The Role of the NIH in the Development and Promotion of Citation Indexing and Analysis as Scientometric Tools

The Role of the NIH in the Development and Promotion of Citation Indexing and Analysis as Scientometric Tools

Meghan A. Jendrysik, Ph.D.

A Thesis in the Field of History of Science, Technology and Medicine

for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

March 2020

Abstract

The NIH played a pivotal role in supporting Eugene Garfield's creation of the Science Citation Index (SCI). As well as being a valuable tool helping scientists to keep abreast of the current literature in their chosen fields, the SCI has for many years also served as the gold standard for a wide range of journal, article and even author rankings within the natural sciences. While supporting the development of the SCI advanced the NIH's mission as a research center and funding agency, its initial involvement owes more to the NIH's role and responsibilities as an agency of the federal government. Beginning in the early 1960s, the NIH was pressed by members of Congress to improve its communications programs; shortly thereafter, the NIH (jointly with the NSF) approved funding for Garfield's proposal. A decade later, the NIH began to use analyses of the data in the index it had funded to evaluate the efficacy of its peer review system of ranking grant proposals, and to compare the efficiency of different funding mechanisms and programs. Correspondence and other material preserved in the archives of the NIH show that these actions were taken primarily in response to pressure from Congress, rather than the scientific community.

Author's Biographical Sketch

Meghan A. Jendrysik was born and raised in the Boston suburb of Melrose, where she was valedictorian of her class at Melrose High School. She attended the Massachusetts Institute of Technology as a National Merit Scholar, and graduated in 1997 with a BS in Biology with a Minor in History. She continued her studies at the University of Illinois at Chicago, and earned a PhD in the field of Pathology in 2005. She completed a post-doctoral fellowship at the National Institutes of Health in 2011, and returned to Cambridge to join the publishing company Elsevier as a Scientific and Managing Editor in their biochemistry and applied biology journal portfolio.

She is an officer of the MIT Club of Boston, and in her spare time enjoys reading, sailing, and spending time in libraries.

Dedication

For my two families: the one I was fortunate enough to be born into, and the friends and loved ones who have become a second family over the years. Your encouragement and support made this easier than it looked.

And for Gram. It's about time.

Acknowledgments

I would like to thank my Thesis Director, Dr. Alex Csiszar, and my Research Adviser, Dr. James Morris, for their guidance (and patience). The Deo-Bennett family was kind enough to host me while I conducted my research in the National Archives facility in College Park, MD.

My employer, Elsevier, Inc. provided financial support for this work in the form of tuition assistance. My colleagues there have supported and encouraged me during the entire degree program.

Table of Contents

Chapter 1 Introduction

Scientific output is difficult to define, much less measure. While the benefits of scientific breakthroughs are easy to point to, it is harder to identify the many small steps that lead up to them. The polio vaccine is the result of years of patient, diligent research that eventually led to very public success and significant benefit to the world. It would have been very difficult to quantify the countless small successes and failures along the path that led to that first effective vaccine at the time they were being made. And yet, this is what researchers must do in order to advance their careers and justify the continued expenditure of grant monies—grants that for American scientists in the second half of the 20th century are largely provide by federal agencies. As a major funder and administrator of research, the U.S. government has played a part in fostering a growing demand for indicators of scientific progress, or scientometrics.

Scientometrics have been utilized to gain in some insight into the effectiveness of various programs and funding mechanisms. Some national governments began conducting surveys to determine the extent of R&D resources available in their countries, and how they were being applied, as early as the 1920s. These early scientometrics had mainly focused on input measures—such as how many scientists and engineers were working in which field, and how much money was being spent on it. In the 1950s, surveys of the U.S. and certain other countries of interest (notably Eastern Bloc states that were seen as potential threats) were carried out by the National Science Foundation (NSF) and the National Research Council (NRC). Economic indicators of scientific output started to make an appearance in the early1970s, and science indicators became more policy-oriented beginning in the 1990s[1]

At the same time, Social Science Indicators were being developed in the U.S. and around the world. This new field sought to measure the "the quality of life," by identifying indicators that could serve as proxies for happiness, prosperity, and similar intangible but highly desired qualities in order to evaluate specific programs, develop a "balance sheet or system of social accounts," and to determine appropriate goals and program priorities.[2] These indicators came to include the quantification of the output of the scientific community both as a national resource, as well as specific individuals and institutes. In 1963, the Organisation for Economic Co-Operation and Development (OECD) conducted its first international survey of R&D efforts, and "conventionalized" the use of the Gross Domestic Expenditures on R&D as a percentage of Gross National Product (GERD/GNP) as an indicator. This indicator was used both to compare the relative amounts of investment in research, and to justify increased funding for agencies that were spending less than their peers.[3]

These attempts at estimating the size of the scientific efforts were largely focused on the input—that is to say, the resources (including manpower) that were dedicated to scientific and engineering enterprises. Eventually, there began to be calls for better measure of the outputs, but the OECD instead focused on "impacts," defined by Benoît Godin as "indirect effects upon society and the economy."[4] Direct outputs were significantly more difficult to measure; surveys

---

[1] Benoît Godin, *Measurement and Statistics on Science and Technology : 1920 to the Present*, 1st ed. (New York: Routledge, 2005), 7–8.
[2] Kenneth C Land, "On the Definition of Social Indicators," *The American Sociologist* 6, no. 4 (1971): 322.
[3] Benoît Godin, "The Numbers Makers: Fifty Years of Science and Technology Official Statistics," *Minerva* 40, no. 4 (2002): 391–92.
[4] Godin, 394.

met with limited success, and available bibliographic and patent databases had been compiled by independent groups largely for administrative purposes, and the OECD was skeptical of their value.[5] The OECD did not consider publications by themselves to be output, and therefore did not include them in their statistics. The OECD's "Guidelines for Collecting and Reporting Data on Research and Experimental Development" (the Frascati manual) recommended that documentation activities (including publication of results) should be counted and reported in parallel with the other statistics, but this recommendation was not acted upon. Instead, scientific publication was relegated to the category "related scientific activities."[6] Not all agencies agreed with the OECD that publications were not worth considering as a type of output. The NSF included a section on metrics associated with publications and citations, or bibliometrics, in its first Science Indicators report in 1972, but since so few countries accepted this as a reliable indicator, the OECD did not adopt it in a systematic manner (though it did collect bibliometric data on occasion).[7]

The development of bibliometrics is of particular relevance, as its growth was fostered and encouraged by the National Institutes of Health (NIH). The NIH, together with the NSF, financially supported Eugene Garfield's efforts to create "maps" of scientific frontiers based upon citation patterns and networks as part of an effort to improve communications between scientists. Garfield's citation index would eventually give rise to the Journal Citation Reports (JCR) and the Journal Impact Factor (IF). Bibliometrics was not a new field at the time, but the advent of the Computer Age made it possible to apply this type of data-driven analysis across larger numbers of journals, articles and authors. Starting in the 1950s, Eugene Garfield had

[5] Godin, *Measurement and Statistics on Science and Technology : 1920 to the Present*, 39.
[6] Godin, 75–75.
[7] Godin, 197.

advocated for the creation of a citation index for scientific publications, similar to Shepard's Citations, and index widely used for case law. His intent was not to quantify output or rank researchers or journals initially, but to create a map of science that could be used to identify the most fruitful areas of research. In 1960 he was granted funds by both the NSF and the NIH to create such an index for the field of genetics research, and he realized it was easier to first index the whole field of natural science research, then extract the information relevant to the specific area of genetics, rather than build the index out of field-specific blocks like Legos. This effort led to the construction of the first Science Citation Index (SCI) in 1963. The first JCR, based on the data collected in the SCI, was published in 1972, and became an annual publication in 1975. Though the SCI was not initially created for the purposes of ranking journals, articles, researchers, or institutions, the Impact Factor (IF) as reported in the JCR has become the most commonly used indicator for journal quality in the U.S.

It is nearly impossible to overstate the role of the NIH in the field of life science and health-related research today. Not only has the NIH been directly responsible for the funding of basic and clinical research both on its own campus and around the world, a significant portion of academicians and physician-researchers have spent some part of their early career being trained on campus or in workshops. The priorities and requirements set by the NIH can therefore hold sway over the careers and research programs around the country. In order to understand why the NIH extended support to Garfield when it did, it is helpful to look at how the NIH rose to its position of prominence, and how the expectations of it changed over time.

It was not always the case that the federal government was the main source of funding for research—it only became so in the second half of the 20th century. The NIH traces its roots to the founding of the Marine Hospital Service (MHS) in 1798, but that was strictly concerned with commercial ventures. The MHS was part of the Treasury Department, and was initially tasked with providing for the health care of merchant seamen. Later Congress gave them the additional responsibility of examining arriving passengers for signs of epidemic diseases such as cholera and yellow fever. In the late 19th century, Joseph Kinyoun expanded the MHS by adding a "laboratory of hygiene" based on the facilities he had trained in in Germany. Within months, he succeeded in isolating the bacillus causing cholera, paving he way for a definitive diagnosis of cholera. Despite this milestone, the funding for the research activities of the MHS and Hygienic Laboratory remained very modest.

During the early 20th century, interest in the progress of science led to an expansion of the funding and support for such endeavors. Between 1900 and 1939, more than 40 scientific agencies were established within the US federal government, and private philanthropy also increased, notably with the foundation of the Rockefeller Institute for Medical Research (RIMR) in 1901. RIMR was established along similar lines of the Pasteur Institute in Paris, with permanent staff researchers who were largely left to their own devices to determine the scope and focus of their research. This was followed in 1913 by the establishment of the Rockefeller Foundation (RF), which would shortly begin providing support for research taking place outside of the walls of the RIMR. This program would later serve as an example for the NIH extramural programs.

Meanwhile, the MHS had been reorganized into the Public Health Service (PHS), and had been authorized to conduct research into non-contagious diseases and pollution of lakes and streams, but true to its name it remained primarily concerned with maintaining a healthy environment for the U.S.'s citizens and industries. During World War I, it was primarily responsible for sanitation at military installations, and treating patients in Washington, D.C. during the 1918 influenza pandemic.

Between the wars, the Rockefeller philanthropies expanded their programs both in scope and size, adding a fellowship program administered by the NRC, block grants to universities, and creating research institutes at universities (to complement the RIMR). Changes in the tax laws in the late 1920s and the onset of the Great Depression led to a decrease in available funds, but even so, in 1934 the Rockefeller philanthropies were responsible for roughly two-thirds of all funding in the fields of medicine and public health, as well as 72% in the physical and biological sciences.[8]

While the RF served as a model for later government programs, it could not provide much guidance in evaluating the success of such programs. Public support for research was also growing at this time. In 1926 two influential acts were introduced in Congress: the Parker Bill and the Ransdell Bill. The Parker Bill would have expanded the existing PHS facilities, while the Ransdell Bill sought to establish an entirely new agency "…to study all diseases of human beings."[9] While the Parker Bill did not pass, the Ransdell Act was signed into law in 1930, formally establishing the National Institute of Health. Together these two bills initiated an ongoing debate about the role of the federal government in directing and supporting research at both the basic and the translational levels. The National Cancer Institute was founded as an independent institute in 1937, establishing what would become the general model of the future NIH (which it would join in 1944) with a categorical-disease structure, and the ability to provide

---

[8] For information on the administration of the RF and the early days of NIH, see William H Schneider, "The Origin of the Medical Research Grant in the United States: The Rockefeller Foundation and the NIH Extramural Funding Program.," *Journal of the History of Medicine and Allied Sciences* 70, no. 2 (April 2015): 279–311, https://doi.org/10.1093/jhmas/jrt074.
[9] Victoria A. Harden, *Inventing the NIH: Federal Biomedical Research Policy, 1887-1937*, 1st ed. (Baltimore: The Johns Hopkins University Press, 1986), 99.

grants for non-federal scientists working at universities. While this move was extremely popular inside the halls of government (the Ransdell Bill was sponsored by every Senator unanimously), the American Medical Association (AMA) was concerned about the dangers inherent in having the government in control of medical research.[10]

During and immediately following World War II, the question of the federal government's support for scientific research became less a question of if, but of how. During the war, the needs to be addressed were immediate, urgent, and somewhat self-evident. These included military objectives in terms of the development of new weapons and communications technology, and treating wounds received in battle, as well as initiatives to improve the general health of soldiers and draftees who were deemed unsuitable due to poor health related to infectious and communicable diseases or lifestyle issues like poor nutrition and dental hygiene. The NRC's Division of Medical Sciences formed a total of 13 committees and 43 subcommittees to address these issues.[11]

Towards the end of the WWII, President Franklin D. Roosevelt asked Vannevar Bush how the lessons he had learned during his time Director of the Office of Scientific Research during wartime "…could be applied in the days of peace that lay ahead."[12] Bush's report contained several recommendations on how investment in research and education in the sciences could contribute to the military and economic strength of the nation. The predicted benefits included longer lifespans, better salaries, and of course, international preeminence. While it is easy to understand that these would lead to a general improvement in the quality of American

[10] Stephen P. Strickland, *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy*, 1st ed. (Cambridge: Harvard University Press, 1972), 14.
[11] Vannevar Bush, *Science, the Endless Frontier; a Report to the President on a Program for Postwar Scientific Research* (Washington: National Science Foundation, 1960), 53.
[12] Bush, vii.

life in the decades to come, there was no specific mention of how these benefits were to be measured. Bush felt strongly that science could not be planned, and should be directed by scientists, rather than bureaucrats. "In Bush's ideal universe… science was not the servant of man; instead, researchers were accountable only to each other rather that to their public patron."[13] The benefits of research were clear to Bush. The immediate challenges as he saw them were fairly well-defined: a generation that had delayed college needed education, and so would their children, and with infectious disease soon to be a thing of the past (thanks to antibiotics), heart disease and cancer were the next targets in the medical arena. The war had highlighted ways in which the American science had depended on European institutions, which was no longer considered tenable. It was time for America to develop its own institutions and talent base. Although very little of what Bush proposed was ever implemented, the ideals described in *Science—The Endless Frontier* were deeply influential in post-war era policy.[14]

While it took a few more years for the NSF to be established to act on the recommendations described in Bush's report to Truman (Roosevelt having died before it could be completed), the NIH and the NCI (still separate institutes) had already spun up their research programs, snapping up wartime contracts and turning them into peacetime grants. Between 1943 and 1959, the benefits medical research could be seen in improvements in the overall health of Americans. In addition to the first polio vaccine, advances began to be seen in the areas of heart

[13] Michael Aaron Dennis, "Historiography of Science: An American Perspective," in *Science in the Twentieth Century*, ed. John Krige and Dominique Pestre (Amsterdam: Harwood Academic Publishers, 1997), 17.
[14] Dennis, 7.

disease and cancer, as better understandings of the causes led to better treatment and prevention strategies. Throughout this period, lifespans of Americans increased by 6.4 years.[15]

The amount of money being spent on research by the federal government in this period was significantly increasing. While some of this could be attributed to increased overhead, and the vast expansion of the physical campus of the NIH and creation of the extramural research program, there was a growing concern that perhaps this spending was not leading to any benefits. Beginning in the early1960s, the government began looking more closely at how public funds were being spent in research. Their investigations failed to uncover any evidence of significant mismanagement, and were generally favorable, but they did have the effect of encouraging the NIH to find ways to demonstrate that their procedures were sound, and that they were acting as responsible stewards of the public's money.

While the decision of what to publish, when, and in which journals has always been solely the purview of the individual researchers, such decisions are not made in a vacuum. Tenure committees consider publication records in making decisions that have a lasting impact on the career of a researcher, and the phrase "publish or perish" has been used to describe the pressure felt by academic researchers in many fields, including the natural and life sciences. The visibility gained by publishing in high profile journals such as *Cell*, *Nature* and *Science* goes a long way to establish a young researcher's career. Conventional wisdom holds that it is extremely difficult to find funding for research that has scant chance of being published, such as confirmatory studies, or reports of negative results.

---

[15] Data from the National Office of Vital Statistics, as cited in: National Health Education Committee, *Facts on the Major Killing and Crippling Diseases in the United States Today; Heart Diseases, Cancer, Mental Illness, Arthritis, Blindness, Neurological Diseases and Other Health Problems* (New York, 1961), sec. Does Medical Research Pay Off?

The NIH's (and, to a lesser extent, NSF's) role in the emergence of modern bibliometrics had two separate stages; both in response to calls for increasing accountability to the American taxpayer and government. The first stage was in the early 1960s, in response to a House of Representatives Committee's identification of the communications program as a weakness within the NIH organization. This led to stronger support for programs associated with the indexing and tracking of publications and citations, including reversing the previous decision to not fund Garfield's indexing project. The second stage began in the late 1960s and stretched into the early 1970s, and was motivated by two different factors: lists of Social Indicators were beginning to include measures of scientific productivity and capacity, and the federal research budget was shrinking. These two factors led to two demonstrations of the faith the NSF and the NIH had for bibliometrics. The first Science Indicators report compiled by the NSF in 1972 included a section of bibliometrics as indicators of the relative size and output of scientific endeavors in the U.S. compared to other nations as the main measure of output. Meanwhile, the NIH had commissioned the RAND Corporation to investigate its peer review system of awarding grants, to confirm that it was effective at selecting projects that resulted in highly cited papers, and to estimate the impact of the smaller predicted budgets. RAND's researchers relied heavily on their own citation analysis (using data from the SCI) to define successful or highly valued research projects. The implied acceptance of bibliometrics and citation analysis for both the NSF's and RAND's reports is an indication that these were accepted as legitimate indicators of scientific productivity and value, respectively.

