small-government advocate and a founder of the American Conservative Union, began pushing for more fiscal accountability from the NSF. Eventually he proposed an amendment to the 1976 NSF appropriations request requiring the NSF to receive Congressional

Donna Haraway, "The Transformation of the Left in Science: Radical Associations in Britain in the 1930s and the USA in the 1960s," *Soundings*, 1975, 58:441–462; Kelly Moore, "Organizing Integrity: American Science and the Creation of Public Interest Organizations, 1955–1975," *American Journal of Sociology*, 1996, 101:1592–1627; Moore, *Disrupting Science: Social Movements, American Scientists, and the Politics of the Military, 1945–1975* (Princeton, N.J.: Princeton Univ. Press, 2008); and Sigrid Schmalzer, Daniel S. Chard, and Alyssa Botelho, eds., *Science for the People: Documents from America's Movement of Radical Scientists* (Amherst: Univ. Massachusetts Press, 2018). See also *Science's* entertaining (though perhaps not entirely neutral) account of the 1970 Chicago meeting: Philip M. Boffey, "AAAS Convention: Radicals Harass the Establishment," *Science*, 8 Jan. 1971, 171:47–49.

³⁹ See, e.g., Editors to Dr. James Bonner, 9 Nov. 1971, Abelson Papers, Box 90, Folder: Mayfield, S. E., Chromosomal RNA.

⁴⁰ Conlan would go on to lose one of the most bitterly contested primary races in U.S. history. His opponent in the 1976 Arizona Republican Senatorial primary, Sam Steiger, was a Jewish Congressman. Conlan portrayed himself as a leader who would be the first in a wave of Christians elected to public office, and his campaign was accused of both subtle and overt anti-Semitism. Steiger won the primary, but Arizona Republicans were deeply divided by the viciousness of the campaign. Ultimately, despite Arizona's heavily Republican voter base, Steiger lost the Senate seat to inexperienced Democratic candidate Dennis DeConcini. On Conlan see John Turner, "Selling Jesus to Modern America: Campus Crusade for Christ, Evangelical Culture, and Conservative Politics" (Ph.D. diss., Univ. Notre Dame, 2006).

⁴¹ On MACOS see Peter Dow, *Schoolhouse Politics: Lessons from the Sputnik Era* (Cambridge, Mass.: Harvard Univ. Press, 1991); R. W. Evans, *The Tragedy of American School Reform: How Curriculum Politics and Entrenched Dilemmas Have Diverted Us from Democracy* (New York: Palgrave Macmillan, 2011), Ch. 5; Mazuzan, "'Good Science Gets Funded . . .'" (cit. n. 21); and Erika Milam, "Public Science of the Savage Mind: Contesting Cultural Anthropology in the Cold War Classroom," *Journal of the History of the Behavioral Sciences*, 2013, 49:306–330.

⁴² On Bauman see his memoir: Robert Bauman, *The Gentleman from Maryland: The Conscience of a Gay Conservative* (New York: Arbor House, 1986). Bauman's promising political career came to an abrupt end when he was arrested for soliciting a teenage male prostitute.

approval for any grants it wanted to fund. The amendment was approved and added to the appropriations bill.

It was Proxmire, however, whose NSF criticisms made the biggest public splash. Proxmire joined the Senate in 1957 after winning a special election to replace the recently deceased Joseph McCarthy. He would go on to a long and colorful Senate career marked by principled, stubborn stances on many controversial issues. He was an early opponent of the Vietnam War, an early advocate of campaign finance reform, and a fierce critic of anything that he perceived as wasteful spending.

