Garfield and the Impact Factor

Stephen J. Bensman
Louisiana State University

Introduction

This is the first of two chapters on Eugene Garfield and the impact factor, the second part of which will be published in next year's volume of the *Annual Review of Information Science and Technology*. Garfield is the founder of the Institute for Scientific Information (ISI), now Thomson Scientific, which launched the *Science Citation Index* (*SCI*), the *Social Sciences Citation Index* (*SSCI*), and the *Arts & Humanities Citation Index* (*A&HCI*). The impact factor is a citation measure that Garfield created during the process of developing these indexes. Simply defined, the impact factor is the ratio of the total number of citations in the journals covered by ISI during a processing year to the issues a given journal published during the two years preceding the processing year divided by the number of citable source items published in this journal during those two preceding years.

This chapter is historical, comprising an intellectual biography of Garfield. It discusses in detail the political, social, and intellectual influences affecting Garfield in his creation of the impact factor. The focus is on Garfield himself, the small group of intellectuals comprising his mentors and colleagues—Garfield's "invisible college" in Price's (1963, pp. 85–91) terminology—and his company, ISI. The thought processes and analyses surrounding the creation of the impact factor are described and placed in the theoretical structure that Garfield developed. Finally, the way the impact factor was utilized in determining the journal coverage of ISI's indexes is analyzed.

This chapter is not a comprehensive treatment of citation indexing, citation measures, or the use of the latter for the evaluation of journals, academic personnel, and academic programs. The emphasis is upon Garfield's own writings about these matters and upon those of persons who either greatly influenced his intellectual development or to whom he felt compelled to respond.[1] Other writings are discussed only insofar as they highlight or validate the principles on which Garfield based his ideas.

The Life and Career of Eugene Garfield

Eugene Garfield created an essential component of the empirical bases for modern information science through the application of citation indexing to the sciences, social sciences, arts, and humanities. He was born in the Bronx borough of New York City on September 16, 1925, to Henry and Edith (née Wolf) Garfinkle, who were themselves the children of immigrants. Garfield's parents separated while his mother was pregnant. As a result, his upbringing was heavily influenced by his uncles, whom Garfield has described in the following manner:

> Three of my uncles were radicals, involved in labor organizing and socialist-communist politics. There were two sides to the family. Two uncles were capitalists and they were always opposed to that. So there was a lot of turmoil going on. They cultivated my interest in science, classical music—in fact almost everything. Atheism was part of it. I was discouraged from going to any kind of religious school, so I did not go. Most Jewish kids of that age would have been Bar Mitzvahed, but I was not. (Garfield, 1987a, p. 2)

Garfield (1999a) reports being affected by political discussions with his uncles. However, it should be noted that Garfield's surrogate father was one of the capitalist uncles, who was a successful ladies' garment manufacturer and helped support his mother. Garfield's biological father was a successful newspaper-magazine distributor, whose firm's name was the Garfield News Company. Over the father's objections, the uncle had Garfield adopt the name of this company. It seems that the capitalist family influences predominated over the Marxist ones, but this family turmoil made Garfield an ideologically complex character. On the one hand, he was attracted to radical ideas; yet, on the other, he was a consummate businessman capable of implementing these radical ideas through the formation of a private corporation when governmental and societal organizations were too conservative and hidebound to do so.

Garfield came to information science by way of chemistry. Upon his graduation from high school in 1942, he enrolled at the University of Colorado to study chemical engineering. This choice may have been influenced by one of his uncles having a chemistry set. Because of a combination of personal problems and the patriotic fervor of World War II, Garfield lasted at Colorado only one semester. Not yet old enough for military service, he worked as a shipyard welder until he was drafted into the Army. After the war, Garfield obtained a B.S. in chemistry in 1949 from Columbia University in his native New York City and worked for a while as a laboratory assistant at Columbia for Professor Louis W. Hammett. In 1951 Garfield attended the Diamond Jubilee Conference of the American Chemical Society, where he discovered its Division of Chemical Literature and the possibility of a career in scientific information.

The turning point in Garfield's life occurred in 1951, when he was hired as a staff member of the Welch Medical Library Indexing Project. This project was headed by Sanford V. Larkey, director of the William H. Welch Medical Library at Johns Hopkins University in Baltimore. It had begun in 1948 and was sponsored by the Army Medical Library, which eventually became the National Library of Medicine. The Welch Project's mission was to investigate the size of biomedical literature, the extent of this literature's coverage by existing indexes, the structure of medical subject headings, and the application of machine methods to medical indexing. In his accounts of the project, Garfield (1985b; 1987a, pp. 10–13; 1997b, pp. 1–26; 1999c, pp. 240–244; 2000, pp. 1–2; 2002, p. 22) has stated that he was hired for his chemical expertise and that much of his time was spent in subject authority work involving the definition and classification of subject headings. His special task was to standardize chemical nomenclature. A sign of his strong interest in this area is that at the same time he voluntarily worked for *Chemical Abstracts*, summarizing articles on pharmacology in Spanish journals. Also taking an interest in machine indexing, Garfield worked out the "programs" needed to wire the punched-card machines that were used to prepare and tabulate subject-heading lists, and played a key role in the First Symposium on Machine Methods in Scientific Documentation that was organized by the Welch Project. While working on this project, Garfield manifested that combination of scientific interest and entrepreneurship which typified his career. To keep himself better informed on what was happening worldwide in documentation, he started in 1952 the first contents-page service he ever published: entitled *Contents in Advance*, it covered library science as well as documentation journals.

Garfield's participation in the Welch Medical Library Indexing Project set the course of his life, for through it he came to know most of the pioneers in information science. After leaving the project, Garfield strengthened his academic credentials in information science. In 1954 he obtained an M.S. in library science from Columbia University. Garfield (1997b, p. 32) has described the library degree as making him "official or Kosher" for librarians. He further reinforced his academic credentials in 1961, earning a doctorate in structural linguistics from the University of Pennsylvania with a dissertation entitled "An Algorithm for Translating Chemical Names to Molecular Formulas." During this period, Garfield continued his pattern of combining scientific interest with entrepreneurialism. He accepted a position as a consultant with the pharmaceutical firm Smith Kline and French in Philadelphia, a move that led him to make that city his permanent base of operations. At the same time, he launched a firm called Eugene Garfield Associates, whose name he changed in 1960 to the Institute for Scientific Information (ISI). Garfield (1999c, p. 249; personal communication, Oct. 5, 2005) reports that he was inspired in this choice of corporate name by the U.S.S.R.'s All-Union Institute of Scientific and Technical Information, citing two reasons for this selection: to cause the company to be perceived as a nonprofit organization and to highlight his small company's ability to achieve what the huge Soviet organization was attempting to do. This firm published two contents-page services on the model of *Contents in Advance*. These services marked the start of the *Current Contents* publications that were for a long time ISI's financial mainstay. One covered management literature and was secured by a contract with AT&T's Bell Laboratories. The other was dedicated to chemical, medical, pharmaceutical, and life sciences publications.

Garfield had conceptualized the citation indexing of science before he established ISI. While still at the Welch Project, he mentally assembled the various elements of this method of covering scientific journal literature. As Garfield (1998a, p. 68) reports, in early 1954 he wrote a term paper proposing the creation of citation indexes at the Columbia University School of Library Service. With the help of Johns Hopkins biologist Bentley Glass, Garfield (1955) revised this paper and was able to publish it in the prestigious journal *Science* in 1955. The article attracted the attention of Joshua Lederberg (2000), a Stanford University geneticist who received the 1958 Nobel Prize for medicine. Lederberg wrote to Garfield about the latter's idea, starting a vigorous correspondence that opened the way for Garfield to implement his idea of applying citation indexing to science. Under Lederberg's guidance, Garfield obtained in 1961 a National Institutes of Health (NIH) grant to produce a citation index for genetics. The NIH grant had to be converted into a National Science Foundation (NSF) contract because ISI was a private company. ISI approached the construction of the genetics citation index by first preparing a multidisciplinary science index. Garfield proposed that the NIH and NSF publish this multidisciplinary science citation index. When these bodies rejected the proposal, he decided to launch the index himself at considerable financial risk to his own company. The rejection fortified Garfield's (1975, p. 5) belief that private enterprise was "the best and most economical way to serve the scientific community."