Support from the NIH played a significant role in the development and widespread acceptance of citation analysis as a valid indicator of value within the scientific community. This support was motivated in large part by the NIH's need to demonstrate to Congress that it was fulfilling its mission responsibly and efficiently. This thesis will focus on exploring how responding to Congress, and specifically to the Fountain Committee of the House of Representatives, led the NIH to support first the development of Garfield's citation index, and later to the use of citation analysis based on this index as a tool for shaping policy.

I will begin with an overview of the development of scientometrics and bibliometrics, with a focus on how the two fields overlap and have developed in tandem. This will be followed by a description of the relevant findings of the Fountain Committee, and the NIH's responses to it. The final section will address the impact the Fountain Committee and the responses had on the NIH's communications policy and programs. Primary source material includes official reports made by the Fountain Committee as part of the Federal Register, hearing transcripts, correspondence and other archival material from the NIH archives now housed by the U.S. National Archives.

Chapter 2 Bibliometrics, Scientometrics, and Social Indicators

The terms "scientometrics" and "bibliometrics" were coined in 1969 by Vassily Vassilievich Nalimov and Alan Pritchard, respectively, but attempts to measure scientific output and progress and the analysis of the published literature for various purposes began long before then.[17] The sociologists of science of the mid-20th century were building on the work of earlier attempts to catalog scientific knowledge and examples of scientific genius in the early 20th century either to inform government policy or to advocate for additional support. Some of these focused on the economic cost of science (Julian Huxley's *Scientific Research and Social Needs*, 1934, J.D. Bernal's *The Social Function of Science*, 1939, and Frederick Rose's *Report on Chemical Instruction in Germany*, 1901), while others tried to construct a natural history of scientific genius (Francis Galton's *English Men of Science: Their Nature and Nurture*, 1874, and James McKeen Cattell's series *American Men of Science*).[18]

In 1926, the statistician Alfred J. Lotka published a more general look at publications as metric of scientific productivity in a paper titled, "The Frequency Distribution of Scientific Productivity."[19] In this article Lotka defined what is now known as "Lotka's Law": that the number of authors (P) who publish N articles is inversely proportional to N^2. Throughout the rest

---

[17] Yves Gingras, *Bibliometrics and Research Evaluation: Uses and Abuses* (Cambridge: The MIT Press, 2016), 1.
[18] Yehuda Elkana et al., eds., *Toward a Metric of Science* (New York, New York: John Wiley & Sons, Inc., 1978), 17.
[19] Alfred J. Lotka, "The Frequency Distribution of Scientific Productivity," *Journal of the Washington Academy of Sciences* 16 (1926): 317–23, https://doi.org/10.2307/24529203.

of the 1920s and 1930s, the systematic study and analysis of the use of scientific literature became more widespread and more sophisticated, as libraries sought to cope with rising subscription costs and an ever-increasing number of specialized journals. These studies were not geared at directing policy, or providing insight to researchers trying to assess the current state of knowledge on a given topic. Most of these studies were designed to help librarians make decisions about which journals would be worth the price of a subscription.

Libraries were interested in these studies due to the expanding number of scientific publications. The eminent sociologist of science, Derek J. de Solla Price described what he saw as one cause of this expansion in *Little Science, Big Science*, According to Price, the rate of expansion of the population and the numbers of working scientists means that approximately 80%-90% of all scientists ever living are living right now, whenever right now is.[24] This expansion in the size of the scientific community is naturally accompanied by a proportional increase in the number of scientific publications. The impact this growth had on the culture of science was central to Price's hypothesis in *Little Science, Big Science;* this is made clear in the foreword by Robert K. Merton and Eugene Garfield, "…this book is dedicated to establishing and interpreting the magnitudes of growth in 'the size of science': in the numbers of scientists and scientific publications and in the societal resources allocated to the pursuit of science and science-based technology."[25] These two giants in the development of bibliometrics go on to claim that the use of the analysis of citation patterns to define "research fronts" is what drew many new scholars to the quantitative study of science, making it one of the most important

---

[24] Derek J. de Solla Price, *Little Science, Big Science... And Beyond* (New York: Columbia University Press, 1986), viii.
[25] Price, ix.

developments in the sociology of science.[26] It is not surprising that Garfield would include publication and citation analysis so prominently in this foreword. Price and Merton were among the first to comprehend both the usefulness of Garfield's index as a tool for the scientometrics, as well the potential for its misuse.

In the preface to *Little Science, Big Science*, Price outlines his intention of formulating and integrated measure of science: "I shall attempt to develop a calculus of scientific manpower, literature, talent, and expenditure on a national and on an international scale."[27] It is worth noting that of the variables included in Price's proposed "calculus," scientific literature is the only one that can be considered to be on the output side of the equation; the main product being created by the combined application of manpower, talent, and expenditure. Price had long been fascinated by the quantitative study of scientific literature as a means to determine the size of the scientific community or enterprises. He describes his own interest in scientometrics as stemming from his observation that his copy of the complete run of the *Philosophical Transactions of the Royal Society of London*, organized in stacks by year of publication described an exponential curve.[28] He goes on to state that his thinking was influenced by "…two new fields emerging as part of the academic explosion of the 1960s, the sociology of science and library science (as distinct from library trade schools). Those… seemed to react almost alchemically… in what was to become science of science or scientometrics."[29]

Price's stacks of the *Philosophical Transactions* became for him an illustration not just of the expansion of science, but of the length of time that this expansion has been going on. From his assertion in the early 1960's that 80-90% of all scientists ever living are alive right now

---

[26] Price, x.
[27] Price, xvi.
[28] Price, xix.
[29] Price, xx.

(whenever "now" is) he extrapolated that there is also a concomitant exponential increase in the literature published. As a result, "…[s]cientists have always felt themselves to be awash in a sea of scientific literature that augments in each decade as much as in all times before."[30]

Price was enthusiastic about the potential of bibliometrics as a tool for analyzing scientific fields as a whole, and he felt that the quantity of scientific publications overall might have a positive correlation with scientific progress.[31] However, he also believed it was not necessarily the case that merely counting the number of publications authored by an individual researcher is a good indicator of their abilities or quality, although publication lists are used by academic departments and other organizations as part of their tenure reviews, promotion, and other decisions. According to Price, "[t]he scale is bad if for no other reason than that its existence has moved people to publish mere because this is how they may be judged."[32] Merton shared his concerns about the potential for goal displacement with Garfield:

> Whenever an indicator comes to be used in the reward system of an organization or institutional domain, there develop tendencies to manipulate the indicator so that it no longer indicates what it once did. Now that more and more scientists and scholars are becoming acutely conscious of citations as a form of behaviour, some will begin, in semi-organized fashion, to work out citation-arrangements in their own little specialties.[33]

In other words, using citation analysis as a measure of a scientist's productivity could potentially lead to scientists making deals to boost one another's citation rate as a favor. Another potential risk was what Merton called "the Matthew effect," as the scientists who publish more receive more recognition, are more likely to be invited into collaborations, and be provided with more

---

30 Price, 13.
31 Price, 69.
32 Price, 36.
33 Alex Csiszar, "From the Bureaucratic Virtuoso to Scientific Misconduct: Robert K. Merton, Eugene Garfield, and Goal Displacement in Science," in *Annual Meeting of the History of Science Society* (Toronto, 2017).

resources-- all of which will lead to more opportunities to publish.[34] In this way, there might in fact be a relationship between the number of publications and the success of the scientist.[35]

In addition to considering the value of a publication to the author, Price was also interested in determining the value of individual publications to the community. As there are a great number of journals, each with their own editorial standards, it stands to reason that not all papers and journals will be valued equally by the community. In order to get some idea of the relative value of journals to the community, Price turned to a study by D.J. Urquhart in 1958 on the requests for periodicals over the course of a year in the Science Library in London and found that the requests per journal described a Pareto curve, with half of all requests being made for the top 40 journals (out of 9120 scientific periodicals).[37] Over half were not used at all in the time period of the study. More than 80% of the requests were for less than 10% of the available journals. Price drew two conclusions from this: there is a high level of variability in the value to the community of publications (here referring to both articles and journals), and that publishing might be more valuable to the author than to the reader/community. At the time, there was no way to look at similar statistics at the level of individual articles, a gap that would be filled by Garfield's index.[38]

Price interpreted the proportion of highly requested journals to mean that in broad terms a greater number of papers and journals would result in a correspondingly greater number of high value papers in a more-or-less fixed proportion, and that proportion would be small. This

---

[34] Robert K Merton, "The Matthew Effect in Science," *Science, New Series* 159, no. 3810 (1968): 56–63.
[35] Price, *Little Science, Big Science... And Beyond*, 37.
[37] D.J. Urquhart, "Use of Scientific Periodicals," in *Proceedings of the International Conference on Scientific Information* (Washington, D.C.: National Academies Press, 1959), 287–300, https://doi.org/10.17226/10866.
[38] Price, *Little Science, Big Science... And Beyond*, 67.

raises the question of what motivates scientists to publish, given that so many publications appear to be of little value to the community.[39] Price defined two separate roles for the scientific publication: the informational and the social. The informational function is the primary one that people have traditionally associated with the scientific article: communicating new advances in knowledge to the community. The social function related to "…the establishment and maintenance of intellectual property… the need which scientists felt to lay claim to newly won knowledge as their own, the never-gentle art of establishing priority claims."[40] He posits that the primary role of scientific publication has shifted from the informational to the social in the modern world. He attributes the changes in the role of publication due to increased availability of resources:

> In many ways the modern ease of transportation and the affluence of the elite scientist have replaced what used to be the effected by the publication of papers. We tend now to communicate person to person instead of paper to paper. In the most active areas we diffuse knowledge through collaboration. Through select groups we seek prestige and the recognition of ourselves by our peers as approved and worthy collaborating colleagues. We publish for the small group, forcing the pace as fast as it will go in a process that will force it harder yet. Only secondarily, with the inertia born of tradition, do we publish for the world at large.[41]

By "small group," Price means the active researchers who would be potential collaborators for the author. Depending on the size of the field and the prominence of the author, the findings being reported might already be known to some, or all, active researchers by the time of publication. Publication, in this scenario, is more about promoting oneself as a worthy collaborator for others in this select group.

---

[39] Price, 69.
[40] Price, 58–59.
[41] Price, 80.

In 1955 Eugene Garfield proposed "…a bibliographic system for science literature that can eliminate the uncritical citation of fraudulent, incomplete, or obsolete data by making it possible for the conscientious scholar to be aware of criticisms of earlier papers." This notion that science is self-regulating with respect to both intentional fraud and honest error was shared by many, including Vannevar Bush. Bush believed that the scientific community must be permitted to be self-regulating in this regard, as well as in larger questions of resource allocation: "In the event that scientific method failed to keep scientists from making errors, the community would step in to sift the good from the bad. Errors would be weeded out by reviewers or fail the test of replication and be expelled from the body of scientific knowledge."[42]

Garfield was inspired in his citation index concept by *Shepard's Citations*, a research tool that was widely used in the legal profession to quickly identify relevant cases and decisions.[43] Garfield struggled to get NSF support for his vision, but he was unsuccessful for years (he believed that his status as a businessman rather than an academic hurt him in their eyes). In 1959 he began a correspondence with Joshua Lederberg, who was impressed with his 1955 description of the proposed index in *Science*. Lederberg suggested that if the NSF wouldn't provide funding, perhaps Garfield should apply to the genetic study section at NIH. The NIH agreed to provide $300,000, together with the NSF, and the work began on the 1961 Science Citation Index with a subset for genetics citations. [44] Funding from both the NIH and the NSF provided a de facto

---

[42] Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*, Course Boo (Princeton: Princeton University Press, 1996), 218.
[43] Eugene Garfield, "Citation Indexes for Science; a New Dimension in Documentation through Association of Ideas.," *Science (New York, N.Y.)* 122, no. 3159 (July 15, 1955): 108–11, https://doi.org/10.1126/SCIENCE.122.3159.108.
[44] Henry Small, "Citation Indexing Revisited: Garfield's Early Vision and Its Implications for the Future," *Frontiers in Research Metrics and Analytics* 3 (March 2, 2018): 8, https://doi.org/10.3389/frma.2018.00008; Gingras, *Bibliometrics and Research Evaluation: Uses and Abuses*.

endorsement of Garfield's work that might have been lacking had he relied entirely on private capital. Price was one of the early adopters of this new tool, and used it in his 1965 *Science* article "Networks of Scientific Papers" to analyze the literature citation patterns of scientists, and to break it down into two distinct patterns reflecting two different needs: the relatively "tight" network of very recent work representing the "research front," and the more widely distributed network drawing on the totality of the literature in a given field.[45]

Originally proposed as an "association-of-ideas index," or a way to map the structure of scientific knowledge, Garfield also envisioned his citation index as a tool to improve communication between scientists and increase discoverability of their work.[46] Of course, citation analysis could also be used to quantitatively determine the relative impact of journals and individual scientists. Although such uses were controversial, the SCI proved to be a useful tool for sociologists of science and university and program administrators. In 1972 Garfield published the first JCR, ranking journals according to an Impact Factor (IF) based on citation rates. Since then the IF has become an important annual measure of individual journal's overall quality.

The social indicators movement had gained momentum in the 1960s as the demand for reliable measures of what had recently been termed the "quality of life" to inform public policy became more urgent. Previous attempts at this end had begun at least by 1927, when W.F. Ogburn edited a series of special issues of the *American Journal of Sociology* dedicated to recent social changes. These annual reports included lists of the major technological and scientific

---

[45] Derek J D Price, "Networks of Scientific Papers: The Pattern of Bibliographic References Indicates the Nature of the Scientific Research Front.," *Science* 149, no. 3683 (1965): 510–15.
[46] Garfield, "Citation Indexes for Science; a New Dimension in Documentation through Association of Ideas."

developments of the previous year.[47] These lists comprised a sort of qualitative indicator of scientific progress.

NASA made a foray into the field of social indicators by commissioning the Social Science Research Council (SSRC) to demonstrate the benefits of the Space Program to society at large. They were unable to perform such a far-reaching analysis outside of an analysis of the full set of social indicators in our society.[48] The result was a volume edited by Raymond Bauer called simply *Social Indicators* in 1966. This expanded on previous efforts to evaluate the state of American society such as the Hoover administration's *Recent Social Trends in the United States* and the more recent annual *Trends* and monthly *Indicators* produced by the Department of Health, Education, and Welfare (HEW) starting at the beginning of the Kennedy administration.[49]

Congress required the National Science Board (NSB) to measure and provide a status report on American science and technology—the first of a series of Science Indicators reports was published in 1972 (SI-72).[50] It had an ambitious scope, with the goal of presenting:

> …the first results from a newly initiated effort to develop indicators of the state of the science enterprise in the United States. The ultimate goal of this effort is a set of indices which would reveal the strengths and weaknesses of U.S. science and technology, in terms of the capacity and performance of the enterprise in contributing to national objectives. If such indicators can be developed over the coming years, they should assist in improving the allocation and management of resources for science and technology, and in guiding the Nation's research and development along paths most rewarding for our society.[51]

---

[47] Otis Dudley Duncan, "Science Indicators and Social Indicators," in *Toward a Metric of Science: The Advent of Science Indicators*, ed. Yehuda Elkana et al. (New York: John Wiley & Sons, Inc., n.d.), 33.
[48] Elkana et al., *Toward a Metric of Science*, 35.
[49] R. Bauer and American Academy of Arts and Sciences, eds., *Social Indicators* (Cambridge: MIT Press, 1966).
[50] Gingras, *Bibliometrics and Research Evaluation: Uses and Abuses*, 7.
[51] National Science Board, "Science Indicators, 1972: Report of the National Science Board." ([Washington, 1973), iii.