In March 1975 Proxmire began issuing his Golden Fleece Award, a badge of dishonor that he bestowed each month on the government project that he deemed the worst use of taxpayer money. Proxmire would “award” his Golden Fleece every month from March 1975 until his Senate retirement in 1988. The first two Golden Fleece Awards went to NSF projects. In March, Proxmire singled out a sociological study at the University of Wisconsin about interpersonal attraction. “I believe that 200 million other Americans want to leave some things in life a mystery, and right on top of the things we don’t want to know is why a man falls in love with a woman and vice versa,” Proxmire proclaimed. “So National Science Foundation—get out of the love racket.” The April award went to the psychologist Ronald Hutchinson’s study of why humans, rats, and monkeys clench their jaws in moments of stress. “The funding of this nonsense makes me almost angry enough to scream and kick or even clench my jaw,” said Proxmire. “The good doctor has made a fortune from his monkeys and in the process made a monkey out of the American taxpayer.”⁴³

Proxmire’s Golden Fleece won widespread media coverage, and letters complaining about those studies began flooding the NSF offices.⁴⁴ The researchers involved were also singled out for criticism and angry mail—even threats. Proxmire, for his part, would make the NSF a favorite target of the Golden Fleece Awards, insisting that the organization funded frivolous and wasteful projects that were of no use to the American people. The school curriculum controversy—in particular, the criticisms of MACOS—also received a great deal of press coverage.⁴⁵

But the controversy was not only being carried out in the media. Dozens of letters were exchanged between the NSF, Proxmire, Conlan, and Bauman as the three legislators sought further information on particular NSF grants. By March 1975, the controversy had grown so intense that NSF Director H. Guyford Stever announced that no further funds would be approved “either for MACOS or any other precollegiate science course development . . . until we have conducted a thorough review of the NSF effort in these areas and reported to the National Science Board and Congress with recommendations.” That decision did not mollify Conlan, who was determined to learn how the NSF had decided to fund ISIS and was frustrated when the foundation refused to give him copies of the grant’s peer review reports: “I would again remind

⁴³ William Proxmire, quoted in Elaine Hatfield and G. W. Walster, *A New Look at Love* (Lanham, Md.: Univ. Press America, 1978) (Hatfield was one of the researchers whose work was being criticized in Proxmire’s press release); and Proxmire, quoted in Constance Holden, “‘Fleece’ Winner Sues Proxmire,” *Science*, 21 May 1971, 192:769.

⁴⁴ For NSF correspondence with members of the public (often via their representatives) on controversial grants see, e.g., Ralph Kazarian to Mrs. Joseph Long, 12 May 1975, NSF Historian’s Papers 6-27-208; Elizabeth Hunt to Clifford Case, 27 May 1975, NSF Historian’s Papers 6-27-237; Patricia Nicely to Robert Mills, 30 May 1975, NSF Historian’s Papers 6-27-239; Nicely to Sam Nunn, 30 May 1975, NSF Historian’s Papers 6-27-244; and Jack Kratchman to Joe Evins, 12 June 1975, NSF Historian’s Papers 6-27-266.

⁴⁵ The syndicated columnist James Kilpatrick was a particularly visible media critic. See, e.g., James Kilpatrick, “Science Foundation Needs More Scrutiny,” *Sun*, 27 Mar. 1975, p. A19; Kilpatrick, “Fifth-Grade and Sex,” *Atlanta Constitution*, 1 Apr. 1975, p. 4A; and Kilpatrick, “NSF Bureaucracy Hard to Kill,” *ibid.*, 20 Jan. 1976, p. 4A.

you that I am a Member of Congress on a Committee charged with the oversight of the National Science Foundation. . . . Consequently, I do again demand that you make available the peer reviewer comments originally demanded by me—in their original and complete form, not paraphrased.”⁴⁶

NSF leaders, however, argued that peer review depended on the anonymity of the referees and said that the NSF was not obligated to disclose the reports to anyone, including members of Congress. In a letter to Conlan, Stever asserted that referee reports were filed under an “implied promise of confidentiality” that was “well understood and accepted by thousands of reviewers” and that releasing the text of the reports or the names of the reviewers would constitute a breach of trust. In a later letter, he argued that any change to the NSF’s policy of keeping referees anonymous would be at odds with the concept of peer review itself.