ISI began regular publication of the *Science Citation Index* in 1964. It introduced the *Social Sciences Citation Index* in 1973 and the *Arts & Humanities Citation Index* in 1978. By showing how many times a given work has been referenced by other works, these indexes made practical the widespread use of citations in evaluating the importance—or "impact"—of works, scientists, and scholars, as well as research institutions. They also created an opening for new methods in exploring the history, sociology, and relational structure of science and scholarship. ISI added to the analytical capabilities of the *SCI* in 1975 by making a regular component of this index a volume entitled *Journal Citation Reports* (*JCR*) and made the same addition to the *SSCI* in 1977. These *JCR*s contained numerical data on the journals covered by ISI's citation indexes during a given year, such as total number of citations to these journals, the number of source items in these journals, the rapidity with which an average journal item was cited, how many times a given journal cited itself and other journals, as well as how many times other journals cited a given journal. However, the centerpiece of the *JCR*s came to be a measure called the "impact factor." In its original form, it was the ratio of the total number of citations during a processing year to the issues a given journal published during the two years preceding the processing year divided by the total—not just "citable"—number of source items published in this journal during those two preceding years. The impact factor can be, and has often been, considered an arithmetic mean. Thus, in the first *SCI JCR* (Garfield, 1976b, p. 6), the impact factor was defined as "a measure of the frequency with which the 'average cited article' in a journal has been cited in a particular year." In spite of its seeming simplicity, the impact factor is an exceedingly complex measure. It has become widely used for purposes that go far beyond the original intention of the measure and for which it is structurally flawed. Such utilization of the impact factor has occurred primarily in Europe in regard to the evaluation of scientists and scientific programs, raising such a storm of controversy that Garfield (1999a, p. 26) commented:

> I do find it hard to keep up with the large literature involving journal impact factors. I am especially frustrated that I can't respond to the portion containing misstatements or misuses. There is much controversy about the validity of impact factors, which are used for many purposes. ... *SCI* and *Journal Citation Reports* (*JCR*) data have become institutionalized. People often criticize the impact factor because it is so pervasive.

Garfield's (2005) level of frustration on this matter is revealed by the title of his talk at the International Congress on Peer Review and Biomedical Publication: "The Agony and the Ecstasy—the History and Meaning of the Journal Impact Factor." He even considered as an alternative title "Citation Sanity and Insanity—the Obsession and Paranoia of Citations and Impact Factors." However, a major reason for this situation is that Garfield himself made the impact factor seemingly his main measure of scientific value, doing so for reasons peculiar to his own intellectual development. Moreover, many of the problems in the validation of the impact factor arise from actions he took and processes he discovered.

The Theoretical Framework of the Science Citation Index

Early British Intellectual Influences

The theoretical framework within which Garfield developed his ideas was constructed in Britain during a scientific revolution that laid the bases for modern information science. This revolution was a consequence of World War I, which demonstrated the need for a modern scientific information system. The persons driving the information science revolution in Britain can be categorized as either reformists or revolutionaries. The reformists wanted, or were willing, to retain the journal as the basis of the scientific information system, whereas the revolutionaries wanted to make the scientific paper the basis of this system.

The Reformist Program

S. C. Bradford, head of the Science Museum Library (SML) in London, was the main theorist of the reformist camp. He was also a leading figure in information science—or "documentation," as it was known in this period—at the national and international level, being the director of the British Society for International Bibliography. Bradford approached the problem of scientific literature from the perspective of its adequate coverage by indexing and abstracting services. To analyze this coverage, he conducted research at the SML, the results of which showed that, although the 300 abstracting and indexing journals in existence at the time noticed 750,000 articles each year, only 250,000 articles were covered due to duplication of effort (Bradford, 1934). Suspecting that the reason for this oversight lay in the way articles on a given scientific subject were distributed among journals, Bradford investigated this distribution in two subject areas: applied geophysics and lubrication. The distributions in both areas were found to be remarkably similar, leading to the formulation of the Law of Scattering that is associated with Bradford's name. According to this law, the articles on a given scientific topic concentrate in a small nucleus of journals and then scatter across other journals in zones that must increase exponentially in number of titles to contain the same number of articles on the topic as contained in the journals of the nucleus. For example, in 1928–1931 the articles on applied geophysics were distributed across journals in the following manner: a small nucleus of nine journals (2.8 percent) contained 429 articles (32.2 percent), a second zone of 59 journals (18.1 percent) encompassed 499 articles (37.5 percent), and a third zone of 258 journals (79.1 percent) comprised 404 articles (30.3 percent). Bradford pointed out two practical consequences of the Law of Scattering. First, to gain complete coverage of articles on a subject, an indexing service would have to scrutinize for long periods thousands of journals, the vast bulk of which would haphazardly provide only occasional references. Second, no special library could gather all the literature on its subject without becoming a general science library. Bradford noted that in practice only one-third of the content of special libraries was definitely related to their subject scope, with the other two-thirds comprising literature on borderline and less related subjects. He pointed out that this situation led to much duplication in special library holdings.

World War II was a quintessentially technological war that severely stressed the British scientific information system, particularly affecting special libraries. The war exacerbated the problem of handling a scientific literature growing exponentially in both size and complexity. As a result, the Royal Society Empire Scientific Conference of 1946 dedicated a session to scientific information services at which Bradford (1948a) delivered a paper that summarized both the development of the international documentation movement and his lifetime's work. The paper contained a detailed explanation of the Law of Scattering and the practical dilemmas arising from it. Shortly after the conference, Bradford (1948b) gave further prominence to his ideas by publishing his classic book *Documentation*. Together, the conference paper and the book set forth Bradford's recommendations on how to handle the dilemmas resulting from his law. In respect to indexing and abstracting services, Bradford (1948b, pp. 28–29) urged the adoption of the Universal Decimal Classification (UDC), which he believed would improve bibliographical cooperation by enabling the collaboration of many persons in the task of indexing by subject the world's literature. At the conference, Bradford (1948a, p. 745) suggested using the UDC to classify journals according to the subjects treated by existing abstracting agencies; the agencies would then forward any article outside their scope but published in their allotted periodicals to the agency interested in that subject. Aware of objections, he emphasized, "*No interference is involved with the internal work of any abstracting agency*" (italics in original). As for the problems afflicting special libraries, Bradford (1948b, pp. 64–84) in his book made a detailed case for developing one of the great libraries of the U.K. into a comprehensive national central library of science and technology, on which the special libraries could rely for articles in journals dedicated to subjects peripheral to their own area of interest. One of Bradford's life ambitions had been to convert the SML into just such a library: He made it into Britain's main backup library in science, pioneering the concept of document delivery.

The Revolutionary Program

The primary theorist of the revolutionary wing of British information science was J. D. Bernal. One of Britain's most eminent scientists, Bernal pioneered x-ray crystallography and was one of the founders of molecular biology. In 1937 he was elected a fellow of the Royal Society and was appointed to the Chair of Physics at Birbeck College of the University of London, where he spent most of his career. Bernal had been born in Ireland in 1901. His father was a descendant of Spanish Sephardic Jews who had converted to Catholicism; his mother, the daughter of an American Presbyterian clergyman. As an undergraduate at Cambridge University, Bernal became a Marxist and joined the Communist Party of Great Britain in 1923.

Bernal's conversion to Marxism led him to become a pioneer in the study of the relationship of science to society. His biographer Hodgkin (1980, p. 64) states that his interest in this facet of science was powerfully strengthened and sharply focused in 1931 at the Second International Congress of the History of Science and Technology, where a Soviet delegation of theoreticians, historians, and scientists led by Bukharin (1931) set forth before Western intellectuals for the first time the view of science then predominant in the U.S.S.R. The Russian contribution to this conference had a profound impact in Britain—especially upon Bernal (1931, p. 43), who summed up this impact in a report entitled "Science and Society":

> The old conception of the history of science, the bare enumeration of discoveries and inventions, the telling of lives and deeds of great men, and the drawing up of the genealogical tree of present knowledge, is now seen as a partial though necessary basis for the study of the interaction of science with economics and politics, with religion, art and industry, throughout the whole course of history, not least in the present.