This is admittedly an ambitious set of objectives, and the report's authors acknowledge that there are significant shortcomings in the report, mainly due to a lack of reliable metrics for scientific output. As a result, the report focuses largely on the resources allocated to research and scientific manpower. The few outputs that were reported were primarily based on readily available data (including the SCI), and included "quantity and quality of scientific publications, patent output, and trade in technical knowledge."[52] Most of the output measures were used to compare the relative output of U.S. scientists compared to those in the U.K., Germany, France, the U.S.S.R. and Japan. Specifically, the significance of publications was determined by a ratio of citations per publication, as "[t]he general rationale for such an index is the expectation that the most significant literature will be most frequently cited, whereas relatively unimportant research articles will attract few, if any, citations. Support for the validity of this indicator is the high correlation, found in a number of studies, between the significance of papers as judged by researchers in the field."[53]

Shortly after SI-72 was released, Dr. Eleanor Bernert Sheldon, the President of the SSRC gathered an interdisciplinary group of Fellows of the Center for Advanced Study in the Behavioral Sciences at Stanford, CA to a conference in 1974 to investigate the nascent field of science indicators, and SI-72 in particular. The invitation challenged the Fellows (including Merton, Garfield, Lederberg, Price, and Merton's students Stephen Cole, Jonathan Cole, and Harriet Zuckerman) "…to pose the question, 'What must one look at in order to estimate the condition of science as an intellectual activity or as a social institution?' We think of this question within a broad historical and sociological frame rather than from a delimited point of

---

[52] National Science Board, viii.
[53] National Science Board, 10.

view dealing with the present inputs to and outputs of science measured in terms of men, money, and materials."[54]

The sociology of science had been moving in the direction of looking to publication information as an output measure for some time. The innovation of Garfield's index made it considerably more practical, not just for sociologists interested in studying science, but also for administrators and policy agencies trying to quantitate scientific progress. However, the specific motivation behind the NIH's decision to support Garfield did not come from the philosophy or sociology of science, but instead was a response to requests from Congress in its role as the keeper of the public purse. The NIH granted individual researchers a high level of independence in directing their NIH-funded research programs, but the NIH as a federal agency was answerable to Congress and ultimately the American people, and must be able to show that public money is being spent responsibly and in the public's interest. This relationship would foster the development of the SCI, and later provide a tacit endorsement of citation analysis as a sound basis for policy decisions.

[54] Elkana et al., *Toward a Metric of Science*, ix.

Chapter 3 Calls for Accountability

The relationship between the Federal Government and universities was undergoing a sea change in the decades after the World War II. In late 1960, the President's Science Advisory Committee (PSAC) presented a report on that relationship, which included a brief discussion of the differences between work performed under a government contract, and that funded by a grant, concluding that the two could really not be considered equivalent:

> Yet in its essence the concept of 'purchase of services,' which is implied in any government contract, was and is a doubtful one, when applied to basic research. Basic research, almost by definition, has no clearly predictable practical result, and so the Congress and the Federal agencies involved have had to interpret very broadly the notion of 'value received' in return for sums spent on research contracts… The whole framework is somewhat arbitrary and unrealistic. The wonder is that it works as well as it does.[56]

Funding under a contract system is essentially like any other purchase—payment is exchanged for specific goods or services. Grants, in contrast, are given to individuals or institutes in support of a proposed line of inquiry—the endpoints are indefinite and the investigators are allowed a wide amount of latitude to adjust or amend the plan of action. During WWII, most government support for science was managed through contracts. The post-war years saw a shift to grants, along with the increase in basic research funding. Since grants lack defined deliverables, determining the return on investment was less straightforward.

---

[56] President's Science Advisory Committee, "Scientific Progress, the Universities, and the Federal Government" (Washington, D.C.: United States Government Printing Office, 1960), 7.

The NIH's expenditures have always been subject to federal oversight, which is not an easy task. The NIH's mission is somewhat nebulous (as well as ambitious): both to improve health and increase longevity, and to advance our understanding of the science of health and disease. This is a very tall order, and progress is not easy to evaluate beyond broad-stroke measures such as increases in the expected life span or decreases in disease-related morbidity and mortality. However, the NIH enjoyed strong support in Congress due to champions in the Senate such as Lister Hill and Warren Magnuson, and so the budget continued to grow.

Given this, it should not have been terribly surprising when in 1956 the House Committee on Appropriations directed the Secretary of Health, Education and Welfare to make an inquiry into the NIH's programs.[57] The committee was primarily interested in how the money allocated was being spent, and "…expressed more than a casual interest in the Institute program evaluation between Study Section and Council, i.e., what the staff does following a Study Section meeting and prior to a Council meeting."[58] This put the NIH directors in the position of having to demonstrate to non-scientists that their programs and procedures were effective and a wise use of public funds. This question was at the center of an investigation led by Representative Lawrence H. Fountain (D, N.C.), that would eventually lead to the NIH's first investment in Garfield's SCI.

Fountain Committee (1961)

---

57 Charles G. Haynes, "Letter to Secretary Folsom, September 12, 1956," 1956.

58 "Notes for Dr. Shannon on Mr. Siepert's Presentation to Institute Directors on the House Investigations Committee," 1956, 2.

In 1959, Rep. Fountain (D, N.C.), the chair of Intergovernmental Relations Subcommittee of the House Government Operations Committee, began a more in-depth inquiry into the management of financial matters related to NIH external grants after learning that grantees could adjust their budgets without having to get additional approval.[59] His purpose was to find ways to tighten the federal budget without decreasing the effectiveness of the programs being supported by the NIH. After two years, the final report was transmitted to the Speaker of the House on April 28, 1961. Though tasked with looking at both "economy and efficiency" in the NIH extramural operations, the committee did not attempt to make any evaluation regarding the scientific progress or merit of these operations. As stated in the Preface to the report, "…the committee's findings and recommendations are concerned primarily with administrative policies and practices. The committee necessarily leaves to those with scientific competence the task of judging whether the individual projects receiving financial support are scientifically sound and meritorious."[60] The report does include some discussion regarding the benefits of the NIH's programs to the nation: a healthy population, and lower health care costs (which were becoming an increasing burden on the federal government via Veterans' Administration and other hospitals, public assistance and other aid programs). The committee thus concurred that "…support of an intensive effort for the conquest of costly diseases represents a sound public investment,"[61] and the major source of government funding in this area was the NIH.

---

59 Melinda Baldwin, "Scientific Autonomy, Public Accountability, and the Rise of 'Peer Review' in the Cold War United States," *Isis* 109, no. 3 (September 19, 2018): 549, https://doi.org/10.1086/700070.
60 Committee on Government Operations, "Health Research and Training: The Administration of Grants and Awards by the National Institutes of Health: Second Report," vol. 1, 1961, VII.
61 Committee on Government Operations, 1:3.

That support had been expanding since the end of WWII, from $8.5 million representing 10% of the funding for research in this field in 1947 to $285 million, or 40%, in 1960. Most of the increased spending went to extramural programs, which increased 13 to 16 fold during the 1950s. This expansion allowed the NIH's Division of Research Grants to improve the majority of grants that were applied for (81% in 1960). Spreading the grant money out to medical schools and universities around the country helped establish departments, launch careers, and in general build the U.S.'s capacity for conducting this type of research. This was in line with the government's, and NIH Director James A. Shannon's, goal of establishing the U.S. as the global leader in health research. However, with this prominence came a great responsibility, as Director Shannon observed in his statement, "[NIH's] policy decisions, and even the cumulative effect of its routine administration actions, now have a significant impact not only on the course of the Nation's medical-research effort but on the training provided by medical schools—and, increasingly, by university bioscience departments as well—and on the deployment of trained manpower in the bioscience field."[62]

Some in Congress were concerned that such a large and rapid expansion would put too much strain on the scientific review process for evaluating grant proposals. Director Shannon agreed with this in principle, saying "The problems arising in connection with the administration of medical research and research grant programs loom much larger today than they did even a few years ago because of the extraordinarily rapid growth of medical and related biological research in this country."[63]

Vague suggestions were made to involve grantee institutions in vetting and administering proposals, and increasing the allowances for site visits to ensure proper administration of the

---

[62] Committee on Government Operations, 1:102.
[63] Committee on Government Operations, 1:102.

awarded funds. As noted in several places in the 1961 report, the NIH reposes almost complete trust in the grantee to faithfully and appropriately carry out the proposed research. The reporting requirements during the grant period were minimal:

> To renew a grant each year during the period of support commitment, the investigator is required to submit only a summary progress report of a single page or less together with the standard one-page application form. When the project has been completed, a final progress report, *or the reprint of published material instead*, is expected 3 months or so after termination of the grant" [emphasis added].[64]

It is worth noting that reprints of the publications would be accepted in lieu of a written summary of the results, though it seems it was not required. Given the fairly nominal content requirements for the final progress reports, this seems to indicate a desire to not put an additional burden on the investigator, rather than an assertion of any equivalence between published articles and research results.

The committee seems to have been mainly focused on issues related to the administration of the funds. One of the questions asked was how to verify that researchers were not "double-dipping" and applying for grants for the same work from multiple sources, or that efforts were being duplicated in other ways. Other specific areas of concern were the purchase of equipment and the amount allowable for indirect costs. One of the major concerns was that there seemed to be no review of financial requirements during the scientific review of applications; again, the NIH trusted the individual researchers to request appropriate levels of support for their work. The Fountain Committee report concludes with 13 recommendations, ranging from ways to improve the effectiveness of the project review system to establishing guidelines on researcher salaries and travel expenses.

[64] Committee on Government Operations, 1:37.

One of the areas singled out for additional exploration in Appendix 8 was the coordination of research activities across different federal agencies. This is one of the few places in the whole report that specifically mentions tracking the publications resulting from federally supported research, but as an administrative aid:

> Here should also be mentioned the published research flowing from the support of each agency. An agency may publish lists of titles of papers or even abstracts of findings. These are valuable to the administrative staffs of other agencies, and specifically to the staff of the NIH, in better comprehending the programs of the agency putting out such material. The NIH also maintains a file of reprints of scientific papers resulting from its grants, and is prepared to furnish any Federal agency a list of all such publications having to do with any topic.[65]

This mostly speaks to the above-mentioned concerns with duplication and co-ordination of efforts across multiple agencies, rather than tracking research output.

Shannon acknowledged the effect of the increased support for research on publications in his statement before the Subcommittee on Reorganization and International Organizations of the Senate Committee on Government Operations on August 11, 1960 (contained in the second report of the Fountain Committee as an appendix). He pleads for support for indexing systems that would improve communication between scientists:

> With the rapid expansion of the national—and indeed, the international—volume of research and the proliferation of resulting publications, the task of keeping abreast of developments has become increasingly cumbersome and time-consuming. There is an urgent need for better information storage, retrieval and communication systems to shorten the time lag between a laboratory finding and its communication to other scientists, to reduce the number of sources that need to be consulted, and to make the stored data more readily accessible.[66]

---

[65] Committee on Government Operations, 1:98.
[66] Committee on Government Operations, 1:104.

This type of indexing was very much in line with what Garfield had envisioned and proposed in 1955. However, the NIH and the NSF had not been receptive to his proposals, but finally extended support in the year that Shannon made this statement.

In hearings to follow up on the Committee's recommendations a year later, Shannon reframed the overall issue as not just an audit of the fiscal management of the taxpayer's dollar, but as a larger exploration of "…the fundamental tenets of [the NIH's] operation, the nature of the statutory research grant base upon which we operate, and the assurances which other committees of Congress have been given as to our basic administrative approach."[67] Then, in a section of his prepared statement titled, "Productivity of NIH Research Grant Support" he attempts in a very few words to outline the benefits of medical research since the end of WWII: "Open heart surgery, increased survival rates in cancer, remarkable advances in understanding and dealing with metabolic diseases, absolute cures to some conditions earlier considered incurable and often fatal… The probability of advances in the next decade that will far overshadow those of the past decade is almost a certainty."[68] This tallying of specific milestones is reminiscent of the list of accomplishments in earlier reports on scientific progress.

Later in the hearings, during an exchange with Delphis Goldberg (Professional Staff Member of the Intergovernmental Relations Subcommittee), Director Shannon and Dr. Ernest Allen (Associate Director for Research Grants) discussed the role and importance of study sections. Mr. Goldberg cited an example of a proposal given a high priority score despite the outside consultants' opinion that it was not a very interesting proposal, and unlikely to yield

---

67 "The Administration of Grants by the National Institutes of Health. Hearings before a Subcommittee of the Committee on Government Operations, House of Representatives, Eighty-Seventh Congress, Second Session. March 28-30, 1962." (U.S. Government Printing Office, 1962), 12.

68 "Adm. Grants by Natl. Institutes Heal. Hear. before a Subcomm. Comm. Gov. Oper. House Represent. Eighty-Seventh Congr. Second Sess. March 28-30, 1962.," 13.

much insight. He suggested that "…persons who are completely disinterested—bureaucrats, if you will—will…provide… an independent judgment."[69] Dr. Shannon strongly defended the study section system of external reviewers:

> I think that what you have to face up to in this situation is that you place your primary bets either that you can develop a study section system that has the broad competence and can provide the facts so that you can act in a highly objective way, or that you do it the other way—the way the National Science Foundation does—and place the supervision and control in a bureaucracy system… These are two different ways of running a program… I suppose each system has its merits. It seems to us that in a program having the breadth of ours, it would be quite impossible to develop the internal competence that could make these judgments with the assurance that would satisfy our obligations.[70]

This concern about the ability of Study Sections to identify high quality proposals would continue to be a concern throughout the next decade.

While the committee members and Shannon largely agreed that a certain level of objectivity or disinterestedness would be desirable in advisory council members, the position requires a level of highly specific expertise that, at the time, was not to be found outside of those actively involved in the ongoing research. Shannon hoped to remedy this in time by building the nation's capacity for medical research:

> But at this point in the development of science, we have not yet developed in this country a sufficient number of the elder statesmen in the medical science who, having been through many years of professional experience in research and having kept pace with the technological advances in the science, are now sitting aside and viewing more objectively what their somewhat younger colleagues will do. We are beginning to develop a larger number of these at the present time.[71]

---

[69] "Adm. Grants by Natl. Institutes Heal. Hear. before a Subcomm. Comm. Gov. Oper. House Represent. Eighty-Seventh Congr. Second Sess. March 28-30, 1962.," 95.
[70] "Adm. Grants by Natl. Institutes Heal. Hear. before a Subcomm. Comm. Gov. Oper. House Represent. Eighty-Seventh Congr. Second Sess. March 28-30, 1962.," 97.
[71] "Adm. Grants by Natl. Institutes Heal. Hear. before a Subcomm. Comm. Gov. Oper. House Represent. Eighty-Seventh Congr. Second Sess. March 28-30, 1962.," 101.

This attention to building a pipeline of qualified scientists goes to what Shannon saw as a major part of his mandate as director of the NIH.

Wooldridge Committee (1965)

Director Shannon was concerned that the Fountain Committee's report would be seen as less than favorable of NIH's programs as a whole. He met with Jerome Wiesner, Chair of the PSAC, in April 1963 to discuss the aims and scope of a possible Blue Ribbon panel that could carry out a different type of review of the activities of the NIH:

> Thus a review of NIH programs and activities should be viewed the occasion to inquire deeply into the principal matters of public policy which have emerged sharply into the limelight as a consequence of the current difficulties which have confronted NIH in its relationship with the Congress and in the further development of its programs.[72]

Shannon listed several questions that could be addressed by such a committee, most having to do with the relationship between the Federal Government and universities, their respective responsibilities and the importance to the nation of maintaining centers of excellence and continued support of academic researchers and training programs. One item on the list, however, stands out as a particularly thorny issue independent of universities and graduate education programs:

> The general problem presented by the lack of real comprehension on the part of the lay public of the nation and the Congress specifically of the essential character of scientific activity and the conditions essential to its flourishing development. The tendency to equate scientific activity with science spectaculars to the diminishment of the broad, steady, undramatic pursuit of understanding and new knowledge."[73]

---

[72] "Notes for Dr. Shannon's Discussion with Dr. Wiesner, April 3, 1963," 1963, 1.
[73] "Notes for Dr. Shannon's Discussion with Dr. Wiesner, April 3, 1963," 3–4.

This goes right to the heart of the matter of accountability (and Shannon's frustration)—how to explain the painstaking, meticulous work carried out by physicians and scientists over months and years to people who lack the technical background (and possibly lack the patience) to understand its significance, but nonetheless must be told what they have gotten in return for their investment? And in the case of Congress, have the ability to decrease that investment if they are not satisfied.

What Shannon is asking Wiesner to do is to elevate the issue from a subcommittee in the House of Representatives to a White House Blue Ribbon commission composed of "…scientific, university, industrial, and lay representatives," appointed by the President, to conduct a review independent of the House of Representatives.[75] A commission made up of such persons with the appropriate background and understanding of the scientific issues, and how scientific progress is made, would, he believed, produce a report that was more favorable, and make the public more supportive of continued research. Shannon's powerful ally Senator Lister Hill called President Kennedy, reportedly "to indicate his concern about the public image of NIH as a result of the Fountain Committee report… and various other criticism's [sic] of NIH… [He] asked the President if he would appoint a high level group to make another study of NIH, the findings of which would hopefully quiet these criticisms, and 'put the gloss' back on the agency."[76]

President Kennedy was very supportive of American science and especially the NIH. According to Shannon in an interview recorded some 20 years later:

> …Kennedy said he wanted Wiesner to find some device to look at the large, R&D programs in the country. Wiesner told him that he had more confidence in NIH

---

[75] "Notes for Dr. Shannon's Discussion with Dr. Wiesner, April 3, 1963," 2.
[76] E. Packard Anderson, "Letter to James Kelly, May 17 1963," 1963. Letter from Anderson to Kelly.

> than any other program and what he would like to do is use NIH as a guinea pig of how you study a broad complex program, and this is the basis of the [Wooldridge] Committee and the [Wooldridge] Report unfortunately Kennedy died before that could be implemented… which mean that when we go into the mid 60s… we not only had congressional support, support in the presidents office, support in the department, but also formal support in the scientific community. And I think the general feeling of everybody was, the need was tremendous, the operation was superb, let's make the budget to satisfy the need rather than to satisfy any rule of annual increases or decreases. And let's let the program determination be at the election of NIH rather than in response to pressure groups.[77]

Clearly Shannon was very confident in the support of Congress as well as the White House, regardless of the findings of the Fountain Committee.