Maintaining the confidentiality of comments solicited by the Foundation on grant proposals and the identity of the reviewer making these comments is a fundamental principle on which the Peer Review System of the National Science Foundation and also of other agencies is based. . . . If this confidentiality principle were to be changed, the entire process whereby the wisdom and knowledge of outstanding scientists and educators throughout the country is drawn upon in evaluating the worth of proposals would thereby be changed.⁴⁷

Confidentiality would continue to be a bone of contention between the NSF and its critics as the conflict progressed.

The careful phrase “implied promise of confidentiality” suggests that instructions given to NSF reviewers did not always explicitly promise that the reports would be kept confidential; however, Stever’s response to Conlan was consistent with internal NSF policy. NSF correspondence with both grant applicants and Congressmen routinely stated that it was the NSF’s practice not to disclose the full text of the referee reports or the names of the reviewers. “[Referees] are private citizens who serve voluntarily without compensation, often devoting several hours of their own time to review the proposal,” NSF General Counsel Charles F. Brown told one disgruntled applicant. “If the name of a reviewer were made public, the reviewer might be subject to harassment or badgering by the principal investigator.”⁴⁸

THE NSF HEARINGS: SCIENTIFIC AUTONOMY OR PUBLIC ACCOUNTABILITY?

The public debate and behind-the-scenes arguments over the NSF’s funding decisions led to the National Science Foundation Peer Review Special Oversight Hearings, held before the Subcommittee on Science, Research, and Technology in July 1975. Bauman, Conlan, and Proxmire were not members of the committee but were invited to testify at the hearings, as was NSF Director H. Guyford Stever and Deputy Director Richard Atkinson. Over a dozen scientists, sociologists, and other academics were also asked to testify about their opinions on peer review. The hearings generated nearly twelve hundred pages of transcripts and supplemental documents. Those transcripts contain not only discussion of the NSF’s particular review process but much discussion of the concept of peer review itself—its outcomes, its ideal form, and its role in scientific practice.

⁴⁶ H. Guyford Stever to Olin Teague, 17 Mar. 1975, NSF Historian’s Papers 6-27-91; and John Conlan to Stever, 15 May 1975, in *Special Oversight Hearings*, p. 21.

⁴⁷ Stever to Conlan, 12 May 1975, NSF Historian’s Papers, 6-27-205; 16 May 1975, NSF Historian’s Papers, 6-27-220.

⁴⁸ Charles F. Brown to Norman Tallent, 12 July 1974, NSF Historian’s Papers, 6-26-239.

The discussion of the NSF's review process focused on two key points: first, what criteria should be used to choose which NSF proposals to fund, and, second, how much weight should be given to the referees' opinions about a proposal. During questioning, both Bauman and Conlan acknowledged that their interest in the NSF stemmed from concerns that some of the funded proposals were not beneficial to the American public. Bauman argued that "we are dealing here with a finite quantity of money taken in the form of taxes from people against their will in a very difficult time in our economy and being spent by a Federal agency, spent in a manner now questioned," and that the NSF "must demonstrate that it is using tax money in a prudent manner, in such a way that the taxpayers can expect that there will be some payoff from NSF-funded research."⁴⁹

Conlan's complaints were more focused; much of his testimony centered on his objections to MACOS and ISIS and his frustration at the NSF's refusal to provide him with the full, verbatim referee reports. Conlan described the NSF's peer review system as a secretive process that allowed program directors to elicit any decision they wanted: "Here is an amazing system, gentlemen, where individual program managers are given carte blanche authority to select peer reviewers. . . . [It is] a completely arbitrary system that is closed and unaccountable to the scientific community and to the Congress. . . . It is common knowledge in the science community that NSF program managers can get whatever answer they want out of the peer review system."⁵⁰

Conlan's criticisms of the NSF, however, did not lead him to criticize the concept of peer review in general. Instead, he argued that the NSF's review system was designed to allow the directors, rather than the reviewers, to make decisions about the quality of a proposal. Because no one outside the NSF was allowed to see the reports, said Conlan, neither scientists nor Congressmen could verify that NSF staff members were listening to the referees. Conlan claimed that NSF program managers occasionally went so far as to misrepresent their referees' comments. He recounted a conversation with one of the ISIS peer reviewers, Philip Morrison of MIT, and reported that Morrison had been very critical of the proposal and was surprised to learn that it received funding.⁵¹