Hodgkin (1980, p. 64) considers this article as a point of departure for Bernal's subsequent writings on the history of science and science policy.

Bernal's (1940) most important contribution to the study of the relationship of science to society was his book entitled *The Social Function of Science*, first published in 1939. In a Festschrift celebrating the 25th anniversary of its publication, Goldsmith and Mackay (1964, p. 11) point out that *The Social Function of Science* presented the views not only of Bernal himself but also of a large school of scientists and others, who formed a kind of "invisible college," most of whose members were influenced by Marxism to a greater or lesser degree. Outlandish as it may seem in the light of the mass famines of collectivization, forced labor camps, and political purges of Stalin's rule, Bernal (1940, p. xiii) pointed to the Soviet Union as a model of "the possibility of combining freedom and efficiency in scientific organization." As any good Marxist, Bernal (pp. 10–11) considered science to be a part of industrial production and its importance, a result of its contribution to profits. However, although Bernal (pp. 408–416) thought capitalism essential to the early development of science, he now regarded the continuance of this economic system as incompatible with the full development of science in the service of humanity. For Bernal, science implied a unified, coordinated, and conscious control of the whole of social life. He thought that the function of science was to serve as the main engine of social change. According to Bernal, the relevance of Marxism to science was that it showed science to be an important part of economic and social development. Bernal defined science as communism because in science men had learned consciously to subordinate themselves to a common purpose. The doctrine expounded in the book came to be called "Bernalism," which was defined by one of its opponents, Oxford biologist John Baker (1939, p. 174), in the following manner:

> Bernalism is the doctrine of those who profess that the only proper objects of scientific research are to feed people and protect them from the elements, that research workers should be organised in gangs and told what to discover, and that the pursuit of knowledge for its own sake has the same value as the solution of crossword puzzles.

Bernal's role in the effort to nationalize the British scientific information system has been well documented by Muddiman (2003, 2004). Bernal was a leading member of the Association of Scientific Workers (AScW), a trade union of Communist, socialist, and liberal scientists. *The Social Function of Science* contains the most detailed exposition of his ideas on the proper scientific communication system. Bernal (1940, pp. 292–293) summarized the main problem affecting this system and his proposed solution in the following terms:

> The present mode of scientific publication is predominantly through the 33,000 odd scientific journals. It is ... incredibly cumbersome and wasteful, and is in danger of breaking down on account of expense. What can we put in its place? The prime function of scientific publications is to convey information about acquired knowledge, but it is clear that whereas certain information is needed by certain workers in full detail, the great bulk of it is only needed by any given worker in outline, if at all. An adequate system of communication would consist in principle of a limited distribution of detailed accounts, a wider distribution of summaries or abstracts, and the frequent production of reports or monographs covering the sum of recent advances in any given field. Behind this must be a body of readily accessible archives in which reference can be made to the work of the past.

He likened the problem of scientific communication to the problem of distribution and storage that was being solved every day by large businesses and mail order houses.

Bernal's proposed reform entailed the abolition of all existing scientific journals and their replacement by a service that would record, file, coordinate, and distribute scientific information. He regarded the publication of scientific journals as an inefficient method plagued by overlap and lack of coordination for distributing large amounts of scientific information. Bernal stated that a scientific publication was read to an extent of only 10 percent, when taken by an individual, but demanded simultaneously by a dozen persons, when taken by a library. The obvious solution was to make the separate paper, rather than the journal, the unit of communication between scientists. The science service that Bernal envisioned would ensure that all relevant information would be available to each research worker in amplitude proportional to the degree of its relevance and without any effort on the part of the worker. He suggested as a model Watson Davis's (1940) proposal for a science service for the U.S., which Bernal published as an appendix in his book. Davis was an interesting character who, in 1925 as managing editor of Science Service, covered the Scopes Monkey Trial in Tennessee and in 1937 founded the American Documentation Institute, which ultimately became the American Society for Information Science and Technology. Davis's proposal was typical of the era of the Great Depression and New Deal. He called for a central organization, called the Scientific Information Institute (SII), which would take over the publication functions of many existing societies and journals. All scientific bibliographical and abstracting services as well as many of the journals then under financial stress were to be brought under the SII, which, he envisaged, would be a monopoly in the same sense as the post office was a monopoly, that is, one operated for the public benefit without profit. Davis listed three factors he considered essential to his project: (1) centralization of scientific publication and bibliography with resulting economy of operation and improvement of service, (2) substitution of photographic duplication for printing-from-type duplication, and (3) the utilization of a comprehensive scheme of numerical indexing and automatic finding and sorting devices for filing and selecting bibliography.

In *The Social Function of Science*, Bernal essentially endorsed the type of scientific communication system Davis had proposed, notably its organizational and technological aspects. In addition, he emphasized two aspects of such a system that would prove to be of great import. First, Bernal (1940, pp. 297–298) stressed the importance of review literature in scientific communication. In doing so, he recommended following Lord Stamp's suggestion that the responsible body in each scientific discipline periodically review its field and report for each period what it deemed to be the chief discoveries and improvements in its subject. Bernal also desired that qualified authors be persuaded to write up their scientific fields at suitable intervals in monographs and textbooks. He suggested as a model for such works the monumental series of German *Handbücher*, which had conscientiously followed the details of scientific advances in every field. Second, Bernal (pp. 298–301) tried to forestall objections based on the dangers arising from the centralization of science. To counter these, he stated that he did not propose putting editorial functions of publication in the hands of a permanent administrative staff. Instead, this administrative staff was to act merely as a link between the writers of papers and the persons who, under the existing system, were editors of scientific journals. Exemplifying this with the same analogy that Davis had used, Bernal (p. 300) wrote, "The publication service would come to be more and more a kind of convenient post office between scientific workers." He responded to potential opposition to his reforms from the scientific societies in the following manner:

> A more serious difficulty would have to be met in the opposition of the existing scientific societies that undertake the bulk of scientific publication. Although in most cases this publication is a serious financial burden to the societies, it gives them in many cases their main *raison d'être*, and the abolition of scientific journals might also be resented for purely sentimental reasons. (p. 30)

In this assessment Bernal was correct, for his proposals were to be almost universally rejected by the scientific societies.

Toward the end of World War II, the British political left took up the cause of a planned scientific state and within the Association of Scientific Workers interest arose in Bernal's proposals for the centralized publication and dissemination of scientific information. The result was what Muddiman (2003, pp. 393–398; 2004, pp. 261–264) has described as a campaign to nationalize scientific information. Bernal obtained the opportunity to advance his ideas by delivering a paper at the 1946 Royal Society Empire Scientific Conference at the same session where Bradford presented his paper. There, Bernal (1948a, p. 698) outlined a set of reforms, whose main purpose was to provide scientific workers with "the maximum of information relevant to their work and the minimum of irrelevant information." The way to do this, he stated, was to organize better the production and distribution of the basic unit of scientific publication—the individual scientific paper. Envisioning reforms at the international level, Bernal proposed that the scientific societies and other publishing bodies in each country jointly establish central agencies, which would use modern methods of reproduction and distribution to disseminate papers and abstracts. In an effort to win the support of the scientific societies, he stipulated that individual papers would be submitted in each country to the appropriate scientific society, which would accept and edit them; the central agency would then print and distribute the papers to members of the society in its country and to corresponding distribution agencies in other countries. Bernal noted that extra copies of papers could be made available to libraries and that the individuality of journals could be retained, as scientists and libraries would have the option of binding papers published by a given society into volumes in accordance with contemporary practice. He suggested that the cost of such a service be borne by general budgets of the participating countries or through subscriptions to UNESCO.