President Kennedy agreed with Hill, Wiesner and Shannon, and announced the appointment of a panel in the summer of 1963. It seems that directors of the NIH had some input in the composition of the committee, as well as defining what the committee's remit would be, based on the notes of Shannon's meetings with Wiesner. Charles Kidd, the Associate Director for Training, communicated to Dr. Shannon that he felt:

> …the question of the scientific worth of research support by NIH could be disposed of quite quickly and easily by the competent group… It seems to me that where we need help, guidance, and support is in the administrative area—defining administration broadly… It would be helpful to have the group express confidence in the advisory system, including study sections.[78]

Joseph Murtaugh, Chief of the Office of Program Planning, felt:

> …we should seek through this group to consolidate the major gains and advances that have been made through NIH programs and secure an affirmation of the underlying concepts and principles which we feel to be important… In the setting

[77] "Archives and Modern Manuscripts Oral Histories: James A. Shannon : An Oral History / Interviewed by Thomas J. Kennedy, 1984 Jan. 11," 125–26, accessed November 3, 2019, https://oculus.nlm.nih.gov/cgi/t/text/text-idx?c=oralhist;cc=oralhist;g=ammpoh;xc=1;xg=1;q1=shannon;idno=9802979;view=toc;frm=frameset.

[78] Charles V. Kidd, "Notes on Luncheon Discussion next Tuesday with MacLeod, et Al," 1963.

> of this confirmation I think it would be possible to obtain discussion of the broader unresolved issues surrounding the support of science through the use of Federal funds.[79]

Murtaugh's list of specific issues included:

> A critical assessment of the processes of review and selection of projects for support and the maintenance of a high level of scientific merit in supported activities… Such a review ought to consider the adequacy of the study section process and the validity of the advisory council function… [and] …[t]he broad issues surrounding the essential conditions of accountability in the use of public funds.[80]

High on everyone's list of desired outcomes from the White House panel was an expression of support and confidence in the selection process, and investigative freedom for the grantee scientists.

It was decided early on that the chair of this committee should come from neither government nor a university, but should be from an industry not directly connected to the NIH or the medical field at all. Shannon discussed some possibilities with James Hartgering of the Office of the Special Assistant for Science and Technology of the White House including Vice Presidents of Union Carbide, General Electric, General Motors, and American Airlines—all of whom had backgrounds in science or engineering.[81] In the end though, they selected Dean E. Wooldridge, a physicist and co-founder of Thompson Ramo Wooldridge (TRW, Inc.), a firm involved in the ballistic missile program.

On November 20, 1963, Shannon communicated to all Institute Directors, Division Chiefs, Scientific Directors, Clinical Directors, Executive and Administrative Officers that the OST had "commissioned a comprehensive examination of the past, present and prospective

---

[79] Joseph S. Murtaugh, "Discussions with Dr. McCloud Re Review of NIH," 1963.
[80] Murtaugh.
[81] "Notes on Possible Chair for Wiesner Committee, Aug. 20 1963," 1963.

programs" of the NIH, and that members of the committee would be drawn mainly from engineering and physical sciences "to bring the highest degree of objectivity to bear on the issues."[82] Shannon commented, "The NIH should welcome an objective and searching examination of its program. Such an assessment can only increase the effectiveness of our efforts to improve the health of the Nation through research," and that he considered this study to be an "extremely high priority.[83]

The assassination of President Kennedy delayed the first meeting of the Wooldridge Committee until early 1964; The final report was published in February of 1965.[84] The final survey group was composed of 11 separate panels that focused on a specific area: Administration, Anatomy, Behavioral Sciences, Biochemistry, Biophysics, Microbiology, Pathology, Pharmacology, Physical Sciences, Physiology, and Review Proceedures. In the Preface to this report, it is clear that they adhered fairly closely to the aims and scope described in Shannon's notes for his meeting with Wiesner:

> We interpreted our assignment broadly—to study how NIH spends its approximately billion dollar budget, to judge whether the American people are getting their money's worth from the expenditure, and to recommend any changes in organization or procedure that would in our opinion increase the effectiveness of the program.[85]

However, they were at least as concerned with matters of scientific merit, and the overall goals, as they were with the administration or operations component of NIH programs. The panel insisted, "In planning our own program, we placed strong emphasis on an appraisal of the quality

---

[82] James A Shannon, "Wooldridge Committee Survey of the National Institutes of Health," 1963, 1.
[83] Shannon, 1.
[84] NIH Study Committee, "Biomedical Science and Its Administration; a Study of the National Institutes of Health" (Washington, D.C., 1965).
[85] NIH Study Committee, xv.

of the scientific work… we would in addition have to attempt to judge the appropriateness of the work, in terms of the present and projected health needs of the nation."[86]

In March of 1964, while the Wooldridge Committee was still investigating, Charles Miller, Chief of the Management Policy Branch of NIH provided an overview of perceived problem areas in the NIH administration that was largely structured as a response to the concerns raised by the Fountain Committee.[87] While Miller's report focused primarily on administrative issues, it also contained a succinct historical overview of how the rapid growth of the grants system at the NIH had indirectly created some serious issues. Chief among these issues was the lack of personal contact both between grantees and granting agency, and between NIH researchers and Congress (in its role as the source of the funding):

> In the early days of the NIH grant program… it was possible for a relatively small number of NIH scientist administrators in the institutes to review each grant application… to become and to remain familiar with its contents, and to maintain direct personal contact with the investigators… Thus were established firmly in the administration of NIH's grant programs two key elements (1) direct NIH scientist to grantee institution scientist communication and decision making on grant matters' and (2) little, if any, distinction between those matters which required scientific judgment and those which were basically administrative in nature.[88]

The numbers of grantees and projects was already straining this cozy arrangement by the time Miller was writing. In the years since the NIH's founding, scientific and professional groups had demonstrated to Congress that additional funds could be used productively, and now the pressure from the public to combat cancer and other diseases led to what Miller refers to as a period of "laissez-faire in grants administration."[89] This period lasted roughly until the funding level began

[86] NIH Study Committee, xv.
[87] Charles Miller, "Administrative Problem Areas," 1964.
[88] Miller, 1.
[89] Miller, 2.

to approach the level of one billion dollars, when Congress became increasingly concerned with how the money was being spent. The increased amount of funding and number of funded projects also exceeded the level of individual scientist administrators to effectively cope with managing them; a situation that was echoed in the individual grantee institutions. The success of the project grants program also led to the creation of the program project grants, clinical grants, and new research centers. Miller cited these new programs as the major concerns of the Fountain Committee. The final section of his summary, titled "Running a Billion Dollar Enterprise from the Third Echelon," circled back to questions of organization and decision making on the budget side. Unlike other federal research agencies (NASA, NSF, AEC, etc.), which had a flatter reporting structure and were able to go directly to the Bureau of the Budget, the White House, or the Office of Science and Technology, the NIH had to go first through the Office of the Surgeon General, then through the Office of the Secretary of Health, Education, and Welfare. The net result of this was that "…those familiar with the details of these issues are not present to present them to the decision making source and to respond to questions from that source."[90] An example of the reports submitted to Congress for the budget hearings for fiscal year (FY) 1968 includes a collection of research highlights organized by subject area. The very brief Introduction to the report indicates that this was:

> …intended to meet requests and administrative needs—contains representative items selected in light of both scientific importance and public interest. In no sense is it a comprehensive review of NIH activities, nor even the most significant results… The items in this report are based on published work. The scientists' articles appear in a wide range of journals, mainly non-Federal, which are available in medical libraries.[91]

---

[90] Miller, 17.
[91] "Research Highlights National Institutes of Health 1966: Items of Interest on Research Studies Conducted and Supported by the Institutes and Divisions of NIH, as Presented to the Congress of the United States" (Washington, D.C., 1967), v.

It is worth noting that the highlights presented here were described as not comprehensive, or even "the most significant." Instead, it seems that the highlights were selected on the basis of public interest (and possibly, the ability of the lay-congressperson to grasp the importance). It is also worth noting that the research highlights were collected not from an internal department or from a federal record of grant progress, but from published articles.

The Wooldridge Committee's Principal Findings as contained in a final report in 1965 closely adhered to what Director Shannon hoped to gain from this survey. "The first and probably the most important general conclusion of the study is that the activities of the National Institutes of Health are essentially sound and that its budget of approximately one billion dollars a year is, on the whole, being spent wisely and well in the public interest."[92] The second principal finding was that changes would be required in the organization and procedures. Specifically, the committee recommended decreasing the autonomy of individual institute directors and the formation of a new Policy and Planning Council to advise the director.[93] These findings closely echoed Shannon's own impressions as described to Wiesner, and both Shannon and Hill must have felt very gratified to have the support of the White House in this area.

The Wooldridge Committee's report contained a somewhat novel re-interpretation of the NIH's mission, stating confidently that:

> Its broad mission of improvement of virtually all aspects of the Nation's health requires NIH to concentrate most of its effort on basic research… In general terms, the public funds that support NIH activities are intended to 'buy' for the American people a commensurate degree of relief from suffering and improvement of health. To achieve this goal, NIH devotes its principal effort to a

[92] NIH Study Committee, "Biomedical Science and Its Administration; a Study of the National Institutes of Health," 1.
[93] Now known as the Advisory Committee to the Director.

> broad program of investigation of life processes, rather than to search for a direct cure or prevention of specific diseases.[94]

This was an important point to establish, since, as they observed, it was unlikely to be apparent to the lay public that this type of basic research was essential, and that an improved understanding of basic life processes was likely to be much more productive in the long run. However, specific legislation and even the names of the various institutes that made up the NIH seemed to suggest a focus on disease-directed research (National Cancer Institute, National Institute of Allergy and Infectious Diseases, National Institute of Diabetes and Digestive and Kidney Diseases, etc.). The committee was enthusiastic in its praise of the quality of the work being done both intramurally and extramurally, and saw no evidence of a decrease in quality despite the tenfold increase in research support over the preceding eight years.

In terms of the benefits of NIH's programs, the committee deviated from citing any improvements in the health of the Nation's citizens, and instead focused on the effect NIH support had on the research capabilities of the country:

> The NIH activities of recent years have greatly improved the quality and quantity of both research and teaching in our biomedical institutions. In the process, the Federal Government has established an important precedent of financial assistance; withdrawal or substantial curtailment of its support would be disastrous to this important segment of science and education… Current biological discoveries constitute the useful immediate products of the NIH-supported research projects; future discoveries will largely depend on the effectiveness of today' scientist/teachers, frequently aided by NIH training grants, in instructing and inspiring the next generation of health research scientists. Therefore, if we are to appraise the major effects to date of the NIH program, we must seek an answer to the question: "What contribution has NIH made to the capability of the universities and medical schools to perform productive health research now, and to maintain a flow of trained scientists to perform the health research of the future?"[95]

---

[94] NIH Study Committee, "Biomedical Science and Its Administration; a Study of the National Institutes of Health," 2.
[95] NIH Study Committee, 4–5.

This focus on capacity and talent reflected Shannon's goals with respect to developing the US as a leader in scientific research. This also represented the different focus and motivation between the Fountain and Wooldridge Committees. The Fountain Committee seemed to have been determined to find room for improvement in the NIH's programs and operation, whereas the Wooldridge Committee was convened fairly explicitly to highlight the NIH's success, to counteract any negative impact the Fountain Committee's findings might have had on the image of the NIH in the public eye.

Each panel produced a separate, more detailed report on their area. The Administration Panel Report is the most important in the context of evaluating scientific output, as it directly addresses issues of reporting and accountability that were relevant to all institutes. The panel notes that researchers are "…required, formally, to submit a brief annual report on his technical progress, to accompany the annual formal request for interim renewal of the grant, but this rule appears not to be universally and promptly observed."[96] The panel felt that the approach of "…pick out the best investigators with the best ideas, assure them of uninterrupted support for a few years, and then leave them rather strictly alone except when they ask for help,"[97] had been highly successful thus far, but they expressed concerns about its continued effectiveness as the programs continued to grow.

A major "pain point" the panel identified is the "Time and Effort Reporting" requirement:

> Recently, recognizing that the pay of an investigator, as a major item of project expense, deserves close administrative attention, NIH issued regulations requiring that a record of the time or effort spent by investigators on supported projects be

[96] NIH Study Committee, 105.
[97] NIH Study Committee, 105.

> maintained by the institutions and be subject to NIH audit. The complaints aroused by these regulations can readily be imagined by anyone acquainted with the pattern of life in an academic community.[98]

In this context, both *time* and *effort* are measures of *input* on the part of the investigator; there is only one brief mention of what could possibly be considered *output*:

> The productivity and accomplishment of the investigator, as judged by his peers in the scientific community, are unquestionably the ultimate criteria for merit of the project. These are the points on which initial decision to support the project is primarily based, and on which decision to support later projects proposed by the same investigator will be primarily based.[99]

The panel does not provide specifics on how the judgment of one's peers is to be determined—again, this is a need that Garfield was seeking to fill. This is worth noting, as citations will later be described as a metric to quantify the approval of peers in terms of scientific acceptance.

Fountain Committee (1967)

The Fountain Committee conducted a follow up investigation, the results of which were published in October of 1967. This time, the purpose was "…to examine and evaluate the performance of the Public Health Service—and especially of its principal research bureau, the National Institutes of Health—in administering grant programs for the support of health research since the Committee's previous reports…"[100] Overall, the report produced in 1967 was much more critical of the NIH than the previous report from 1961 and the update from 1962, especially in areas of administration and cost-sharing arrangements. They felt the support provided via

---

98 NIH Study Committee, 114.
99 NIH Study Committee, 116.
100 Committee on Government Operations, "The Administration of Research Grants in the Public Health Service, Ninth Report" (Washington, D.C., 1967), vii.

institutional grants (especially the one to the Sloan-Kettering Institute for Cancer Research) to be against the general principles of the agency since it bypassed scientific review of specific programs, and removed a large sum of money from the pool of funds available on a competitive basis. The recently established Health Sciences Advancement Award was derided as "irresponsible, unscientific, and contrary to the best interests of the academic community and the Government… the purpose of this program is still unclear."[101] They found "NIH's administration of the General Research Support program surprisingly casual, with policies and procedures inadequately developed for the equitable and uniform treatment of grantees and with management indifferent to the waste of program funds."[102] They also expressed concern that grants seemed to be concentrated in a relatively small number of schools, with the overall effect of exacerbating the gap between "rich" and "poor" schools, and helping already very good schools become excellent, at the expense of strengthening weaker institutions.

Regarding the project grants and the overall procedure for selecting which ones would be funded, the Committee was concerned that "[r]esearch projects supported by grants from NIH… have, according to the agency's own ratings declined in quality over the years."[103] They were also concerned about potential conflicts of interest and the composition of the study sections in general:

> The Committee is concerned by the tendency in the Public Health Service to use the services of a small group of individuals for long periods on the advisory councils and other major advisory bodies. When some of the same individuals who have served on advisory councils for many years receive substantial NIH grants, and also testify before the Congress in support of the agency's appropriations, the appearance of favoritism is unavoidable.[104]

[101] Committee on Government Operations, 3.
[102] Committee on Government Operations, 3.
[103] Committee on Government Operations, 2.
[104] Committee on Government Operations, 2.

The appearance of favoritism may have been unavoidable, but it would have been difficult to avoid having recipients of NIH grants serving as councilors. The Committee's recommendation to remedy this situation was to rotate members of the advisory councils more frequently, and to encourage younger scientists to serve in this capacity. This approach differs from that suggested following the first Fountain Committee report (to have "elder statesmen" serve as advisors), which had agreed with Shannon's thinking regarding having semi-permanent, somewhat disinterested advisors making funding recommendations to seeing advisory councils as a way to help young researchers build their careers.