Conlan thus proposed that he and his fellow Congressional critics were acting in defense of proper peer review. Referees, not directors, should be trusted to determine which proposals were most deserving of funding—and if all NSF reports were made available, NSF program officers would not be able to manipulate or discount their contents so easily. He therefore argued that all peer review reports, for funded and unfunded grants, should be available to Congress and to grant applicants. In Conlan's view, this would stop the NSF from making any further questionable grants: "if peer review and NSF grants management are open, I think the area of faulty grantmanship or inadvisable or incompetent awards will take care of themselves." Conlan also suggested that the reviewers' names ought to be available to applicants, arguing that this would encourage more objectivity in the refereeing process: reviewers would work harder to be fair if their names would be attached to their reports.⁵²

Unsurprisingly, the NSF leadership defended the foundation against the charges of secrecy and cronyism. Stever vigorously defended the role of program officers, arguing that their expertise and professionalism were crucial to the NSF's work. However, he said, "it is the Foundation's internal

⁴⁹ *Special Oversight Hearings*, pp. 182–183.

⁵⁰ *Ibid.*, p. 9.

⁵¹ *Ibid.*, pp. 6–8.

⁵² *Ibid.*, pp. 36 (quotation), 35.

practice to make sure that no member of its professional staff, no matter how competent, is permitted to make a decision in private on behalf of the Foundation.” Stever attributed the Golden Fleece and curriculum controversies not to flaws in the NSF’s review process but to a gap between public expectations and scientific realities. “Society is expecting more of science,” he said. “Some of these expectations are reasonable, but some are unreasonable expectations of magical results on a faster timetable than ever will be realized.” Because the benefits of scientific research would not always be immediate or obvious, said Stever, the foundation believed that proposals should be judged on the quality of the science. “The foundation never supports frivolous research,” he insisted. “Scientific excellence is the primary criterion for NSF support.”⁵³

The only way to choose proposals based on scientific excellence, Stever and Atkinson argued, was through a peer review process—by placing proposals in the hands of those who were best qualified to know what scientific excellence looked like. The foundation consulted with “members of the scientific community who are expert in the specialty or discipline involved before making a recommendation to support or to decline a specific proposal,” Stever explained. “The persons consulted are expected to be the scientific peers of those proposing to do the research; thus the growth of the term ‘peer review.’”⁵⁴ Stever repeatedly insisted that these peers could give candid advice only if their identities were kept confidential; otherwise, reviewers might fear professional reprisal if they did not recommend particular proposals, especially those by senior colleagues in their field.

The NSF’s critics and defenders disagreed about whether the NSF itself had done a good job with its refereeing procedures, but in many ways the claims they made about peer review itself were remarkably similar. Bauman, Conlan, and Stever all argued that peer review was critical and that referee opinions should be weighed heavily when determining which NSF proposals should receive funding—indeed, Conlan’s major criticism was that referee opinions were not given *enough* weight at the NSF. However, Conlan and Bauman both argued that the NSF’s peer reviewers could perform their most important work if their identities were public knowledge. Stever vehemently disagreed, arguing that if reviewers had to sign their names to their reports they might face pressure to give favorable reviews to particular proposals. Referees could be free to speak their minds only if they knew that they would remain anonymous. His comments cast the referee as a person who was performing a selfless service to science with no hope of reward, but also as a potentially malleable figure who might succumb to external pressure if his or her identity became known. This disagreement over anonymity reveals a major difference between the NSF’s critics and its supporters. NSF leaders supported anonymous peer review because, they argued, it was the only way to ensure candid and accurate feedback about proposals. Legislative critics, on the other hand, felt that the NSF’s decisions should be more transparent to legislators, taxpayers, and the scientists who applied for grants. In other words, the NSF’s priority was to ensure scientific excellence; the critics’ priority was to make the process accountable to the public.