Victory of the Reformist Program

The scientific information session of the 1946 Royal Society Empire Conference recommended that the Royal Society hold another conference specifically dedicated to improving the management of scientific literature. This recommendation was implemented in 1948 with the convening of the Royal Society Scientific Information Conference. By this time the positions of the reformist and revolutionary camps had become fairly clearly defined. As set forth by Bradford (1948a), the main elements of the reformist program were: (1) maintenance of the scientific journal as the basis of scientific communication, (2) maintenance of publication of scientific journals in the hands of scientific societies and other publishing bodies, (3) cooperation among the existing indexing and abstracting services to ensure complete bibliographic coverage of scientific literature, and (4) provision of scientific information through the existing library system improved by the creation of a central document delivery library. In contrast, the revolutionary program advocated by Bernal (1940, 1948a) can be summarized by the following points: (1) replacement of the journal by the individual paper as the basis of scientific communication, (2) centralization of publishing as well as indexing and abstracting functions in a single administrative agency, and (3) transfer of the main distribution function from libraries to the central administrative agency.

Bernal (1948c) had intended to present to the Royal Society Scientific Information Conference a paper outlining a plan for the central distribution of scientific papers. This paper, which was submitted and circulated to the delegates prior to the opening of the conference, was essentially a more detailed elaboration of the proposals and considerations set forth by Bernal and Davis in *The Social Function of Science*. Bernal (1948c) introduced the scheme by describing its aim as the more effective distribution to scientific workers of those papers in which they were most interested. In an attempt to win the support of the scientific societies for the plan, he declared:

> This is to be effected not by radical reorganization of the methods of presentation of scientific papers, but the existing machinery of the scientific societies is to be supplemented by a distributing body functioning as the agency of the societies and in no way interfering with their editorial functions. (p. 253)

Bernal then proposed establishing a small number of central organizations called the National Distributing Authorities (NDA). These NDA were to combine the functions of publishing, abstracting, and distributing scientific papers, so that the whole of scientific publication and distribution would operate together as one unit. Bernal envisioned the establishment of such a system on the national, Commonwealth, and world levels.

In spite of Bernal's attempt to tone down the radicalism of his plan, it provoked a storm of criticism in the British press and infuriated the scientific societies. East (1998, pp. 295–296) has described the campaign in the British press, which was so virulent that it caused Sir Robert Robinson (1948, p. 16), President of the Royal Society, to comment wryly in his opening address that "the writers of certain notices in the newspapers have allowed themselves to dwell, with evident relish, on the prospect of a clash of ideologies and the probable conflict between the planners and those who don't want to be planned." The reactions of the scientific societies and others to Bernal's plan, which were printed in the conference proceedings, were overwhelmingly hostile. Thus, a memorandum signed by sixteen representatives of scientific societies ranging from the Anatomical Society to the Zoological Society of London as well as two journal editors (Memorandum on Section I, 1948, p. 518) deemed Bernal's plan "unacceptable" and declared that "the present system of production of journals by societies ... is not only adequate but essential in principle." Under these conditions Bernal withdrew his plan from consideration of the conference.

Muddiman (2003, p. 396; 2004, p. 263) has interpreted Bernal's withdrawal of his paper in a political sense as marking the defeat of the leftist campaign to nationalize the British scientific information system. However, close examination of the reasons Bernal stated for withdrawing the paper suggests that there may have been very good technical reasons underlying his decision to do so. Bernal's (1948c, p. 258) statement referred to another paper he had submitted for the conference, in which he presented the results of a survey of scientists at leading research institutions in the U.K. on their use of scientific literature (1948b). This survey yielded findings of great import for the further development of the scientific information system. First, it revealed that libraries were the primary source of the scientific journals whose articles scientists either browsed (54 percent) or carefully read (56 percent). In comparison to libraries, personally owned journals and article reprints played relatively small roles. Second, the survey found that scientists were extremely multidisciplinary in their reading habits. Of the journals respondents read, 43 percent were in their own fields, 29 percent classed in related fields, and 28 percent were general journals. Bernal (p. 595) was surprised by this phenomenon. He also (pp. 596–597) noted that the distribution of articles on a given subject across journals conformed to Bradford's Law of Scattering, which stipulated that, although most articles on a topic were in a small number of journals, recourse had to be made to a large number of journals to find all the relevant literature. A third important finding of the survey (p. 599) was that an overwhelming proportion of scientists (76 percent) read reviews and that reviews must form a very important and increasing part of background reading. This finding conformed to the importance Bernal had assigned such literature in *The Social Function of Science*. The first two findings caused Bernal (1948c, p. 258) to make the following explanation in his withdrawal statement:

> It seemed much more profitable to concentrate on improved library systems and on the possibility of copies of papers through libraries rather than from the original publishing body. The scatter of references in journals also

> revealed in the survey, though it could be greatly reduced by good grouping, could never from the very nature of science be altogether eliminated. As a consequence an ideal system would be one which would (a) ensure a wide and rapid spread through all libraries of papers from the few great common journals; (b) use special libraries to distribute copies on request of papers in the middle rank of special journals whose numbers at present only run into hundreds; [and] (c) provide a service from central libraries in conjunction with abstracting agencies and special research institutions of papers from the many thousand small, local, or highly specialized journals.

Muddiman (2003, p. 396; 2004, p. 263) has dismissed this explanation as "a fig leaf" to cover political defeat, but Bernal (1965, p. 456) reiterated these conclusions some 17 years later when he divided physics journals into three categories: the most cited and the most read papers, which should be in every physics laboratory library; the less often cited ones, which could be in the library of the university or large-scale research institute; and, finally, the group containing by far the largest number of journals but not the largest number of papers, which would be sufficiently accessible if they were found in some national science library. It thus seems that, through scientific analysis, Bernal had come to endorse Bradford's reformist position.

In general, the 1948 Royal Society Scientific Information Conference can be considered a victory for Bradford's reformist concepts. Bernal's proposal for the centralized distribution of scientific papers was blocked from consideration and the conference made several recommendations that dovetailed with Bradford's position (Recommendations, 1948). In particular, the conference (p. 199) concluded that more active cooperation among the abstracting agencies would be of benefit to science and made several proposals in that regard. Together with this, it (p. 203) came out in favor of the wider application of the Universal Decimal Classification. The conference (pp. 201–202) also urged further development of information services and special libraries, including increased funding for central scientific libraries such as the Science Museum Library to expand their collections and greater cooperation among libraries to reduce unnecessary duplication and extend access to the world's scientific literature. However, Bernal did win a victory in that the conference (p. 201) recognized the importance of reviews and urged senior scientists to regard the provision of reviews as an important ancillary to the pursuit of new knowledge. In 1956, D. J. Urquhart, who had attended the conference as a representative of the government department charged with promoting scientific research, obtained authorization from his agency to establish the National Lending Library for Science and Technology—the central document delivery library that Bradford had envisaged—which ultimately evolved into today's British Library Document Supply Centre.

Conceptualization of the Science Citation Index

Bernal's Influence on Garfield

Garfield (1982b) was introduced to British developments in information science through Bernal's writings. Upon his graduation from high school, an uncle gave him Bernal's *The Social Function of Science*. He took the book with him to the University of Colorado, where he discussed it intensely with a group of friends, including a Marxist woman who became his wife (Garfield, personal communication, December 5, 2005). The chapter on scientific communication made a great impression on Garfield (1982b, p. 13), who considered it as anticipating the modern revolution in science communication. According to Garfield, Bernal foresaw the need for a reference work that would give scientists access to a large body of past and present scientific literature. Garfield regarded the Scientific Information Institute, which Watson Davis had advocated in the appendix of Bernal's book, as a predecessor to his own company, the Institute for Scientific Information. While working on the Welch Project at Johns Hopkins, Garfield (p. 5) was further heavily influenced by Bernal in the latter's role at the 1946 Royal Society Empire Scientific Conference and the 1948 Royal Society Scientific Information Conference. These conference proceedings became a bible for Garfield as a fledgling investigator. Garfield always considered Bernal something of a "hero figure" (p. 6). In 1962 he sent samples of the *Science Citation Index*, then under development, to Bernal, who agreed in 1964 to serve on its editorial advisory board. Garfield (1976b, p. vii) acknowledged his intellectual debt to Bernal in the dedication to the first *Science Citation Index Journal Citation Reports* (*SCI JCR*): "Dedicated to the memory of the late *John Desmond Bernal* whose insight into the societal origins and impact of science inspired an interest that became a career." As will be discussed, the creation of the *Science Citation Index* itself was partially an attempt to solve the problem posed by Bradford's Law of Scattering for Bernal's plan for the central distribution of scientific papers. Ironically, Garfield himself was to make his greatest theoretical breakthrough by further developing Bradford's law, a breakthrough that was to reveal the complications of employing citation measures such as the impact factor for evaluative purposes.