The Committee presented a list of 17 recommendations for the PHS detailing what they viewed as weaknesses in the administration, most of which dealt with institutional and developmental grants. With respect to advisory committees, the Fountain Committee recommended term limits, and "consideration be given… to obtaining a balanced representation of geographic regions and educational institutions. To the extent possible, consultants should be drawn from among qualified scientists who are not themselves recipients of PHS grants."[105] With respect to the quality of funded projects, the recommendation was vague:

> The committee recommends that the Surgeon General establish a high standard of quality as the basic qualification for research project support, and that he develop adequate procedures for the uniform maintenance of that high standard by NIH… The confinement of research grants to projects in the range of excellent to good should not be breached except in special circumstances where the reasons for supporting a lower quality project are fully documented in a written record.[106]

This is a much stronger questioning of the validity of the priority rankings assigned by the study sections than in the previous reports. The report also states, "NIH and the Public Health Service have never clearly defined the qualitative level expected of applicants seeking support for their

[105] Committee on Government Operations, 7.
[106] Committee on Government Operations, 5.

research. It is unclear whether the objective is to support only high quality research or to extend support to all 'competent' investigators."[107] The committee notes that the number of funded projects with scores in the highest priority class has declined since 1956, and has not improved despite the urgings in the 1961 and 1962 reports, but it is not clear what significance this has, if any. They do, however, maintain that this shift in priority scores reflects a decline in quality, which they take pains to point out is in contradiction to the findings of the Wooldridge Committee. They charge that the findings of the individual panels of that committee do not even support the claim that there has been no decline in quality, and quote Dr. Harold Orlans of the Brookings Institution's criticism of the Wooldridge Committee, "All told, a more accurate summary might be that NIH sponsored research is generally of 'good' quality or… 'no better and no worse' than other work in the same field."[108] The committee expresses anxiety regarding this point. "What is the merit and the national purpose of supporting research that fails to meet a high qualitative level? If pedestrian research contributes importantly to the advancement of science, that fact has not been brought to our attention."[109] They go so far as to suggest that the excessive research expenditures might be diverting scientific manpower away from teaching and the practice of medicine to the glamor of the research bench. They also expressed concern that the volume of scientists reliant on the federal government for support might mean that they can no longer be relied upon for an objective opinion of the NIH's programs. "Perhaps the greatest obstacle to open and objective examination of the undesirable effects of Federal research grants is the fact that PHS now supports most of the biomedical research in the United States. Consequently, investigators are reluctant to 'bite the hand that feeds them'—especially in

---

[107] Committee on Government Operations, 42.
[108] Committee on Government Operations, 44.
[109] Committee on Government Operations, 44.

public."[110] Thus NIH could be seen as something of a victim of its own success—the very fact of this prominence creates a lack of appropriate reviewers.

In regards to its claim that overfunding of research was drawing professors away from teaching, the Committee seems to be relying heavily on a rather alarmist essay by Dr. W.C. Davison, dean emeritus of Duke University School of Medicine in *The Pharos* bemoaning the way American medical schools were beginning to resemble the research institutes of pre-war Germany, and research being done solely for the purpose of obtaining funding. This argument is bolstered by citing a report from the National Academy of Sciences (NAS) that determined "the fruitful combination of research and graduate education" might be hindered by universities offering to release prestigious academics from teaching duties in order to prevent them leaving for pure research jobs—the NAS does not differentiate between government, research institutes or industry, and one presumes that universities would replace faculty members who have left with others willing to take on their duties, including teaching (though that is not stated here).[111] However, the Fountain Committee itself did not investigate the impact of federal research funding on medical school faculties or instruction. They posit that:

> …[s]upport for biomedical research of less than high quality has been rationalized on the grounds that the spillover from this research "enriches the academic environment" and thereby benefits the educational programs of recipient institutions. This trickle-down theory, unfortunately, overlooks the diversionary effects of such support as well as other inefficiencies.[112]

In other words, the Fountain Committee claims that the NIH has justified funding mediocre research as a way to provide financial support to academic institutions for non-research related purposes, and that this could be causing harm in the long run. There were two ways to address

[110] Committee on Government Operations, 45.
[111] Committee on Government Operations, 47.
[112] Committee on Government Operations, 47.

this concern: either show that the grant-supported projects are being selected appropriately, or show the "trickle down" of support is having a positive effect. For the NIH, the former is more in line with their mission, and that is how they framed their responses to this report.

Chapter 4 Program and Policy Impacts

The immediate impact of the Fountain Committee's first report in 1961 was to put the NIH's communications programs into sharper focus. This emphasis on the role of scientist-to-scientist communication helped to define the NIH's role in the dissemination of research results to other scientists, and also to health practitioners and the general public. It also had a more material impact in encouraging the NIH to financially support improved indexing methods, which is of key importance in the field of bibliometrics and scientometrics. This investment in building more comprehensive databases permitted a second wave of impact, when the later reports of the Fountain Committee (and other agencies) led the NIH to use bibliometric and citation analysis to evaluate existing programs and attempt to predict the impact of future policy decisions.

Impact on Communications Programs

In 1961 the House Appropriations Committee's recommendations to the Department of Health, Education, and Welfare included a specific mention of the lack of attention being paid to scientific communications at the NIH. Out of four recommendations in total, this was the only one that did not address a specific and broad-reaching public health issue. Under Secretary Ivan Nestingen summarized the recommendation in a memo to Surgeon General Luther Terry: "The Committee comments at some length on the importance of prompt dissemination of research findings to physicians. The Committee then states it was 'disappointed in what appears to be a very restricted point of view on the part of the National Institutes of Health in respect to this very important field.'"[114] Shannon responded to Terry that the communications issue was much broader than indicated in the Committee's report, encompassing communication between scientists, communication from scientists to clinical practitioners, and communication of health information to the general public. Shannon notes that the Committee report appeared to place the full burden of communicating advances in research to physicians on the NIH, which he felt was inappropriate. He felt that the responsibility should be shared between the different arms of the entire PHS. He went on to describe the NIH's activities in each of the three previously identified categories of scientific communication. Of the three, he considered communication between scientists to be the area of the most immediate concern to the NIH. He listed the ways in which the NIH supported communication directly between scientists (through conferences, symposia, and meetings), through publications (both original and in translation), and the then-new digital arena ("…the adaptation of electronic computing and data processing equipment to the storage and retrieval of scientific information").[115] He also named other agencies that involved in the same endeavor: the National Library of Medicine (NLM), NSF, the Office of Technical Services, the Library of Congress, and "the full universe of professional and scientific publications."[116] He stated that there was some difficulty in distinguishing between the role of the NIH and the NLM, and identified delineating those roles as the main problem for these two agencies at the time. Shannon placed the main responsibility for communication between scientists and health practitioners, which was the main concern of the Committee with the Public Health Service's

---

[114] James A Shannon, "Communication of the Results of Research," 1961, 1.
[115] Shannon, 2.
[116] Shannon, 3.

Bureau of Community Health, with the NLM and NIH playing supporting roles in maintaining "bibliographic and serial materials," and determining which findings were relevant to practitioners, respectively. Communication of health information to the general public was entirely within the purview of the Bureau of Community Health.[117]

In a memorandum to Joseph S. Murtaugh, Chief of the Office of Program Planning, George.L. Payne, Chief of the Special Projects Branch, summarized the House committee's requests and "scattered references to programs which will 'speed up the clinical application of laboratory findings'" from reports from both houses.[118] He believed that "[p]reoccupation with the application of research results feeds and fosters the growing interest in the problems of communication."[119] He suggested that these concerns indicated a need for a "full-dress review of the subject,"[120] and NIH's role in the solution.

The report proposed by Payne was delivered to the Subcommittee on Labor, Health, Education, and Welfare Appropriations of the House Appropriation Committee on March 9, 1962, and was adapted for publication in the Journal of Medical Education in November of that year.[121] It begins with an explanation of the origin of the "Communication Problem," which the authors attributed to two main sets of circumstances: the major changes that had occurred in the previous 25 years, and the gap between the infinite amount of knowledge, and the limits of human comprehension and retention. While the second set of circumstances is beyond the

---

117 Shannon, "Communication of the Results of Research."
118 G.L. Payne, "Notes on Special Reports Requested by Congressional Committees," 1961, 1.
119 Payne, 1.
120 Payne, 2.
121 National Institutes of Health (U.S.) Office of Program Planning, "Communication in the Bio-Medical Sciences: A Report by the National Institutes of Health to the Subcommittee on Labor, Health, Education and Welfare Appropriations of the House Appropriation Committee" (Bethesda, 1962); J.S. Murtaugh and G.L. Payne, "Communication in the Biomedical Sciences," *Journal of Medical Education* 37, no. 11 (1962): 1169–82.

influence of federal science policy, and the first is not something that could be reversed (nor would anyone want to), it is helpful to keep these concerns in mind in looking at issues related to policy related to scientific communication and support. Murtaugh and Payne divide communication into the same three areas that Shannon had identified: scientist-to-scientist, scientist-to-health practitioner, and scientist-to-public. Again, they echo Shannon's contention that the "scientist-to-scientist" communication is the main concern of the NIH, though they stress that the solutions to the problems in that area will likely be common to all three areas.[122]

Murtaugh and Payne subdivided the communication of research results within the scientific community into four categories: Personal Interchanges, Publication of Findings, the Library Process, and Training and Educational Process. "Personal Interchanges", the first is described as "…the fundamental and undoubtedly the most powerful, immediate and effective means of conveying information in the scientific and professional world."[123] The NIH actively supported both formal and informal meetings to facilitate this type of exchange, as well as the Study Sections, Review Committees and Special Advisory Groups that primarily evaluated grant proposals but also "…initiate and sponsor conferences and symposia to explore a new development, discuss the state of knowledge and needed areas of research effort in a particular field or to examine in detail a specific problem."[124] Individual researchers were supported in attending national and international conferences by providing travel funds in the supported grants. The NIH also provided travel grants to individual researchers to attend both national and international conferences.

---

122 Murtaugh and Payne, "Communication in the Biomedical Sciences," 1170.
123 National Institutes of Health (U.S.) Office of Program Planning, "Communication in the Bio-Medical Sciences: A Report by the National Institutes of Health to the Subcommittee on Labor, Health, Education and Welfare Appropriations of the House Appropriation Committee," 6.
124 National Institutes of Health (U.S.) Office of Program Planning, 8.

Regardless of the immediacy and power of interpersonal communication of research results, publication of findings was described as an essential part of the research process:

> Research… is not complete until the investigator's findings are made known to those who can apply them or are added to the reservoir of knowledge upon which the further progress of science depends… The publication of his findings in a scientific or professional journal is a definitive, though interim objective towards which a scientist works… It also is an event from which the scientists derive some measure of personal satisfaction that is both personally meaningful to him and on which in no small measure his personal reputation and the progress of his career depends.[125]

This is in fairly close agreement with Price's interpretation of the reasons scientists publish: the informational ("findings… are added to the reservoir of knowledge") and the social ("[publishing] is an event… on which in no small measure his personal reputation and the progress of his career depends"). Both functions support the NIH's mission to contribute to human knowledge and improve health, and to enhance the U.S.'s reputation as a leader in science and medicine.

While interpersonal communication between scientists had gotten easier since the end of WWII, the increases in manpower and resources devoted to research was creating challenges in the publication of research findings. The post-war expansion of the scientific community led to a corresponding expansion in the volume of scientific publications, in both numbers of articles published and journal titles. The increase in the number of journals was due not just to the need for more space to publish more articles, but also to increasing specialization and interdisciplinary research. The report mentions that this form of publishing is expensive, and then as now, not supported by advertising like most general interest periodicals, and describes some of the ways in which the NIH helped with the costs. Murtaugh and Payne cited the availability of grant

---

[125] National Institutes of Health (U.S.) Office of Program Planning, 10–11.

money to pay for page charges as one way in which the NIH was providing material support for the publication of research findings, but they also suggested that new forms of direct support would be needed, as well as new publication devices (such as review journals, bibliographies, indexes, etc.). At the time this report was written, the NIH published an annual list of the approximately 1600 scientific papers per year produced by its staff (the Scientific Directory-Bibliography), which served both as a concise survey of the research interests and outputs of the various labs as well as staff directory.[126] In FY 1961, $743,000 was granted externally for the production and distribution of primary research journals, but far more ($3,527,000) was spent on bibliographic and reference services.[127]

The third area, libraries, is intertwined with the publications area. Several of the issues facing libraries were directly related to the expansion of the publication universe, such as the needs for more space to house the larger collections of journals, improved indexing, and trained staff. Although the NLM was the primary agency to be assisting with these issues, the NIH was involved too. In addition to providing funds for buildings and staff NIH supported was the development of the MEDLARS-- Medical Literature Analysis and Retrieval System project. MEDLARS was developed under contract to the General Electric Company, and would use computers to manage bibliographic indexing. This would provide "...increased high-speed capacity for the production of the Index Medicus… recurring bibliographic listings of references selected in accordance with the predetermined requirements of particular fields… search and retrieval capacity to answer, on demand, queries from individual research installations concerning newly published information bearing on immediate problems."[128] The Advisory

---

[126] National Institutes of Health (U.S.) Office of Program Planning, 13.
[127] National Institutes of Health (U.S.) Office of Program Planning, 14.
[128] Frank B. Rogers, "MEDLARS Meeting," 1962, 1.

Committee on Computers in Medical Research was also tasked with "exploring the possibilities of electronic information retrieval systems for the specific needs of research in the biomedical sciences."[129]

The final section of Murtaugh and Payne's report, Planning and Development Data, looks at larger issues of the role of data in policy and the conduct of research. This section seems to be asking the questions that citation analysis would eventually be deployed to answer:

> Research administrators… have a need for data that will display in an easily assimilable manner the whole spectrum of work in progress… and that will give some measure of the relative effort going into each of the various research areas. These data are usually most conveniently presented in terms of dollars expended, though this is often not the most meaningful yardstick… More precise information is usually necessary—and much more difficult to obtain—to identify gap areas, to spot over-concentration of effort, and to determine the relationship of existing or proposed projects to the total institutional or national research program in a broad discipline or disease category.[130]

The "most conveniently presented" data referred to here is a measure of input (funding). As Murtaugh and Payne observe, it is not necessarily the most accurate measure of actual effort. Much of the information they feel is necessary would be available via citation analysis.

This need for data about research programs is related to the NIH's "…special responsibilities for deploying its tax-derived resources for the most effective attainment of specific primary national objectives… and for full accountability for the expenditure of public funds."[131] Murtaugh and Payne asserted that the NIH faced a complex challenge in meeting this responsibility due to the state of the science being essentially basic research, and therefore less

---

[129] National Institutes of Health (U.S.) Office of Program Planning, "Communication in the Bio-Medical Sciences: A Report by the National Institutes of Health to the Subcommittee on Labor, Health, Education and Welfare Appropriations of the House Appropriation Committee," 20.
[130] National Institutes of Health (U.S.) Office of Program Planning, 28.
[131] National Institutes of Health (U.S.) Office of Program Planning, 29.

amenable to timetables and reasonable assurances of eventual success (the authors cited the atomic bomb and space capsule as examples).[132] The report also emphasized that apparent duplication of effort is not necessarily something to be avoided at this stage, as it is likely that different researchers would bring different approaches and insights. They expressed the view that:

> …the Federal government can not and should not attempt to dominate the nation's biomedical research effort but should make its support available on the basis of the scientific excellence and promise of the projects submitted to it by qualified investigators from an alert and informed community of scientists, and that the direction of research effort should, by and large, be determined by the collective judgment of the scientific community itself rather than the arbitrary decisions of scientist-administrators in the Federal service.[133]

This reasserted the NIH's commitment to letting scientists direct their own research with minimal interference.

Another component of the NIH's communication mandate was communication with the public. While still maintaining that communicating information relevant to maintaining and improving health was the responsibility of other divisions of the PHS, and not the NIH, "…it is clearly incumbent on Federal administrators of this research effort to acknowledge the public's confidence and support with candid and lucid reports of research achievements and of future program needs."[134] The NIH fulfilled this obligation through direct communication to the public and by cooperating closely with other parts of the PHS that had more direct contact with the public.

The report recommended a few "Primary Areas of Attack" to improve the communications efforts of the NIH. First among these aimed at communication within the

---

132 National Institutes of Health (U.S.) Office of Program Planning, 30.
133 National Institutes of Health (U.S.) Office of Program Planning, 30–31.
134 National Institutes of Health (U.S.) Office of Program Planning, 33.

scientific community, by facilitating "the prompt publication of research findings in the scientific and professional literature" (including the acceptance of page charges by all granting agencies), supplemented by secondary publications and bibliographic devices."[135] The second area called for enhancements in the nation's medical libraries, to ensure maximum accessibility to the literature, including "the most effective means for its bibliographic management." These means were to include "…applying the potentialities of the new electronic technologies to the handling of bibliographic materials…"[136] The third prong of the attack addressed the need for continuing medical education. Again, the responsibility for creating and offering such programs is again "assigned" to the medical schools and professional organization.[137]

By the time this was written, Eugene Garfield had already received funding from both the NIH and the NSF to build a citation analysis map. The Advisory Committee to the Study of Problems in Biomedical Communication of the NAS and the NRC also felt that use of new computing technology could contribute significantly to this area: "Research on means of communication can make a very significant contribution to the national biomedical effort. Generous support is warranted, in particular, for studies that seek to exploit modern information technology and to adapt it to the habits and needs of biomedical investigation."[138]

Eugene Garfield had long proposed constructing an index of citations to scientific literature. He saw this not as a measure of productivity, but as a tool for researchers and administrators to be able to identify the most productive areas of collaboration, and for researchers to be able to gauge the trustworthiness of the data in a given paper. As Garfield had

---

135 National Institutes of Health (U.S.) Office of Program Planning, 33.
136 National Institutes of Health (U.S.) Office of Program Planning, 34–35.
137 National Institutes of Health (U.S.) Office of Program Planning, 36–37.
138 Bruce P. Carson, "Report on October 3-4 Meetings of NAS-NRC Advisory Committee on Scientist-to-Scientist Communication," 1963.

written in 1955: "I propose a bibliographic system for science literature that can eliminate the uncritical citation of fraudulent, incomplete, or obsolete data by making it possible for the conscientious scholar to be aware of criticisms of earlier papers."[139] Simply put, access to lists of published works that had cited a given reference would make it feasible for researchers to easily check what had been said about it, and hopefully therefore once a source had been flagged as suspicious, other researchers might be warned off. In this way, unreliable or problematic data would eventually disappear from the active literature.