WE ASK TOO MUCH OF IT: THE LIMITS OF PEER REVIEW?

Not everyone at the 1975 hearings, however, was convinced that peer review ensured either scientific excellence or public accountability. William D. Carey, the Executive Officer of the AAAS, argued that Congress and the NSF might be asking too much of peer review. Peer review, he told the committee, “is not a fail-safe procedure, nor one that should silence disagreements on

⁵³ *Ibid.*, pp. 141, 215.

⁵⁴ *Ibid.*, p. 142.

the theory that there is something sacramental about it." Carey endorsed the older system of using NSF peer review reports as advisory documents, saying, "Peer review has its uses as a first round of proposal screening, but it does not absolve the Government program manager from full responsibility for the decision to fund or reject a proposal. . . . There are some things that we should not ask of peer review. We should not ask it to take Government agencies off the hook on the question of protecting the public purse."⁵⁵

It was Proxmire, however, who expressed the most negative ideas about peer review. In his submitted testimony, he wrote, "I have received a number of letters pointing out that the peer review system serves to perpetuate the funding of established people, ideas and institutions." Proxmire seemed to hold out little hope that peer review could be anything other than "incessant." He argued that NSF reviewers would favor proposals whose research had been published in peer-reviewed journals, that peer-reviewed journals would look favorably on articles that had received NSF funding, and that "we come full circle, when we realize that many of the top researchers in a scientific field have been recipients of NSF grants, reviewers of NSF grants and finally editors of their technical journals."⁵⁶

But interestingly, any limitations of peer review received little discussion during the testimony of the NSF officials. Almost all of the actors at the 1975 hearings agreed that peer review was the appropriate mechanism for evaluating grants. Peer review, according to those at the hearings, was "a fair and effective system," "an indispensable component of the decisionmaking process for allocating funds," and "an integral feature of 'the scientific method.'" Subcommittee chair James Symington seemed to sum up the hearings' attitude toward peer review when he said: "Witnesses overwhelmingly agree that some form of peer review should continue to be used to assist in the allocation of funds for scientific research. To be sure, no witness—including the NSF—has claimed that the Foundation's peer review system and its decisionmaking ability are perfect. . . . Nevertheless, it appears that, as a base to work from, the peer review concept is seen as fundamentally sound."⁵⁷ In the 1960s, the term "peer review" was only beginning to be introduced; by 1975, peer review was seen as a central part of scientific knowledge making in the United States.

The belief that peer review was "fundamentally sound" strongly shaped the NSF's reactions to the criticisms. Most significantly, Stever announced at the hearings that as of 1 January 1976 applicants would receive complete, verbatim copies of their referee reports with the reviewers' names redacted. He argued that keeping the referees anonymous was the only way to ensure their candor and incorruptibility but offered the verbatim reports as a compromise.

Stever and Atkinson also indicated that in the future the NSF would lean more heavily on peer review, and less heavily on the judgment of NSF staff, when deciding which grants to fund. A new audit office would ensure that funding decisions placed appropriate weight on positive and negative referee reports. Finally, the NSF commissioned a report on its peer review process from the RAND Corporation to learn more about scientists' views as to its efficacy and fairness.⁵⁸ Meanwhile, although the NIH was not directly involved in the 1975 controversy, NIH leadership—including the new director, Donald H. Fredrickson—watched the debate closely. In response to the uproar, the NIH took some precautionary steps to reform its own

⁵⁵ *Ibid.*, p. 780.

⁵⁶ *Ibid.*, p. 987.

⁵⁷ *Ibid.*, pp. 1092, 460, 390, 887.