Leake and Review Articles

The initial impetus for Garfield's conceptualization of the *Science Citation Index* came from Chauncey D. Leake, whom he met in 1951 while working on the Welch Project. Leake, who was the chairman of the project's advisory committee, was a polymath, being both an accomplished scientist and a poet: Over the course of his career, he served terms as president of the American Association for the Advancement of Science, the American Association for the History of Medicine, the History of Science Society, and the American Pharmacology Society. He had an enormous influence on Garfield (1976b, p. vii), who placed his name immediately after Bernal's and right before Lederberg's in the dedication of the first *SCI JCR*. Leake admonished Garfield (1970a, 1974b, 1978) to study review articles and try to understand why they were so important in science. He regularly called Garfield's attention to the failure of conventional abstracting and indexing services to take advantage of the real bibliographic significance of the references in review papers. This advice probably fell on willing ears because, as has been noted, Bernal had also believed in the importance of review papers and had validated this belief in the paper he had presented to the 1948 Royal Society Scientific Information Conference, which, in turn, had officially endorsed this view and urged senior scientists to write such papers. As a result of his study, Garfield came to recognize that the sentences in review articles are implicit indexing statements and that the process of producing a scientific index could be automated by making these sentences the grist for such an index. This idea eventually led to creation of the *SCI*. Garfield continued to hold review articles in high esteem because he remained convinced that it is impossible for artificially intelligent machines to produce the indexing statements that authors contribute when writing critical reviews.

Adair and Legal Citators

Although Leake had pointed Garfield toward the review article as a possible pathway to the automation of indexing scientific literature, the latter still lacked a method for connecting the review article to an index. This was provided by William C. Adair, who had retired as executive vice-president of the Frank Shepard Company. In 1873, this company had begun publishing a system of legal citation indexes called *Shepard's Citations*. U.S. common law is based upon the principle of *stare decisis*, which dictates that all courts must follow precedents laid down by higher courts and each court also follows its own precedents. This makes it essential that lawyers know how a particular case, statute, or other document has been cited by subsequent legal documents. *Shepard's Citations* provided such information. Its legal indexes became popularly known as "citators," and the term "Shepardize" became the common terminology for tracing how a given legal document had been cited in the legal literature. Garfield came into contact with Adair through the First Symposium on Machine Methods in Scientific Documentation sponsored by the Welch Project. As Garfield (1987a, p. 13) has reported, the vice-president of Johns Hopkins University, Lowell J. Reed, delivered the symposium's opening address, in which he stated, "Man is going to be drowned in a flood of paper." This statement was given national coverage by the press, resulting in a flood of letters from all over the country asking for information about the symposium. One letter came to Garfield from Adair, who explained the principle of citators. Garfield (pp. 13–14) has related the moment of discovery thus:

> I didn't know what *Shepard's* was so I went down to the Enoch Pratt Free Library and went into the reference room. I found *Shepard's Citations* and I literally screamed, "Eureka." I had been trying to devise a system around review articles which Chauncey Leake had been pushing me to do. He kept saying, "Review articles are extremely important to scientists. Study them carefully. Find out why they are so successful." I had done a primitive kind of linguistic analysis of reviews. Essentially if you parse a review article, each sentence becomes an indexing statement. ... I was taking the review article and analyzing each sentence and tagging it with the article it had cited. When I saw the *Shepard's Citations* I found the methodology that I needed for linking all these things, for indexing all these cited references that were cited in the review.

Garfield had become the youngest associate editor of *American Documentation*, the journal of the American Documentation Institute founded by Watson Davis. He was so excited by his discovery that he invited Adair (1955) to write an article for *American Documentation* on the possibility of using citators for scientific literature.

First Major Theoretical Paper and the Impact Factor

Garfield put together these elements in an article entitled "Citation Indexes for Science," which appeared in the journal *Science* in 1955. In this article, which Garfield (1987a, p. 16) himself has characterized as "my most important paper," he listed the advantages of a citation index over conventional alphabetical and subject indexes, using *Shepard's Citations* as a tested and successful model. The first advantage, he said, was that its different construction allowed it to bring together material that would never be collated by means of the usual methods of subject indexing. Garfield described a citation index as "an association-of-ideas" index that allowed readers as much leeway as they needed (Garfield, 1955, p. 108). In his opinion, conventional indexes were inadequate because scientists were often concerned with a particular idea rather than a complete concept. The basic problem was to build subject indexes that could anticipate the infinite number of possible approaches that scientists might require in order to bridge the gap between the subject approach of those who created the documents and the subject approach of those who were seeking the information. Garfield stated that the utility of a citation index had to be considered from the viewpoint of the transmission of ideas. He pointed out that scientists could not rely on conventional indexes alone to establish the history of an idea but also had to do much eclectic reading because it was impossible for any one person (the indexer) to anticipate all the thought processes of a user.

Garfield (1955, p. 110) described the advantage of a citation index from this perspective:

> By using authors' references in compiling the citation index, we are in reality utilizing an army of indexers, for every time an author makes a reference he is in effect indexing that work from his point of view. This is especially true of review articles where each statement, with the following reference, resembles an index entry, superimposed upon which is the function of critical appraisal and interpretation.

This perspective, together with the technical discussion in the article about how to code entries for machine processing, shows that Garfield viewed the citation index as a vehicle for using automation to capture more fully the multiplicity of thought processes operative in the scientific information system.

Flexibility in collating materials from different subject areas underlay another advantage that Garfield (1955, p. 109) saw in citation indexes. In his 1955 *Science* article, he alluded briefly to this advantage, stating that the listing of articles that had cited a given article could "provide each scientist with an individual clipping service." This was only the tip of a very deep iceberg, whose full dimensions Garfield (1956, p. 11) revealed one year later:

> An intriguing application of the Citation Index is its potential use in disseminating scientific information as well as for retrieval. Bernal proposed some time ago that a centralized reprint clearing house be established. Each scientist would then regularly receive papers in designated areas of interest. The proposal is excellent in its simplicity. Its execution is not so simple. How would one spell out his interests? By decimal class numbers of index headings or specific compounds? In time any conventional system of classification would break down even if the individual did decide on class numbers or headings. However, a reprint distribution plan based on the principle of the Citation Index could overcome this difficulty. The flow of reprints to each scientist would be reasonable and geared to his individual specialized needs. His changing frame of reference would not periodically disrupt the entire classification scheme.

Thus, Garfield saw in citation indexes a way to circumvent the rock on which Bernal's proposal to nationalize Britain's scientific information system had foundered: Bradford's Law of Scattering. It should be noted that it also offered a way to upgrade the contents-page services on which Garfield was building his business. Garfield (1982b, p. 5) has written that this passage was his first public acknowledgment of Bernal's impact on his career. Shortly after this public acknowledgment, Garfield (1999a, p. 28) first met Bernal in person in Washington, DC at the 1958 International Conference on Scientific Information, which was the successor to the 1948 Royal Society Scientific Information Conference. Both men presented papers there, Garfield's (1959, p. 461) being on the construction of a unified index of science, by which he meant "a single interdisciplinary index to *all* documents, primarily periodical literature in *all* fields of science."