What is a Citation Index? According to Garfield, it is "…an ordered list of cited articles each of which is accompanied by a list of citing articles. The citing article is identified by a source citation, the cited article by a reference citation. The index is arranged by reference citations. Any source citation may subsequently become a reference citation."[140] This differed from other indexes available at the time, which generally listed papers by either subject or author, in that it provides some context and indication of the relative value of a reference to the community, and leads the researcher to additional related papers published subsequently. Garfield termed this use of this index "cycling." "By a technique called "cycling" you can quickly find the other papers on the subject... 'Cycling' means examining the bibliographies of the papers you start with, and of the source papers obtained in order to locate additional relevant works. By looking up the latter in the citation index you find new citing sources."[141] The index was arranged by author name, although use of computers made it less important since users could search by author, or title, or keywords, or other indexed metadata. It was assumed that the

---

139 Garfield, "Citation Indexes for Science; a New Dimension in Documentation through Association of Ideas.," 108.
140 Eugene Garfield, "'Science Citation Index' --A New Dimension in Indexing," *Science (New York, N.Y.)* 144, no. 3619 (May 8, 1964): 650, https://doi.org/10.1126/science.144.3619.649.
141 Garfield, 651.

researcher has at least one paper or author to use as a "starting point" to conduct a search in this style of index. In this scenario, a researcher finding a "negative result" (i.e.—a lack of citations to the original reference paper) would know that they had found a knowledge gap that was ripe for investigation. Articles that had several citations could, in turn, lead a researcher to more recent work in that area, which would quickly bring them up to speed with the current state of the science for that specific area.

While Garfield's primary purpose in creating his citation index was to help researchers, Dr. Joshua Lederberg recognized the potential for it to help determine the NIH's impact on research during a debate at the Genetics Study Section. He recalled Garfield's 1955 article in *Science*, and reached out to him to encourage him to reapply to the NIH for funding.[142] The desire to evaluate the impact of the NIH's funding on genetics research was only one component driving the decision to fund Garfield's proposed index. According to Paul Wouters, there was also the perceived "information crisis" that had arisen as a result of the increasing amount of scientific publishing in the last decade:

> This information crisis shaped the way the central problems in the realms of science, science management and science policy were defined. Government agencies provided funds to find solutions to this information crisis and thereby created a new labour market, asking for people with both scientific and librarian skills… The crisis, made more urgent by the Sputnik crisis, finally gave citation indexing the official approval it needed to take off.[143]

Finally, in December 1960 he was given support in the form of a single project, funded jointly by the NIH and the NSF. The two agencies had different objectives: the NIH was mostly interested

---

142 Paul Wouters, "The Citation Culture" (University of Amsterdam, 1999), 39–40.
143 Wouters, 76.

in getting a genetics citation index, wherease the NSF wanted a "sound test of the value of citation indexes as a bibliographic tool."[144] According to a press release announcing the project:

> Research scientists will soon be consulting a more precise and specific literature index that links together subject material that would never be collated by usual indexing systems. Concerned with new starting points for scientific literature searches, the unique concept uncovers sometime-buried associations, relating important works and authors, yet keeps the researcher abreast of the masses of current published scientific information… Better scientist-to-scientist communication is expected, for authors can see at a glance what literature has been published since their works, in their own and related fields, that refer back to their own works.[145]

This statement of the purpose of the Citation Index describes both the way it would be used as a tool for researchers to identify research fronts and opportunities, and the expectation that it would improve the communication of research progress between scientists, which was a very specific objective for the NIH.

Policy Impacts

The details of science policy in the US through the 1960s was largely left in the hands of the scientists themselves:

> Quality control stayed in the hands of the scientists themselves: after all who else could judge the often esoteric research results? As a consequence, assessing the state of affairs in a given scientific domain was based on the predominantly qualitative judgements [sic] of the experts in that domain. Quantitative indicators were mostly restricted to budget figures and personnel estimates.[146]

[144] Wouters, 57.
[145] Wouters, 219.
[146] Wouters, 132.

What Wouters means here is that the decision about what projects and directions to pursue was determined by active researchers, based on their assessments of the merit of the work (and the merit of the investigator), rather than by systematic, objective analysis.

At the NIH, this quality control was largely carried out by the peer review of grant proposals and renewal requests. While the reviewers did assign numeric scores to proposals, the process was largely qualitative in the sense that it rested largely upon the individual reviewer's assessment of the overall quality and merit of the proposal. In the early 1970s, the NIH sought confirmation that their process to identify the highest quality proposals was effective. In order for the analysis to be as impartial as possible, an outside agency, the RAND Corporation, was contracted to perform the work.

The RAND Corporation became involved in the evaluation of the effect of federal research on academic health centers through a contract with the Bureau of Health Resources Development of the Health Resources Administration and the Office of the Assistant Secretary for Planning and Evaluation, HEW to do a study to evaluate the effectiveness of the peer review system for awarding grants. Much of the work was performed by Grace Carter (including her Ph.D. thesis). In the earliest published product of this contract (December 1974), Carter describes the first challenge in performing the analysis as identifying metrics for the output of research to use to benchmark the scores given to renewal applications by the study sections. Citation count was selected as the independent output measure:

> In this report counts of citations of articles produced under NIH grants are used as a surrogate for the usefulness of the sponsored research to other researchers. The scientific merit of the proposal to renew the grant and citation data are crude measures only of the value of the research to science. The progress of science is only an intermediate step in the attainment of the long range goal of NIH—

> improvement of health care. Thus, the basic assumption of this report is that in the long run science will be an effective instrument in the alleviation of disease. [148]

Citation counts were being used here as an indirect indicator of the value of the research to the community. This rests on the assumption that citations are generally positive in nature, from within the community, and not subject to confounding influences. A larger assumption is that research of value translates into improvements in health care. There are specific reasons for using citations: a lot of research funded by the NIH does not directly (immediately) lead to changes in how health care is delivered, but informs other research (an impact which can be mapped by looking at citations), and it can be assessed in a much shorter time frame.[149]

In October of 1976, Carter presented the second progress report on RAND's study "Analyses of NIH Policies for Support of Extramural Research" at the Small Office of the [NIH] Director staff meeting detailing their progress from April to August of 1976.[151] Most of the report is concerned with the details of the methodology behind the analyses they had conducted, how they selected the grants from each institute to include in their analyses and case studies, and which "variables" they would be using. In addition to the Administrative Report, the NIH Office of the Director was also provided with a Draft (WN-9624), "Sample Selection and Data Collection for a Study of Research Performed Under NIH Grants."[153] This draft goes into a little more detail on the rationale behind the methodology and selection process presented in the Administrative Report. The report justifies the use of citation analysis by citing the results of

---

148 Grace M. Carter, "Peer Review, Citations, and Biomedical Research Policy" (RAND Corporation, 1974), v.
149 Carter, 1.
151 Grace M. Carter, Clara S. Lai, and Carolyn L. Lee, "Sample Selection and Data Collection for a Study of Research Performed Under NIH Grants," 1976.
153 Carter, Lai, and Lee, "Sample Selection and Data Collection for a Study of Research Performed Under NIH Grants."

their previous study in which it was shown that"…citation data, particularly exceptionally large numbers of citations, can provide a crude indicator of the value or research to the scientific community. Our hypothesis here is that publication and citation data can also be used to describe other aspects of research performed under the grant."[154] The NSF reported a similar finding in an internal study in 1972:

> …an internal study of NSF awards in chemistry shows a striking correlation between the distribution of citation rates and the allocation of NSF funds. Of NSF awards in chemistry, 80 to 85 percent go to departments whose faculty averaged more than 60 citations per author during the 5-year period ending in 1972; the top four or five departments… averaged around 400 citations per author.[156]

This statistic also demonstrates the way in which top departments get the lion's share of citations, as well as grants.

We see here that citation and publication data is presented as a widely acceptable proxy for productivity in terms of scientific research. Even in such a thorough, well-funded study, looking at a relatively small sample of grants, the researchers apparently did not even consider making an independent judgment or evaluation of whether or not the investigators had achieved their stated goals, proved or disproved any hypothesis being tested in their proposal. The reliance on the publication of results shaped the selection of grants included in their analysis, as they were limited to grants that had been funded several years prior to their study.[157]

The goal of Carter's work was to determine if there was a relationship between the number of citations to articles produced by grantees seeking renewal of funding and the priority scores assigned to the proposal at the initial application, and again at renewal. The peer review

---

154 Carter, Lai, and Lee, 5.

156 Thane Gustafson, "The Controversy over Peer Review," *Source: Science, New Series* 190, no. 4219 (1975): 1063.

157 Carter, Lai, and Lee, "Sample Selection and Data Collection for a Study of Research Performed Under NIH Grants," 7.

procedures as practiced by the study sections were regarded as an effective method to identify high quality research proposals both within the scientific community and the government overall, a position which was further strengthened by the addition of a peer review requirement for NCI-supported projects in the National Cancer Amendments of 1974.[158] However, there were questions as to the objectivity of peer review, perhaps especially because the proposed research was so advanced that only other members of the small group of experts can be called upon as evaluators, and the general public (including most members of Congress and the Executive Branch) must trust the scientific community to disburse public funds fairly and appropriately. Given this, some kind of objective measure of the success and/or usefulness of NIH-supported research was highly desirable. As Carter stated in a presentation of her work at the annual meeting of the American Association for the Advancement of Science in 1978:

> Both citations and the peer review judgment are imperfect measures of the usefulness of research to the scientific community which we cannot measure directly. In presenting this data, I have assumed that citation counts are correlated with the usefulness of the research to the scientific community, and used this assumption and the data to infer that the judgments of the peer review process are also related to the usefulness of the research to the scientific community. However, it is possible to turn the argument around. If you begin with the hypothesis that the priority scores are measures of scientific merit, then this data strongly suggests that counts of citations are correlated with scientific merit.[159]

In other words, both priority score and citation count are related to research output. Carter selected the average citation rate in predicting priority scores for renewals because it was a convenient continuous variable. She found that the renewal application score was related to the

---

[158] NIH Study Committee, "Biomedical Science and Its Administration; a Study of the National Institutes of Health."
[159] Grace M. Carter, "A Citation Study of the NIH Peer Review Process:" (RAND Corporation, 1978), 20.

number of citations to the articles published in the grant period (but was unrelated to the investigator working at research-intensive school).[160]

There are of course a variety of specific ways to analyze the raw counts of citations, and indeed many different types of publications. Carter et al. purchased the citation data from Eugene Garfield's Institute of Scientific Information for publications listed in the *Research Grants Index* from FY 1967- FY 1970 for 747 research projects and 51 program project grants awarded by NIH in FY 1967.[161] This resulted in a list of 5,783 total publications (including 4,351 journal articles) and 41,339 total citations.[162] Based on the distribution of the number of citations/article, Carter determined that the top 5% "most cited" articles were results of 14.5% of the grants in the sample—and she hypothesized that these 116 grants were "exceptionally useful pieces of research," and it should "…then be possible to discover those characteristics of grants that are likely to produce exceptionally useful research results."[163] The approval rate for renewals was significantly higher for grants that were categorized as having produced at least one "most cited" article ($p = 0.02$), but that does not capture any potential differences between the scientific merit or perceived value of the original proposal and the renewal application. To attempt to answer this concern, Carter performed a regression analysis to quantify the relationship between initial proposal priority score (from 1967), production of at least one "most cited" article, and renewal score. The analysis showed that on average, renewal applications from the "most cited" group

---

[160] Carter, 21–23.
[161] Carter, "Peer Review, Citations, and Biomedical Research Policy," 20. The sampling was also restricted to the ten representative medical schools identified for the larger study, plus a random sampling of grants to other medical schools, excluding those from study sections with 9 or fewer grants available.
[162] Carter, 22.
[163] Carter, 24–25.

received scores 47 points better than would have been expected based on the initial scores from 1967.[164]

This could be interpreted to lead to a somewhat circular argument that the higher-than-predicted priority scores on these renewal applications indicate that the initial priority scores did not correlate well with value to the scientific community, or that the study sections' evaluation of the renewal applications may have been influenced by early (and recognized) success on the part of the supported researchers. As Carter puts it:

> Since the production of a highly cited article results in a significantly better renewal priority score even after one controls for the priority score received in 1967, it would follow that part of the low correlation between the scores received on a funded grant and its subsequent renewal is due to an adaptation by the second study section toward the merit of the research and away from the earlier appraisal of its potential… the existence of a strong correlation between the two research output measures [production of at least one "most-cited" article and renewal priority score] boosts the face validity of each of them.[165]

Based on this, it seems that renewal scores are a better indicator of the usefulness of the funded work, if one accepts that the citations are good indicators of the value the community assigns to the work described in the published article. Accordingly, the rest of the report uses the criteria of having produced at least one "most-cited" article as the defining characteristic of an "exceptionally useful" grant.

Of course, simply dividing grants into "most-cited" and "other" is a fairly crude measure of the scientific value of the work being reported. Carter's more detailed analysis of citations included looking at overall average numbers of citations by scientific field and the time pattern. Time patterns of citations did not vary significantly between the fields, and that citations to the most frequently cited articles occurred later, suggesting that citation counts are only reliable

---

164 Carter, 28.
165 Carter, 29.

indicators of value for research that is at least 5 to 6 years old—making this a reasonable variable to consider when creating broad management policy, but not when considering individual grant renewals (which typically occurred earlier than 5 years).[166]

In attempting to identify a reasonable measure of research output, Carter looked at number of publications, citations, and average citation rates of each kind of publication, for both basic and medical science research fields. Among the various output measures, Carter found the average citation rates were the best correlate with the renewal priority scores. This correlation was refined further by eliminating publications that received no citations (a filter that also possibly removes errors introduced into the indexing by misspellings in the authors' names, also the majority of non-journal publications such as books and conference talks). Eventually the variable she chose to represent research quality for grants in further regressions was the average number of citations of all publications that were cited at least twice in the six years following publication.[167] "Since citations are cheap to obtain compared with peer judgments, they should be useful for questions of research policies."[168]

In attempting to evaluate the relationship between publication/citation data and the length of support awarded, Carter performed a regression looking at these variables as well. For this analysis she used two measures of research output: the number of articles which had been cited at least twice, "…on the grounds that only such publications communicated useful research results," and the average citation rate for these articles. On the other side of the comparison, she looked at the number of years of support and the total dollars awarded, controlled for new vs. renewal applications, and for research carried out at "very research oriented" schools (in this

[166] Carter, 39–40.
[167] Carter, 44.
[168] Carter, 45.

case the top 14 medical schools in terms of NIH support).[169] Both the number of publications and their average citation rate were related to the number of dollars granted and the years of support, though the relationship of the citation rate to years of support was weaker than the number of publications. In both cases, renewals and applications from research-intensive institutes produced more publications and higher-cited publications.[170]

A key question for Carter was the relationship between the priority score for the initial funding application (in this case made in 1967), number of years and dollars awarded, and the same research output measures (publications cited at least twice, and the average citation rate for those publications). In contrast to the previous analysis, there was no relationship found between the priority score in 1967 and the number of publications, only with the citation count, which Carter interpreted as an indication that the publication count was not an indicator of research quality.[171] Citation counts were subsequently used in Carter's evaluation of the NIH's process for selecting which grants to fund, and at what level:

> …at least for grant applications from most of the larger basic science and clinical departments of medical schools, the judgments of the peer review process are significantly related to an objective measure of research output derived from citations to articles describing the results of the grand. The first relationship points to the existence of criteria for scientific merit that are shared by different study section members. The second relationship shows that these criteria are related to objective measure of research output.[172]

This was a critical question during the time this work was being performed, as the level of resources being put into research was declining for the first time since WWII. Carter applies her models to "predict" what the effect would be if the level of support in 1967 had been lower. She

[169] Carter, 47.
[170] Carter, 48.
[171] Carter, 49.
[172] Carter, 51.

determined that if the funding cutoff for 1967 was the same as that as in 1973, only 40% of the funded projects would have been supported. However, the 10% of grants that produced the most highly cited papers would have been in the 40% that would have been funded at the lower level—the loss to science would have been the proposals that were on the margin of the funding cutoff.[173]

The work performed by Carter as part of her thesis research laid the groundwork for much of the overall work done by RAND in determining the effectiveness of the study sections in identifying projects that merited support, and the amount of that support. This work provided the analytical underpinning for using citation analysis as a tool for evaluating scientific merit, and was liberally cited in subsequent RAND reports looking at the value of investment in research and the allocation of manpower.