⁵⁸ The report was a fairly straightforward survey of NSF grant applicants and reflected a general approval of the NSF's procedures. See Deborah Hensler, *Perceptions of the National Science Foundation Peer Review Process: A Report on a Survey of NSF Reviewers and Applicants*, prepared for the Committee on Peer Review, National Science Board, and the Committee on Science and Technology, U.S. House of Representatives, Dec. 1976.

peer review system. Between 1975 and 1978, the NIH gradually gave external reviewers a more central role in deciding which grant proposals should be funded. Like the NSF, the NIH also began providing more detailed comments about proposals to NIH applicants.⁵⁹

Notably, both the NSF's and the NIH's reforms placed *more* emphasis on the opinions of external referees. The contrast with the NIH hearings in the 1960s is striking. In 1962, even with NIH grants under fire, few questioned the NIH's internal review system. In 1975, however, NSF employees' power to determine which grants should be funded was a major point of criticism. Funding body employees were individuals with biases; peer review, however, could be relied on to produce answers about what was and was not good science. The judgment of the anonymous referees could serve as a stand-in for the judgment of the scientific community as a whole. Individuals could not be trusted, but the system of peer review could. Peer review was thus elevated from an optional bureaucratic process to a system that was supposed to ensure the quality and trustworthiness of science.

CONCLUSION

Few walked away from the 1975 NSF hearings completely satisfied with the outcome. The NSF's education programs were significantly downsized, and funding for MACOS and ISIS was largely eliminated. However, Conlan and Bauman did not get the Congressional oversight of NSF grants they had advocated, and the Bauman amendment was removed from the NSF's appropriations request. Furthermore, the NSF reforms were not quite as sweeping as they might have seemed. NSF program officers retained a great deal of influence over funding decisions—influence that grew when the NSF changed the name of its review process from “peer review” to “merit review” in the 1980s, emphasizing that concerns besides scientific excellence (such as national interests) would be part of the decision-making process.⁶⁰

The 1975 peer review controversies are most noteworthy not because of their effect on the actual *practice* of refereeing at funding bodies but because of their implications for public conceptions of peer review. The hearings provided a moment for stakeholders with different views about science and its funding to debate what the practice of peer review was for. What emerged at the hearings was a general (though not unanimous) consensus that the NSF—and any other organization—had to rely on external referees in order to judge “good science” properly. At the hearings, peer review was cast as a process that was crucial to the way science worked, one that had to be preserved and defended in order for science to work properly in the future.⁶¹ That vision of peer review has endured and expanded, resulting in headlines like the one dubbing the Higgs boson “actual science.”

However, Carey's and Proxmire's concerns about peer review have also endured. In fact, peer review seems to be in a moment of crisis. In recent years, high-profile papers have passed peer review only to be heavily criticized after publication or retracted amid allegations of fraud.⁶² Some studies of peer review outcomes have suggested that women and underrepresented minor-

⁵⁹ Mandel, *Half Century of Peer Review, 1946–1996* (cit. n. 19), pp. 149–151.

⁶⁰ Marc Rothenberg, “Making Judgments about Grant Proposals: A Brief History of the Merit Review Criteria at the National Science Foundation,” *Technol. Innov.*, 2010, 12:189–195.

⁶¹ There are interesting resonances between this episode and other points in the history of science when a moment of crisis or debate pushed historical actors to make previously unspoken processes explicit, thus cementing their importance. See, e.g., William Deringer, *Calculated Values: Finance, Politics, and the Quantitative Age* (Cambridge, Mass.: Harvard Univ. Press, 2018); and Evan Hepler-Smith, “A Way of Thinking Backwards: Computing and Method in Synthetic Organic Chemistry,” *Historical Studies in the Natural Sciences*, June 2018, 47:300–337. I thank Hepler-Smith for this observation.