Garfield's (1955) *Science* paper is notable also because it was the first time that he used the term "impact factor," mentioning this term in two places, which are quoted here:

> In effect, the system would provide a complete listing, for the publications covered, of all the original articles that had referred to the article in question. This would clearly be particularly useful in historical research, when one is trying to evaluate the significance of a particular work and its impact on the literature and thinking of the period. Such an "impact factor" may be much more indicative than an absolute count of the number of a scientist's publications. (p. 108)

> Thus, in the case of a highly significant article, the citation index has a quantitative value, for it may help the historian to measure the influence of the article—that is, its "impact factor." (p. 111)

Garfield is clearly speaking here of measuring the "impact factor" of a given scientific paper in terms of the total number of citations to it. He was to change this meaning when he created a measure he called the "impact factor" to determine which journals should be covered by the *SCI*. This term came to be defined as the average number of citations to the papers of a given journal. Nevertheless, the importance Garfield assigned to this measure was determined by his prior intellectual development, which was conditioned by two major determinants in his thinking. First, Garfield considered the paper, not the journal, as the chief vehicle of scientific information. Therefore, he would be inclined to consider important a measure designed to evaluate the significance not of a given journal but of the articles of a given journal. Second, he developed his idea of the citation index on the basis of the review article, which he regarded as the epitome of scientific writing. Therefore, he would favor a measure that ranked the review article higher than the research article. This view of scientific value was at considerable variance with the opinion of many—if not most—scientists. Given that the impact factor was used to rank journals in the *Journal Citation Reports*, there was always inherent in this measure a certain ambiguity as to what actually was being ranked—the journal or the article—and, if the article, then what type of article.

Citation Measures, Bibliometric Laws, and the Journal Citation Reports

Early Statistical Analyses

Definition of the Basic Method of Calculating the Impact Factor

From the very beginning, Garfield combined the development of the *Science Citation Index* with statistical analyses of the citation data being collected for this index. These analyses afforded him great insight into the social structure of science and how it related to the scientific journal system. As part of this process, he created citation measures and formulated bibliometric laws, which he used to improve the journal coverage of the *SCI*. The culmination of this process was the launching of the *Journal Citation Reports* in the mid-1970s as integral parts of the *Science Citation Index* and the *Social Sciences Citation Index*. These *JCR*s were compilations of journal citation measures, of which the two most important qualitative measures were total citations and impact factor.

The first report of such research was written by Garfield and Sher (1963) while the *SCI* was still in its embryonic stage as the Citation Index Project, which, under the sponsorship of the National Institutes of Health and the National Science Foundation, was focused on genetics. This paper is notable as the first one in which the term "impact factor" was defined in the form that the Institute for Scientific Information was to employ it and the reasoning behind this form was stated. Garfield and Sher (p. 199) reported that their concern was not so much the "vital statistics" of scientific publishing; rather, they were "more interested in certain 'impact' factors such as how often a particular paper, author, or journal is cited compared to corresponding average values in a given Citation Index file" (p. 199). They then proceeded to demonstrate that the larger the number of citations taken into account, the more positively and highly skewed became the distribution of the citations to the papers within the data set. Thus, Garfield and Sher calculated that, whereas 95 percent of the papers were cited merely one to three times, certain papers were highly cited, with the range running from the vast bulk of the articles cited only once to an article by Oliver H. Lowry cited 305 times. They also reported that the first year preceding the citing source was the most heavily cited and that half the references were to the eight-year period preceding the source. Garfield and Sher (p. 200) then introduced what they termed "journal impact factor":

> One of the most interesting correlations is the "journal impact factor." In the usual citation count methods ... the importance of a journal is determined by the absolute number of citations to it. The J AM CHEM SOC ranks first on such a list. However, this count is largely a reflection of the fact that more articles are published in this journal than

> most. This approach is not much more sophisticated than ranking the importance of a journal by the quantity of articles published. The first step in obtaining a more meaningful measure of importance is to divide the number of times a journal is cited by the number of articles that journal has published. This linear relationship is valid at least as long as self-citations are not eliminated ... nor multiple citations omitted. ... When this calculation is performed, the J AM CHEM SOC no longer ranks first. In our own citation counts, the PROC NAT ACAD SCI, NATURE, SCIENCE and other journals move towards the top—including some journals which publish many articles and others which do not.

Here, the concept of "impact factor" has undergone an interesting and important evolution from its initial form in Garfield's 1955 *Science* paper. First, the "impact factor" is no longer the total number of citations to a given scientific paper but the mean number of citations per paper in a given journal to be used for explicitly comparative purposes. Second, Garfield explicitly rejects size as a component of "importance." However, the focus is still on the paper instead of the journal, as is evident in a statement by Garfield and Sher (pp. 200–201) that clearly reflects Bernal's influence:

> Librarians and information scientists can organize collections of frequently used papers. It seems utterly foolish to be sending out bound volumes of journals which are being borrowed for a small group of frequently cited articles. The citation data now available makes such a determination possible without intimate knowledge of the subject matter. Thus, we can say with reasonable certainty that any biochemistry librarian would be well advised to have Lowry's article on protein analysis readily available, since it is the most frequently cited paper in the field.

Garfield and Sher then issued a caveat that such information should be used with caution for personnel selection and evaluation such as that done by the Nobel committees, even though many people were interested in this application of citation indexing. They concluded with a declaration that the basic purpose of their project was not "to take a statistical inventory of scientific publication" but "to develop an information system which is economical and which contributes significantly to the process of information discovery—that is, the correlation of scientific observations not obvious to the searcher" (p. 201) through the new insights provided by citation indexes that are not possible through descriptor-oriented systems.

Initial Application of Citations to Evaluating Scientific Research

Immediately after this paper's appearance, Garfield (1963) published another one on the use of citation indexes in sociological and historical research. Here again he stressed the dangers of using citations to evaluate scientists and scientific research, declaring that his purpose in this paper was "to record my forewarning concerning the possible promiscuous and careless use of quantitative citation data for sociological evaluations, including personnel and fellowship selection" (pp. 289–290). In particular, Garfield wanted to disassociate himself from the idea that one could measure the importance of a paper by citation counting, declaring, "*Impact* is not the same as *importance* or *significance*" (p. 290, italics in original). He wrote that citation indexes could be used to facilitate personnel and fellowship evaluation because they synthesize a consensus of scientific opinion needed in a careful appraisal of research. But, he argued, it would be preposterous to conclude blindly that the most cited author deserves a Nobel prize because on this basis Lysenko might have been judged the greatest scientist of the last decade. The reference to Lysenko, the Stalinist scientist who destroyed genetics in the U.S.S.R., possesses a certain amount of irony, because Garfield's mentor Bernal was held in disrepute for openly and steadfastly supporting the Soviet scientist.

Two years later Sher and Garfield (1966) took up the sensitive issue of using citations to evaluate scientists and their work. In this paper, presented at a conference on research program effectiveness sponsored by the U.S. Office of Naval Research, they touted ISI's products as being able to "provide, for administrators, interesting capabilities that can be used in studying, evaluating, and improving the effectiveness of research programs" (p. 137). Cautioning that ISI was on record as being against the promiscuous use of quantitative citation data for research evaluation, Sher and Garfield then declared that "this by no means implies that such evaluations are not possible" (p. 138). The method they selected for demonstrating the validity of the impact factor involved an extreme case—the comparison of the citation rates of Nobelists to the citation rates of average scientists. To do this, Sher and Garfield tallied the citations in the 1961 *SCI* to the persons awarded Nobel Prizes in physics, chemistry, and medicine in 1962 and 1963. They then constructed the "impact factor" of the Nobelists by calculating the number of citations per prize winner, finding that as a group there were thirty times as many citations per average Nobelist as there were per average cited author in the 1961 *SCI* database. Having done this, Sher and Garfield then stated that administrators could use the "individual journal impact factors" (p. 140), which they had just created, to evaluate the effectiveness of specific publications.