The methodology used in Carter's thesis is expanded upon in a second report, "A Comparison of Large Grants and Research Project Grants Awarded by the National Institutes of Health."[174] As the title states, the purpose of this analysis was to compare the overall quality of research performed under the auspices of different grant mechanisms such as the center grants and program project grants that were instituted in 1960. As such, it was again necessary to define and develop "measurable variables," or "surrogates that would describe research" in order to compare research output.[175] This report focused on two main surrogates for research quality: The peer-produced evaluations of applications for project grants produced by the NIH Initial Review Groups, and the number of times a journal article was cited. Another comparison looked directly at the difference in citations to articles published in clinical and basic science journals that were

---

173 Carter, viii.
174 Grace M. Carter, Clara S. Lai, and C. Lee, "A Comparison of Large Grants and Research Project Grants Awarded by the National Institutes of Health" (RAND Corporation, 1978).
175 Carter, Lai, and Lee, v.

supported by R01 Research Project Grants and larger program grants. They sought to compare the quality of work being conducted in the research centers that had been established at the NIH in the recent years, as compared to the research going on in academic centers that had been funded through a more competitive process.

In justifying their use of publication and citation analyses, the authors state: "In our earlier work we found that counts of publications and citations are linearly related to priority score and the logarithm of the dollar amount of the award."[176] While they do not provide a specific reference for the "earlier work," it is likely that they are referring to the report "Peer Review, Citations, and Biomedical Research Policy: NIH Grants to Medical School Faculty" by Carter. This focus on publications and citations as a research quality surrogate is supported by the authors' assertion that, "[s]ince much research is used primarily as input to further research, one aspect of the scientific quality of a research product is the extent to which it is used in subsequent research. The use of research should be highly correlated with citation counts for articles reporting the research."[177] This statement casts the entire goal of the NIH as the advancement of science, and omits the goal of improving health and longevity.

In this study, they expanded their analysis, looking not just at the number of publications and citations, but also at the journals published in to determine if contributions are being made to basic versus clinical science, and in the expected target fields versus others.[178] They also examined the co-authors of articles for evidence of interdisciplinary research (defined here as having collaborators with different disciplinary backgrounds).[179] The analysis of citation patterns revealed that the selection of references cited in a given article had more to do with the topic of

---

[176] Carter, Lai, and Lee, 9.
[177] Carter, Lai, and Lee, 58.
[178] Carter, Lai, and Lee, 36, 40.
[179] Carter, Lai, and Lee, 42.

the article than the field of the journal, and this citation pattern could be used to identify new "problem fields" and which kind of research would be the most productive in these fields.[180] Furthermore, they demonstrated that they were able to discriminate between research articles published by authors working under different NIH-backed programs in the same fields. This is more than just verifying the robustness of the model; the ability to benchmark the "productivity" of different programs would allow administrators and policymakers to direct resources towards the most successful programs, as well to the most pressing national priorities in medical research.[181]

Another motivation for commissioning the RAND was the shifting levels of available funding and competence in the medical research community. As stated in the Summary of the 1976 report prepared for the President's Biomedical Research Panel, "Policy Analysis for Federal Biomedical Research," between the end of WWII and the mid-1960s, there was a good working relationship between the executive and legislative branches of the federal government and the biomedical research community. As a result, the government appropriated sufficient funding for almost all proposals that were deemed sufficiently sound by their peers. That equilibrium shifted when the biomedical research community expanded to the point where there were more scientifically sound proposals than can possibly be funded. This added urgency to the need for an answer to the question of whether or not the protocols for determining which proposals would be funded were functioning effectively.[182] In other words, "[d]ecsionmakers could have greater confidence in the current practices and a better basis for initiating improvements if they knew the reliability of the priority scoring process for proposals, the

---

180 Carter, Lai, and Lee, 62.
181 Carter, Lai, and Lee, 63–64.
182 Albert P. Williams et al., "Policy Analysis for Federal Biomedical Research" (RAND Corporation, 1976), v.

assignment of proposals to IRGs, and the effectiveness of the categorical structure for funding Institutes."[183]

A simple cost-benefit analysis was judged to be inappropriate for such a complex question, due to difficulties in predicting scientific progress, a lack of understanding of the effects of federal and private programs on scientific activity, difficulties in translating the outcomes of scientific research into medical practice and improvements in health, and the difficulty in assigning a "social value" to improved health (even if the first three issues could be solved).[184] Several models of measuring scientific progress and its benefits were described; including what they call the "NIH" model as described in HEW's plan for FY 1975. This model imagined a linear flow of results "…from basic research to applied research and development to clinical investigation, clinical trials, and demonstration programs—in parallel with control programs and professional and public education programs."[185] This model is in line with the popular imagination of how research is done, but it lacked detail in their estimation. Another model put forth by Lewis Thomas distinguished between "high technology" which is based on understanding of a disease process and leads to a cure or prevention, and "halfway technology" which is more empirical and directed towards mitigating the damages of disease. The Sabin vaccine and the iron lung epitomize the two technologies, respectively, in the example of polio treatment and prevention.[186] At the time, this model was also considered to need more details to be considered useful.

The more complex models they described were more concerned with the interconnections between different streams of research and development, and between different areas of research.

---

[183] Williams et al., vi.
[184] Williams et al., 2.
[185] Williams et al., 7.
[186] Williams et al., "Policy Analysis for Federal Biomedical Research."

In order to do this, they turned to previous works that had established the use of citation analysis to define and understand behavior within the scientific community. This was not novel; as they state in a footnote, "Citation analysis of scientific literature has led to the development of a number of techniques for analyzing the underlying patterns of behavior within the scientific community. Citation analysis is based upon the assumption that the scientific literature represents the revealed nature of scientific development."[187] They specifically cite the work of Small and Griffith, which defines the scientist or single work group as the "basic unit" of science, units which are then organized into clusters defined by frequency of co-citations among published papers.

In their analysis of interactions between basic science, clinical science, and clinical practice, they found that articles on cancer research cited basic and non-cancer research journals more frequently than they did other cancer journals. This differed from other fields, such as applied chemistry—which cited other journals within the field more often than those in other fields. While the authors of the report do not attempt to answer the question of what might account for the different citation patterns in applied chemistry and cancer research, they posit that "…we have reached a point where empirical tests of theory treating questions like [these]… would be useful to federal policymakers."[188] The President's Biomedical Research Panel seems to have agreed, and commissioned cluster reports in an effort to assess the state of various fields, and if funding and manpower were being appropriately allocated.[189] This approach provided some degree of objectivity, as it did not rely on the opinions of experts in the field (who would likely be in one way or another beneficiaries of this government largesse). This method could

[187] Williams et al., 10.
[188] Williams et al., 15.
[189] Williams et al., 16.

also be used as a benchmark to compare different methods of peer review (such as the study sections at NIH and the "mail review" system at the NSF). In the various studies proposed to evaluate the worth of scientific funding and initiatives and priorities, citation data was routinely identified as the best surrogate for research quality and/or output.

The Policy Analysis report produced by RAND sought to answer two questions:

> First, is it possible to perform analysis—cost benefit analysis—that will provide *answers* to the questions of how much and where the federal government should make its investment in biomedical research? Second, if outright answers are not possible, is it possible to perform analysis that will lead to *better informed judgments* as to how much, where, and in what manner the federal government should commit funds to biomedical research.[190]

Overall, they concluded that a cost-benefit analysis was not possible. As to the second question, they concluded that such an analysis was possible, but there was no single approach that would yield a confident answer. They identified four areas of analysis that would be useful in determining government policy: the predictability of scientific advances, the links between federal program expenditures and scientific advances, the relationship of scientific advances to improved health, and assessments of the value society places on improved health and longevity. Analysis of the scientific literature and citation clusters was a key part of the first three of these four areas.[191]

At the same time as the RAND Corporation was conducting their analyses, a separate study was undertaken by Dr. Robert McGinnis's group at Cornell University's Research Program on Social Analyses of Science Systems under a two-year contract with NIH (OPEE

[190] Williams et al., 47.
[191] Williams et al., 47–49.

PA).[192] The objectives of this study were laid out in the highlights of a presentation given to the Office of the Director staff on June 1, 1976: "The research is concerned with measuring flows of R&D funds, trained scientists, *published knowledge*, the organizations through which they flow and their mutual impacts. A major goal of this research is to construct and evaluate quantitative models of the economic, social and intellectual processes involved in biomedical research."[193] They focused on identifying and characterizing the component processes of biomedical research, focusing on two in particular: "flows of new knowledge through the biomedical periodical literature and flows of trained scientists through their early careers."[194] For simplicity, they limited their analysis to a single, active area of research and discovery, reverse transcriptase, and analyzed several different variables associated with publication such as authors, citations, keywords, and funding sources. They believed that this provided a fuller picture of the growth in knowledge than looking at a single factor such as was typical in co-citation studies. They defined the most active researchers in the field as those that were the most highly published, and discovered that these authors formed a single "cosmopolitan" co-authorship network.[195] In addition to looking at the flow of knowledge, McGinnis's group also looked at patterns of early career researchers. Here too they considered the publications as an important metric of scientific productivity, and tracked authorship and publication histories along with employment histories to assess the impact of changes in the historical environment on job mobility.[196]

---

[192] Robert McGinnis, "Highlights of a Presentation of Research in Progress by the Program on Social Analyses of Science Systems, Cornell University. For Staff, NIH/OD, June 1, 1976," 1976.
[193] McGinnis, 1.
[194] McGinnis, 1.
[195] McGinnis, 3.
[196] McGinnis, 3–4.

Following these examples of citation analysis being used as at least a partial basis for policy recommendations, listing specific grant numbers in publications that were the result of NIH-supported research became a requirement. A request to include grant numbers in MEDLINE (the NLM's premier bibliographic database) was made in 1979 by the Evaluation Oversight Committee on Science Base Activities, and the Director of NIH agreed to it in principle. In a memo composed by Helen Hofer Gee, Chair of the Evaluation Oversight Committee on Science Base Activities, the reasons for this request were explicitly stated: "If we wish to improve the quality of program evaluation at NIH we must provide the tools that are needed by program staffs to monitor the effects of programmatic change. Further, if predictions of budgetary restrictions are realized we should be in a position to track the effects of reduced funding on the NIH role in bioscience research output."[197] This request can be viewed as an acknowledgement that research articles (and the citations to those articles) were an accepted indicator of scientific output that would need to be regularly tracked. This addition to the database enabled researchers outside of NIH to analyze the impact of NIH funding on other output indicators, such as pharmaceutical and biotechnology patents, and the career development of NIH-funded researchers compared to those who received funding from other sources.[198]

The importance of citation tracking as a metric of an individual researcher's reputation and standing in the scientific community is also seen in an "Action Item" recorded from the NIH Office of the Director's Staff Meeting of Jan. 5, 1979 to include the fact that 15% of the 300

---

[197] Helen Hofer Gee, "NLM Response to Request for the Addition of Grant Numbers to the Medline Data Base," 1979.

[198] Pierre Azoulay et al., "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules," *The Review of Economic Studies* 86, no. 1 (January 1, 2019): 117–52, https://doi.org/10.1093/restud/rdy034; P Azoulay, J S Graff Zivin, and G Manso, "Incentives and Creativity: Evidence from the Academic Life Sciences," *RAND Journal of Economics* 42, no. 3 (2011): 527–54, https://doi.org/10.1111/j.1756-2171.2011.00140.x.

most cited authors in the field had worked at NIH in the opening statement at House and Senate Appropriation Hearings.[199]

The need for such evaluations can be seen in the reactions to campus visits by Government Accounting Office representatives working on a project for the Senate Committee on Labor and Public Welfare on August 2, 1976. Their main objective was "…the establishment of a set of data having to do with programs and related impact-oriented statistics that would prove useful to the Senate Committee in its considerations and decisions having to do with budget, program authorizations, and periodic evaluations of programs."[200] William Rhode, the Acting Director of the Division of Resources Analysis, was bothered by "the assumption on the part of the representatives that biomedical research lent itself to the same kind of clear identification of direct and immediate outcome and impact measures that might be more appropriate for health services delivery activities…"[201]

Given this need to be able to provide data and statistics to Congress, it is not surprising that the NIH invested both financially and intellectually in the development of publication and citation analysis tools as a means to estimate research output and the relative value of federal programs and support.

---

199 J. Leonard Hooper, "Action Items - OD Staff Meeting, January 5, 1979," 1979.
200 William E. Rhode, "Visit by Representatives of the General Accounting Office on Monday, August 2" (Bethesda, 1976), 1.
201 Rhode, 2.

Chapter 5 Conclusions

Many factors contributed to the development of bibliometrics, and citation analysis in particular, as an important tool in determining the relative scientific merit of individuals and institutions. This thesis has focused on the role of the NIH in particular because of the unique role the NIH has played in the development of medical research in the U.S., and its direct involvement in supporting the work of Eugene Garfield. It is difficult to overstate the impact of either the NIH or Eugene Garfield in their respective communities, so it is worth considering how the stories of their development intersect.

The federal government's role in supporting research in the natural sciences was fairly limited before WWII. Post-WWII, there was a rapid expansion of this role, partly due to the rapidity with which wartime contracts were converted into peacetime grants,[202] and the existence of the pattern established by the Rockefeller Institute for project-based grants. By the time the NSF was established, fulfilling at least in part the vision contained in Vannevar Bush's *Science: The Endless Frontier*, the NIH was well-established and expanding.

The massive expansion in terms of funding, personnel, and infrastructure that took place during the late 1940s and 1950s yielded great benefits in terms of the health and wellbeing of US residents. Some of these were described, glowingly, in publications by the National Health Education Committee (founded and chaired by activist, lobbyist and philanthropist Mary

---

[202] Strickland, *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy*, 28.

Lasker). These reports aimed to demonstrate that research spending was, in fact, paying off for the American people. The reports list the spending in terms of the entire federal budget (of which it was of course a very small proportion), and the numbers of Americans who died of or were severely impacted by the major "killing and crippling diseases." From there the report tabulates the money lost due to lost productivity, and how much the government spent to "defend us" from these diseases. This amount is then compared to spending in other arenas, such as defense, forestry, and even the space program. This financial account is followed by descriptions of the "major research pay-offs" in these areas, as well as the number of lives saved, and how much, in dollars, the improved health of Americans has added to the national income and the tax revenue.[203]

At the time the Lasker reports were being compiled and circulated in the late 1950s-early 1960s, the NIH's budget was approaching the billion dollar mark, much of which was going to investigators at premiere academic research centers based on the evaluations of their peers, with fairly little administrative oversight. Concerns about potential mismanagement of funds led to the formation of the Fountain Committee, and its investigation of the financial management at the NIH.[204]

Overall, the Fountain Committee did not find any serious mismanagement of how the NIH selected projects to fund via the grants program. The Committee's final report did, however indicate that it felt there was not a sufficient focus on ensuring that the results of their research were communicated to other researchers, to health practitioners, and to the public at large. The

---

[203] National Health Education Committee, *Facts on the Major Killing and Crippling Diseases in the United States Today; Heart Diseases, Cancer, Mental Illness, Arthritis, Blindness, Neurological Diseases and Other Health Problems*.

[204] Baldwin, "Scientific Autonomy, Public Accountability, and the Rise of 'Peer Review' in the Cold War United States," 549.

directors of the NIH took this to heart, and while they felt that communication to the public and to practitioners was not entirely within their remit, they did feel that scientist-to-scientist communication was their responsibility, and they could do more in that regard. The NIH and the NLM began dedicating more resources, including computing resources, to creating better tools to allow researchers to search, identify, and access the latest publications by NIH-supported authors. It was at this time that Eugene Garfield received the critical financial and moral support from NIH that allowed him to create the first SCI.

Initially conceived as a tool for researchers and administrators to identify the "active front" of research, and to predict the most fruitful areas for additional work, the SCI data was used for many different types of evaluations. Garfield himself cautioned against some of these uses, notably that it would not be able to predict the winner of the Nobel prize, and that it should not be used to determine the quality of journals or institutes. However, some years after the development of the SCI, Garfield introduced the JCR, including the IF, which ranked journals based on the number of citations received in a year to the articles published over the previous two years. The IF has since become one of, if not the, most widely used indicator of journal quality in the natural sciences. In some countries, professors are explicitly rewarded financially for publishing in high impact journals, and some students are required to publish their thesis work in journals that meet a minimum IF.[205]

---

[205] Péter Vinkler, *The Evaluation of Research by Scientometric Indicators*, *The Evaluation of Research by Scientometric Indicators* (Elsevier Ltd, 2010), 202, https://doi.org/10.1533/9781780630250.

The IF has spawned imitators and competitors, such as Elsevier's CiteScore,[206] SCImago's Journal Rank,[207] and the Source-Normalized Impact per Paper, developed by Henk Moed at the University of Leiden.[208] Web of Science, which was originally developed at Garfield's Institute for Scientific Information but has since become part of Clarivate Analytics, has developed more sophisticated indicators as well.[209] Some of these have expanded on the basic IF by looking at a 5-year time period instead of only two years, others have applied weighting systems to the citing journals. Though these various indicators are all very similar in terms of evaluating the overall value of a journal to its target community, none of them has managed to gain the same level of prominence as Garfield's original IF.