⁶² See, e.g., Felisa Wolfe-Simon *et al.*, “A Bacterium That Can Grow by Using Arsenic Instead of Phosphorus,” *Science*, 3 June 2011, 332:1163–1166, which faced widespread criticism of its conclusions. See also Andrew Wakefield *et al.*, “Ileal-Lymphoid-

ities are more likely to receive unfavorable referee reports than their colleagues. Other observers have argued that current peer review procedures suppress innovative research, thus contributing to a public perception that most scientific research is irrelevant.⁶³ In 2011, Great Britain’s House of Commons commissioned a report on the state of peer review and concluded that while peer review “is crucial to the reputation and reliability of scientific research,” many scientists believe the system stifles progress, is often biased, and that “there is little solid evidence on its efficacy.”⁶⁴ There are also significant gaps between peer review in theory and peer review in practice. Although the system of peer review is allegedly supposed to determine what gets published and funded, program officers at funding bodies still shape the type of science their organizations support, and journal editors still retain tremendous power over what appears in their journals’ pages.⁶⁵ Some scientists wonder if systematic peer review is still the best method of evaluating scientific research and have argued that peer review could be eliminated with little cost to the quality of the scientific literature.⁶⁶

Scientists and laymen alike are frustrated when peer review does not instantly reward innovative papers with publication in a top journal or bestows grant money on a decorated scientist while leaving more ambitious work by a younger colleague unfunded. Many of the complaints about peer review’s perceived failures tend to assume that it is supposed to distinguish good science from bad science with perfect accuracy. But that is not the purpose for which refereeing was initially designed. The current “crisis” of peer review arguably has its origins in this moment in the 1970s, when the process was cast as the only acceptable method of evaluating scientific quality. The more we have expected of peer review, the more its opportunities to disappoint have expanded. Peer review’s perceived failures may have their roots in the gap between modern expectations of refereeing and the more modest functions it was initially designed to fulfill.

Nodular Hyperplasia, Non-Specific Colitis, and Pervasive Developmental Disorder in Children,” *Lancet*, 1998, 351:637–641, and Michael Lacour and Donald Green, “When Contact Changes Minds: An Experiment on Transmission of Support for Gay Equality,” *Science*, 12 Dec. 2014, 346:1366–1369, both of which were retracted by their publishing journals owing to accusations of fraud.

⁶³ For studies suggesting that women and minorities fare poorly in the peer review process see Lutz Bommann, Rüdiger Mutz, and Hans-Dieter Daniel, “Gender Differences in Grant Peer Review: A Meta-Analysis,” *Journal of Infometrics*, 2007, 1:226–238; Donna Ginther, Walter Schaffer, Joshua Schnell, Beth Masimore, Faye Liu, Laurel Haak, and Raynard Kington, “Race, Ethnicity, and NIH Research Awards,” *Science*, 19 Aug. 2011, 333:1015–1019; and Christine Wennerås and Agnes Wold, “Nepotism and Sexism in Peer-Review,” *Nature*, 22 May 1997, 387:341–343. For the concern that peer review procedures suppress innovative research see Paul Nightingale and Alister Scott, “Peer Review and the Relevance Gap: Ten Suggestions for Policy-Makers,” *Science and Public Policy*, 2007, 34:543–553; and Kyle Siler, Kirby Lee, and Lisa Bero, “Measuring the Effectiveness of Scientific Gatekeeping,” *Proceedings of the National Academy of Sciences*, 2015, 112:360–365.

⁶⁴ “Science and Technology Committee, Eighth Report: Peer Review in Scientific Publications,” United Kingdom Parliament, London, 2011, <http://www.publications.parliament.uk/pa/cm201012/cmselect/cmsctech/856/856.pdf> (accessed 29 July 2016).

⁶⁵ Baldwin, “Credibility, Peer Review, and *Nature*, 1945–1990” (cit. n. 5). For recent studies of editorial influence see Tommaso Colussi, “Social Ties in Academia: A Friend Is a Treasure,” *Review of Economics and Statistics*, 2018, 100:45–50; and Chad Wellmon and Andrew Piper, “Publication, Power, and Patronage: On Inequality and Academic Publishing,” *Critical Inquiry*, 2017, https://criticalinquiry.uchicago.edu/publication_power_and_patronage_on_inequality_and_academic_publishing/.

⁶⁶ Larry Wasserman, “A World without Referees,” *ISBA Bulletin*, 2012, 19:7–8.