The ISI Citation Analysis of Scientific Journals of 1971

Garfield conducted ISI's most important research project on citation analysis as a means to evaluate the significance of scientific journals in 1971. The aims and progress of this research were reported in his essays entitled "Current Comments," which appeared as a regular feature of the weekly *Current Contents*. Garfield (1972a) summarized the results of this project in an article published in *Science* in 1972. The gist of this article was succinctly stated in the following sentence beneath the article's title: "Journals can be ranked by frequency and impact of citations for science policy studies" (p. 471). Garfield began his description of the citation data for the project by stating that the *SCI* had international and multidisciplinary coverage that included the world's most important scientific and technical journals in most disciplines. Data for the study were collected by extracting from ISI's database all the references published during the last quarter of 1969 in the 2,200 journals then covered by the *SCI*. To ensure that this three-month sample was representative of the year as a whole, it was compared to a random sample of every twenty-seventh reference from the approximately four million references collected over all of 1969. The two samples were sufficiently similar to validate the three-month sample, from which three listings were produced. The first listing showed the total citations received by each title and the distribution of these citations over the journal's backfile by year. The second displayed how many times each title was cited by the other titles and the distribution of these citations over the titles' backfile by year. The third was a listing of how many times each journal referenced the other journals and the distribution of these references over the journals' backfile by year. The purpose of these listings, which were to become incorporated into the *JCR*s, was "to map the network of journal information transfer" (p. 472).

Total Citations and Garfield's Law of Concentration

Garfield (1972a) split his discussion of the ISI project into two basic parts: the ranking of journals by total citations and the ranking of journals by impact factor. Concerning the first, the major finding was that, in terms of absolute counts, scientific journal frequency distributions are highly and positively skewed. Pointing out that the majority of all references cited relatively few journals, Garfield (p. 474) described the findings on total citations:

> [The plot of the distribution of citations among cited journals] shows that only 25 journals (little more than 1 percent of *SCI* coverage) are cited in 24 percent of all references; that only 152 journals … are cited in 50 percent of all references; that only 767 journals are cited in 75 percent of all references; and that only 2000 or so journals are cited in 85 percent of all references. In addition, the data … show that only

> 540 journals are cited 1000 or more times a year, and that only 968 journals are cited even 400 times a year.

He also demonstrated that this same pattern held for the distributions of journals by number of articles and number of references to other journals. According to his figures, of the 2,200 journals covered by the *SCI* in 1969, about 500 published approximately 70 percent of all the articles, while a small group of 250 journals provided almost half of the 3.85 million references processed for the *SCI* in 1969. These figures caused Garfield to conclude that "many journals now being published seem to play only a marginal role, if any, in the effective transfer of scientific information" (p. 475).

In analyzing the *SCI* data structure in his *Science* summary article, Garfield (1972a, p. 475) observed, "The predominance of cores of journals is ubiquitous." This observation marked his greatest theoretical breakthrough and provided a solution to the problem that Bradford's Law of Scattering posed for abstracting and indexing services. Garfield (1971) discussed this problem in a "Current Comments" essay written while the ISI project was in progress. Here he noted that Bradford's law dictated that "no matter what the specialty, a relatively small core of journals will account for as much as 90 percent of the significant literature, while attempts to gather 100 percent of it will add journals to the core at an exponential rate" (p. 5). He then stated that "[a]ny abstracting or indexing service that ignores Bradford's law in attempting to realize the myth of complete coverage does so at its great financial peril" and echoed Bradford himself in noting that "no special library can gather the complete literature of its subject *without becoming a general scientific library*" (p. 5). Garfield (p. 5) declared that Bradford's law explains why a multidisciplinary index like the *SCI* is generally more effective than any discipline-oriented index, no matter the specialty:

> At ISI, we are completing a study which has resulted in a generalization of Bradford's law which, in a sense, "unifies" the demonstration of its validity in studies of individual fields. Allow me the eponymic shorthand of calling this unified theory or generalization "Garfield's law of concentration." The name is intended to suggest that, in opposition to scattering, a basic concentration of journals is the common core or nucleus of all fields.

He described his law as postulating for science as a whole what Bradford's law had postulated for a single discipline, declaring that his law held true no matter whether journals were considered as a source of citing articles or as a collection of cited articles. In his *Science* article Garfield (1972a, p. 476) stated his law thus:

> The data reported here demonstrate the predominance of a small group of journals in the citation network. Indeed, the evidence seems so conclusive that I can with confidence generalize Bradford's bibliographical law concerning the concentration and dispersion of the literature of individual disciplines and specialties. Going beyond Bradford's studies, I can say that a combination of the literature of individual disciplines and specialties produces a multidisciplinary core for all of science comprising no more than 1,000 journals. The essential multidisciplinary core could, indeed, be made up of as few as 500 journals.

In his monograph on citation indexing, Garfield (1979a, pp. 21–23, 160) used as a physical analogy for Bradford's law a comet, whose nucleus represents the core journals of a literature and whose tail of debris and gas molecules widening in proportion to the distance from the nucleus depicts the additional journals that sometimes publish material relevant to the subject. According to Garfield, his Law of Concentration postulates that the tail of the literature of one discipline largely consists of the cores of the literatures of other disciplines. He had theoretically solved the problem posed by Bradford's Law of Scattering, but the processes described by his Law of Concentration ensure that any citation analysis seeking to validate the impact factor as a measure of scientific value will be plagued by exogenous subject citations, which render accurate estimates of parameters almost impossible and cause extreme outliers that distort the results.

The Meaning, Methodology, and Consequences of Impact Factor

Garfield's (1972a, pp. 476–477, 478–479; 1972c) discussion of the impact factor in his "Current Comments" essay on this measure and his *Science* article on the 1971 ISI project was more tentative than his discussion of total citations. Indeed, no major conclusions were presented with respect to the impact factor. This was because at the time of this project, Garfield had not fully worked out how to calculate the impact factor, even though he had conceptualized the measure some eight years earlier. Therefore, the discussion of impact factor in both the essay and the article was focused on various ways to construct the impact factor and the reasoning underlying them. These writings provide evidence that Garfield was still keeping a close eye on work in Britain as a source of ideas. For example, in his essay, Garfield (1972c, p. 5) cited a letter in *Nature* written by Sandison (1971), who had been working on the problem of journal obsolescence. Sandison (p. 368) emphasized that in studies of citations or library use the following rule had to be observed:

> To be useful as parameters of the relative value to scientists of groups of volumes, the data must be presented as the

> number of references per available item, and not as the numbers from groups of differing size. The need to correct "obsolescence rates" for the fact that there is much less of the older literature to cite or read is becoming generally recognized.

Garfield (1972c, p. 5) termed Sandison's rule a "recent rediscovery of the impact factor."

However, Garfield (1972a, p. 476) seemed to be particularly influenced by Martyn and Gilchrist (1968), whose work on evaluating British scientific journals he cited in his *Science* article. For their work, Martyn and Gilchrist utilized data from the 1965 *SCI,* which they obtained from ISI. In structuring their data, their first decision was to restrict the sample to the citations made by the journals covered by the 1965 *SCI* to issues of the British journals under evaluation that were published in 1963 and 1964. Their reasoning for this echoed Sandison's logic:

> We decided that our most practical course would be to confine our study to citations made during 1965 to journals published in the two preceding years. It was already known that 26.1 percent of all the 1965 citations were to the literature of 1964 and 1963, so in terms of number of citations this would give us an adequate sample. There is no reason to suppose that, so far as the more important journals are concerned, the ranking we obtained would have been materially altered had our sample covered a greater time span, and by confining ourselves to the two years prior to 1965, we avoided the problem of correcting for cited journal age. (Martyn & Gilchrist, 1968, p. 2)

Then, citing Garfield and Sher's (1963) article on the impact factor, Martyn and Gilchrist (1968) stated that they required, for the journals cited, some measure of their relative sizes in terms of the number of citable items. In this way, adjustments could be made to the number of times each title was cited in order to allow for the increased probability of citations of journals with the greater number of citable items. For this purpose, they deemed it sufficient to count the number of citable items contained in each journal in 1964. Of great import for the future was the fact that they found it difficult to define just what a "citable item" was and had to proceed on an ad hoc basis. To capture various aspects of importance, Martyn and Gilchrist ranked British scientific journals in three different ways: (1) by total number of citations, (2) by ratio of 1964 items cited to 1964 items published, and (3) by number of citations per cited item. There is evidence that Garfield carefully considered Martyn and Gilchrist's work, for he explored their options and incorporated, with important modifications, much of what they did in his construction of the impact factor.