In addition to journal-level metrics, today's natural scientist can also track the reach of their work by turning to author-level metrics (such as the h-index) or even article-level metrics.[210] Article-level metrics like Altmetric and PlumX have expanded the range of citations to include not just other scientific publications, but traditional and social media mentions as well.[211]

Providing seed money for Garfield was not the only time the NIH turned to bibliometrics to respond to Congressional concerns. When the Fountain Committee convened again in 1967 to

---

[206] "How Are CiteScore Metrics Used in Scopus? - Scopus: Access and Use Support Center," accessed January 1, 2020, https://service.elsevier.com/app/answers/detail/a_id/14880/supporthub/scopus/.
[207] "SJR - About Us," accessed January 1, 2020, https://www.scimagojr.com/aboutus.php.
[208] "CWTS Journal Indicators," accessed January 1, 2020, https://www.journalindicators.com/.
[209] "Web of Science - Web of Science Group," accessed January 1, 2020, https://clarivate.com/webofsciencegroup/solutions/web-of-science/.
[210] J. E. Hirsch, "An Index to Quantify an Individual's Scientific Research Output," *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 46 (November 15, 2005): 16569–72, https://doi.org/10.1073/pnas.0507655102.
[211] "Discover the Attention Surrounding Your Research – Altmetric," accessed December 18, 2019, https://www.altmetric.com/; "PlumX Metrics - Plum Analytics," accessed December 18, 2019, https://plumanalytics.com/learn/about-metrics/.

assess the NIH's progress on their 1961 recommendations, they were much less satisfied with the peer review system for determining which proposals to support than they had been in 1961. After these hearings were completed, the NIH contracted the RAND Corporation to conduct an independent inquiry into the effectiveness of the peer review system, the impact of different funding mechanisms, and the impact of the now shrinking federal budget on future research progress. The researchers at RAND relied largely on citation analyses to determine the answers to these questions—and they used Garfield's index to do it. Similarly, when the NSB was asked for an overall account of the status of American scientific endeavors, they included publications and citations to determine magnitude and predominance of American science relative to that of other nations. Again, they used Garfield's SCI for this information.

While the initial drive behind both Garfield's ambition to create a "map" of science through his index, and the NIH supported him out of a desire to provide scientists with a valuable tool to aid in their research and decision-making, the SCI also allowed the creation of new ranking systems for individual researchers, institutes, and journals. Citation analysis in the past was tedious work that had to be done "by hand," and was limited in scope. The introduction of the computerized database now allowed for broader analyses, across more fields, which could be updated frequently.

Publication of papers has long been closely tied to the reward system in science, as one of the main ways scientists communicate their discoveries, and establish priority and originality. Efforts to track publications by author was likewise not a new concept in the 20th century. In the late 1860s the Royal Society of London began publishing the *Catalogue of Scientific Papers*, which helped to establish the idea of scientific publishing as a distinct subset of publishing in general, as well as the idea that a scientist's productivity could be quantified. This led to the concept of "…the scientific paper… as a standardized unit, not only uniform enough, but also important enough, that it was the kind of thing that might be counted and compared."[212] If the system works well, then the scientists doing the best, most original and worthwhile work will receive the highest accolades. Or, as Robert Merton put it: "When the institution of science works efficiently… recognition and esteem accrue to those who have best fulfilled their roles, to those who have made genuinely original contributions to the common stock of knowledge. Then are found those happy circumstances in which self-interest and moral obligation coincide and fuse."[213]

Much of this rosy optimism about the self-regulation of science rests upon an unexpressed assumption that scientists cite papers based entirely on merit. While it is fairly safe to assume that poor-quality articles are not cited often, there are different reasons behind why individual articles are cited. A survey of authors citing one of the four most highly cited biochemistry articles produced by a single lab found that "…papers are perceived as having different kinds of significance, and these perceptions do not correspond directly to the functions that the cited texts play in the argumentation of citing texts."[214] In other words, an author might cite a specific claim, but give a vague reason for why it was cited—so one should be skeptical of efforts to categorize or analyze citation behavior in this field.

It is not my intention here to make a judgment on whether or not citation analysis or bibliometrics as a whole is an appropriate measure of scientific value, or to attempt to put

---

[212] Alex Csiszar, "How Lives Became Lists and Scientific Papers Became Data: Cataloguing Authorship during the Nineteenth Century," *British Journal for the History of Science* 50, no. 1 (March 1, 2017): 54, https://doi.org/10.1017/S0007087417000012.
[213] Robert K Merton, "Priorities in Scientific Discovery: A Chapter in the Sociology of Science," *American Sociological Review* 22, no. 6 (December 1957): 635, https://doi.org/10.2307/2089193.
[214] Loet Leydesdorff and Olga Amsterdamska, "Dimensions of Citation Analysis," vol. 15, 1990, 326.

forward any conclusions regarding the behavior of scientists as authors or as collective mediators of what is, or is not, scientifically sound and valuable. My goal was to examine the circumstances that contributed to the NIH's support of the creation of the most important citation index in use today, and indirectly the indicators based upon it. While the NIH is first and foremost a medical research center and a funding body for basic and clinical research, it is also an agency of the federal government. It is this role and the accountability that it entails that provided the impetus for the NIH's support and utilization of citation maps and analysis. The NIH's initial involvement as a funder of Garfield's SCI was followed over a decade later by the NIH's use (through a contract with RAND Corporation) of citation analysis to evaluate its own funding programs and practices. Since then the IF (and its competitors and imitators) has become an important consideration not just for librarians deciding which journals need to be on their shelves, but for scientists deciding to which journals to submit their work. This is one example of the quality being measured influencing behavior.

On a darker note, several cases of scientists (as authors, reviewers, and even editors) manipulating citation behavior have come to light in recent years, representing another way the act of measurement may be influencing behavior.[215] As with any indicator, it is important to keep in mind that citation-related metrics are not direct measures of what they are indicating. Looking back at the forces that contributed to the development of the first large-scale citation index and the adoption of citation analysis (and its offshoots) can provide a more complete understanding of the ways in which this indicator may have influenced the field it was seeking to measure.

---

[215] Retraction Watch lists several cases representing a wide range of citation-related misconduct across multiple fields (https://retractionwatch.com/?s=citation+manipulation).

Bibliography

Anderson, E. Packard. "Letter to James Kelly, May 17 1963," 1963.

"Archives and Modern Manuscripts Oral Histories: James A. Shannon : An Oral History / Interviewed by Thomas J. Kennedy, 1984 Jan. 11." Accessed November 3, 2019. https://oculus.nlm.nih.gov/cgi/t/text/text-idx?c=oralhist;cc=oralhist;g=ammpoh;xc=1;xg=1;q1=shannon;idno=9802979;view=toc;frm=frameset.

Azoulay, P, J S Graff Zivin, and G Manso. "Incentives and Creativity: Evidence from the Academic Life Sciences." *RAND Journal of Economics* 42, no. 3 (2011): 527–54. https://doi.org/10.1111/j.1756-2171.2011.00140.x.

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat. "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules." *The Review of Economic Studies* 86, no. 1 (January 1, 2019): 117–52. https://doi.org/10.1093/restud/rdy034.

Baldwin, Melinda. "Scientific Autonomy, Public Accountability, and the Rise of 'Peer Review' in the Cold War United States." *Isis* 109, no. 3 (September 19, 2018): 538–58. https://doi.org/10.1086/700070.

Bauer, R., and American Academy of Arts and Sciences, eds. *Social Indicators*. Cambridge: MIT Press, 1966.

Bush, Vannevar. *Science, the Endless Frontier; a Report to the President on a Program for Postwar Scientific Research*. Washington: National Science Foundation, 1960.

Carson, Bruce P. "Report on October 3-4 Meetings of NAS-NRC Advisory Committee on Scientist-to-Scientist Communication," 1963.

Carter, Grace M. "A Citation Study of the NIH Peer Review Process:" RAND Corporation, 1978.

———. "Peer Review, Citations, and Biomedical Research Policy." RAND Corporation, 1974.

Carter, Grace M., Clara S. Lai, and C. Lee. "A Comparison of Large Grants and Research Project Grants Awarded by the National Institutes of Health." RAND Corporation, 1978.

Carter, Grace M., Clara S. Lai, and Carolyn L. Lee. "Sample Selection and Data Collection for a Study of Research Performed Under NIH Grants," 1976.

Committee on Government Operations. “Health Research and Training: The Administration of Grants and Awards by the National Institutes of Health: Second Report.” Vol. 1, 1961.

———. “The Administration of Research Grants in the Public Health Service, Ninth Report.” Washington, D.C., 1967.

Csiszar, Alex. “From the Bureaucratic Virtuoso to Scientific Misconduct: Robert K. Merton, Eugene Garfield, and Goal Displacement in Science.” In *Annual Meeting of the History of Science Society*. Toronto, 2017.

———. “How Lives Became Lists and Scientific Papers Became Data: Cataloguing Authorship during the Nineteenth Century.” *British Journal for the History of Science* 50, no. 1 (March 1, 2017): 23–60. https://doi.org/10.1017/S0007087417000012.

“CWTS Journal Indicators.” Accessed January 1, 2020. https://www.journalindicators.com/.

Dennis, Michael Aaron. “Historiography of Science: An American Perspective.” In *Science in the Twentieth Century*, edited by John Krige and Dominique Pestre, 1–26. Amsterdam: Harwood Academic Publishers, 1997.

“Discover the Attention Surrounding Your Research – Altmetric.” Accessed December 18, 2019. https://www.altmetric.com/.

Duncan, Otis Dudley. “Science Indicators and Social Indicators.” In *Toward a Metric of Science: The Advent of Science Indicators*, edited by Yehuda Elkana, Joshua Lederberg, Robert K. Merton, Arnold Thackray, and Harriet Zuckerman, 31–38. New York: John Wiley & Sons, Inc., n.d.

Elkana, Yehuda, Joshua Lederberg, Robert K. Merton, Arnold Thackray, and Harriet Zuckerman, eds. *Toward a Metric of Science*. New York, New York: John Wiley & Sons, Inc., 1978.

Garfield, Eugene. “Citation Indexes for Science; a New Dimension in Documentation through Association of Ideas.” *Science (New York, N.Y.)* 122, no. 3159 (July 15, 1955): 108–11. https://doi.org/10.1126/SCIENCE.122.3159.108.

———. “‘Science Citation Index’ --A New Dimension in Indexing.” *Science (New York, N.Y.)* 144, no. 3619 (May 8, 1964): 649–54. https://doi.org/10.1126/science.144.3619.649.

Gee, Helen Hofer. “NLM Response to Request for the Addition of Grant Numbers to the Medline Data Base,” 1979.

Gingras, Yves. *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge: The MIT Press, 2016.

Godin, Benoît. *Measurement and Statistics on Science and Technology : 1920 to the Present*. 1st ed. New York: Routledge, 2005.

———. “The Numbers Makers: Fifty Years of Science and Technology Official Statistics.” *Minerva* 40, no. 4 (2002): 375–97.

Gustafson, Thane. “The Controversy over Peer Review.” *Source: Science, New Series* 190, no. 4219 (1975): 1060–66.

Harden, Victoria A. *Inventing the NIH: Federal Biomedical Research Policy, 1887-1937*. 1st ed. Baltimore: The Johns Hopkins University Press, 1986.

Haynes, Charles G. “Letter to Secretary Folsom, September 12, 1956,” 1956.

Hirsch, J. E. “An Index to Quantify an Individual’s Scientific Research Output.” *Proceedings of the National Academy of Sciences of the United States of America* 102, no. 46 (November 15, 2005): 16569–72. https://doi.org/10.1073/pnas.0507655102.

Hooper, J. Leonard. “Action Items - OD Staff Meeting, January 5, 1979,” 1979.

“How Are CiteScore Metrics Used in Scopus? - Scopus: Access and Use Support Center.” Accessed January 1, 2020. https://service.elsevier.com/app/answers/detail/a_id/14880/supporthub/scopus/.

Kidd, Charles V. “Notes on Luncheon Discussion next Tuesday with MacLeod, et Al,” 1963.

Land, Kenneth C. “On the Definition of Social Indicators.” *The American Sociologist* 6, no. 4 (1971): 322–25.

Leydesdorff, Loet, and Olga Amsterdamska. “Dimensions of Citation Analysis.” Vol. 15, 1990.

Lotka, Alfred J. “The Frequency Distribution of Scientific Productivity.” *Journal of the Washington Academy of Sciences* 16 (1926): 317–23. https://doi.org/10.2307/24529203.

McGinnis, Robert. “Highlights of a Presentation of Research in Progress by the Program on Social Analyses of Science Systems, Cornell University. For Staff, NIH/OD, June 1, 1976,” 1976.

Merton, Robert K. “Priorities in Scientific Discovery: A Chapter in the Sociology of Science.” *American Sociological Review* 22, no. 6 (December 1957): 635. https://doi.org/10.2307/2089193.

———. “The Matthew Effect in Science.” *Science, New Series* 159, no. 3810 (1968): 56–63.

Miller, Charles. “Administrative Problem Areas,” 1964.

Murtaugh, J.S., and G.L. Payne. “Communication in the Biomedical Sciences.” *Journal of*

Medical Education 37, no. 11 (1962): 1169–82.

Murtaugh, Joseph S. "Discussions with Dr. McCloud Re Review of NIH," 1963.

National Health Education Committee. *Facts on the Major Killing and Crippling Diseases in the United States Today; Heart Diseases, Cancer, Mental Illness, Arthritis, Blindness, Neurological Diseases and Other Health Problems*. New York, 1961.

National Institutes of Health (U.S.) Office of Program Planning. "Communication in the Bio-Medical Sciences: A Report by the National Institutes of Health to the Subcommittee on Labor, Health, Education and Welfare Appropriations of the House Appropriation Committee." Bethesda, 1962.

National Science Board. "Science Indicators, 1972: Report of the National Science Board." [Washington, 1973.

NIH Study Committee. "Biomedical Science and Its Administration; a Study of the National Institutes of Health." Washington, D.C., 1965.

"Notes for Dr. Shannon's Discussion with Dr. Wiesner, April 3, 1963," 1963.

"Notes for Dr. Shannon on Mr. Siepert's Presentation to Institute Directors on the House Investigations Committee," 1956.

"Notes on Possible Chair for Wiesner Committee, Aug. 20 1963," 1963.

Payne, G.L. "Notes on Special Reports Requested by Congressional Committees," 1961.

"PlumX Metrics - Plum Analytics." Accessed December 18, 2019. https://plumanalytics.com/learn/about-metrics/.

Porter, Theodore M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Course Boo. Princeton: Princeton University Press, 1996.

President's Science Advisory Committee. "Scientific Progress, the Universities, and the Federal Government." Washington, D.C.: United States Government Printing Office, 1960.

Price, Derek J. de Solla. *Little Science, Big Science... And Beyond*. New York: Columbia University Press, 1986.

Price, Derek J D. "Networks of Scientific Papers: The Pattern of Bibliographic References Indicates the Nature of the Scientific Research Front." *Science* 149, no. 3683 (1965): 510–15.

"Research Highlights National Institutes of Health 1966: Items of Interest on Research Studies Conducted and Supported by the Institutes and Divisions of NIH, as Presented to the Congress of the United States." Washington, D.C., 1967.

Rhode, William E. “Visit by Representatives of the General Accounting Office on Monday, August 2.” Bethesda, 1976.

Rogers, Frank B. “MEDLARS Meeting,” 1962.

Schneider, William H. “The Origin of the Medical Research Grant in the United States: The Rockefeller Foundation and the NIH Extramural Funding Program.” *Journal of the History of Medicine and Allied Sciences* 70, no. 2 (April 2015): 279–311. https://doi.org/10.1093/jhmas/jrt074.

Shannon, James A. “Communication of the Results of Research,” 1961.

———. “Wooldridge Committee Survey of the National Institutes of Health,” 1963.

“SJR - About Us.” Accessed January 1, 2020. https://www.scimagojr.com/aboutus.php.

Small, Henry. “Citation Indexing Revisited: Garfield’s Early Vision and Its Implications for the Future.” *Frontiers in Research Metrics and Analytics* 3 (March 2, 2018): 8. https://doi.org/10.3389/frma.2018.00008.

Strickland, Stephen P. *Politics, Science, and Dread Disease: A Short History of United States Medical Research Policy*. 1st ed. Cambridge: Harvard University Press, 1972.

“The Administration of Grants by the National Institutes of Health. Hearings before a Subcommittee of the Committee on Government Operations, House of Representatives, Eighty-Seventh Congress, Second Session. March 28-30, 1962.” 1962.

Urquhart, D.J. “Use of Scientific Periodicals.” In *Proceedings of the International Conference on Scientific Information*, 287–300. Washington, D.C.: National Academies Press, 1959. https://doi.org/10.17226/10866.

Vinkler, Péter. *The Evaluation of Research by Scientometric Indicators*. *The Evaluation of Research by Scientometric Indicators*. Elsevier Ltd, 2010. https://doi.org/10.1533/9781780630250.

“Web of Science - Web of Science Group.” Accessed January 1, 2020. https://clarivate.com/webofsciencegroup/solutions/web-of-science/.

Williams, Albert P., Grace M. Carter, Alvin J. Harman, Emmett B. Keeler, Willard G. Manning, C. Richard Neu, L. Pearce, and Richard A. Rettig. “Policy Analysis for Federal Biomedical Research.” RAND Corporation, 1976.

Wouters, Paul. “The Citation Culture.” University of Amsterdam, 1999