Garfield (1972c) first discussed the findings of the 1971 ISI project on the impact factor in one of his "Current Comments" essays. There, he introduced the impact factor by pointing out the need of librarians for objective criteria in the selection of journals. He also noted that citation frequency is biased in favor of large journals and that ISI had developed the concept of a journal's impact some 10 years earlier. According to Garfield, impact could be measured in a number of ways, of which two had been used by Martyn and Gilchrist. First, one could use the ratio of citations to the number of articles actually cited one or more times, disregarding those that were not cited. Garfield described this as "the *putative* impact factor" (p. 5). Second, a very different ratio could be constructed by calculating the fraction of articles cited. Garfield also observed that one could try to discount the "inbreeding" effect of journal self-citations and that one could compare "*journal utilization* factors" indicating number of different citing journals involved (p. 5, italics in original). The method of calculating the impact factor Garfield chose during the 1971 project became ISI's standard method. This method, some of whose main features had been pioneered and justified by Martyn and Gilchrist, computed the ratio between citations to particular years of a journal and the number of articles published in those years. Garfield noted that this method differs from that used to calculate the putative impact factor, which discounts the negative influence of articles that are never cited, and that such discounting can have a significant effect for certain journals, depending on the definition of an article or citable item. Garfield considered the ratio of citations to sources as providing an overall measure of impact, but he cautioned that the ratio can be skewed by a few super-cited classics unless limited by chronological criteria, observing that a single article cited 500 times has the same effect as 100 articles cited five times. Garfield established these chronological criteria in the same way as Martyn and Gilchrist had, by dividing the number of citations of the source year of the ISI project—1969—to the issues of a given journal published during the two years preceding the source year—1967 and 1968—by the total number of articles published in these issues in 1967 and 1968. Like Martyn and Gilchrist, he justified the two-year limitation by pointing out that about 25 percent of all citations were made to the two-year period prior to the source year chosen. However, years later, he would elaborate further that this figure of 25 percent of all citations to the two years preceding the source year held true basically for the two fields—molecular biology and biochemistry—that were of greatest interest to the users of *Current Contents* and the *SCI* (Garfield, 2003, p. 366). Garfield (1972c, p. 6) declared that, by establishing this two-year limitation, ISI had chosen "a current impact factor" that discounts the effects of most super-cited classics.

At the end of the "Current Comments" essay, Garfield (1972c) published two lists: (1) the 50 most-cited journals ranked in descending order by total citations and (2) the fifty journals highest in impact factor ranked in descending order. His analysis of these lists made two important points. First, the list of 50 journals highest in impact factor was quite different from the list of the 50 most-cited journals: Indeed, only 11 titles appeared on both lists. Second, almost half of the high-impact journals could be categorized as review journals, none of which appeared among the top 50 most-cited list. For example, the title highest in total citations was the *Journal of the American Chemical Society*—the society's main research journal—whereas the title highest in impact factor was *Accounts of Chemical Research*—a review journal of the society. This was to be a consistent feature of all such citation rankings and, given Garfield's intellectual development, the impact factor's capturing of the importance of review articles and journals was bound to influence his evaluation of what was to be considered the most valid citation measure.

Garfield (1972a, pp. 476–477, 478–479, nn. 27–28) developed these considerations concerning the impact factor more fully in his *Science* article on the 1971 ISI project. He began by emphasizing the relationship of citation frequency to journal size, writing that he had very rarely found among the 1,000 most frequently cited journals one that is not also among the 1,000 journals that are most productive in terms of articles published. Garfield (p. 476) carefully distinguished scientific significance from size, declaring that "[i]n view of the relation between size and citation frequency, it would seem desirable to discount the effect of size when using citation data to assess a journal's importance." He then outlined the method by which the effect of size had been discounted for the ISI project:

> We have attempted to do this by calculating a relative impact factor—that is, by dividing the number of times a journal has been cited by the number of articles it has published during some specific period of time. The journal impact factor will thus reflect an average citation rate per published article. (p. 476)

Garfield went on to spell out the details of the calculation of "relative impact factor":

> With the *SCI* data base, it is easy to determine how frequently a journal has been cited within a given period of time, but it is much more difficult to agree on a total-items-published base to which such citation counts can properly be related because the items may have been published at any point in the journal's history. In selecting an items-published base for each journal, I have been guided by the chronological distribution of cited items in each annual edition of the *SCI*. An analysis of this distribution has shown that the typical cited article is most heavily cited during the 2 years after its year of publication. ... Therefore, since my sample consists of references made in 1969, I have taken as the items-published

> base for each journal the number of items it published during 1967 and 1968. To calculate an impact factor for each journal, I divided the number of times 1967 and 1968 articles were cited in 1969 by the number of articles published in 1967 and 1968. (p. 476)

He also pointed out that Martyn and Gilchrist had used a similar method in ranking British journals in their analysis of 1965 *SCI* data.

In presenting the method he chose for calculating journal impact factors, Garfield (1972a, p. 476) wrote that "the development of impact factors that fairly relate the size of a journal during the cited years to its current citation rate is a formidable challenge to statistical analysis." He then analyzed the difficulties of calculating this measure in two lengthy footnotes to the article. In the first footnote, Garfield (p. 478, n. 27) stated that the impact factor discussed in the paper—which he defined as the "average citation rate per published item"—gives some idea of the frequency with which the "average" paper in a particular journal is cited. He then stated that the impact factor is both adversely affected by papers in the journal that are not cited at all and favorably affected by papers with unusual citation frequency. According to Garfield, the influence of uncited and very frequently cited papers can be discounted either by considering the total number of citations in relationship to cited items only (rather than in relation to all published items) or by considering only the number of cited items (rather than total citations) in relation to all published items. It is not without interest that these two ranking methods were used by Martyn and Gilchrist. Although Garfield acknowledged the potential usefulness of these methods for assessing the impact of journals, he noted their impracticality, stating that their derivation would require enormous amounts of computer time.

Garfield's second footnote dealt specifically with the problem of defining what should be in the denominator of the impact factor equation. This problem was to play such a significant role in the subsequent history of the impact factor that it is worth directly quoting Garfield's (1972a, pp. 478–479, n. 28) own words on the matter:

> The problem of selecting an items-published base is further complicated by the variety in the kinds of items published in scientific journals. Many journals publish only full-length reports of original research. Many others publish, in addition, editorials, technical communications, letters, notes, general correspondence, scientific news surveys and notes, book reviews, and so on; all of these are potentially citable items. I have not attempted in this article to limit the definition of items-published to lead articles, original communications, or the like. Even assuming it were possible to construct an acceptable classification that would accommodate all of the different kinds of published material, it would

> have been impossible for me, within the resources available for this article, to have examined individually each of the approximately 600,000 items that I use for the items-published base.

In Garfield's view, it was reasonable to assume that, had such a differentiation among kinds of material been made part of the analysis, the lead articles of such journals as *Science*, *Nature*, *Lancet*, and the *Journal of the American Medical Association* would have had higher impact factors than those calculated for these journals.

Garfield's Constant and Its Implications

In addition to the aforementioned statistical problems with the impact factor, Garfield (1972a) encountered another difficulty that he did not fully discuss in his *Science* article on the 1971 ISI project. Although he referred, somewhat obliquely, to this problem, he did not explicitly analyze its possible consequences. In the article, Garfield (pp. 474–475) observed that, as a result of the highly skewed distribution of citations over journals, the impact of most papers was relatively slight: The average paper is cited only 1.7 times per year, he noted, citing statistics that showed that from 1964 to 1970 the number of *SCI* citations per cited item per year was consistently around 1.7 (p. 478, n. 19). Later, Garfield (1976a) was to call this ratio "Garfield's constant" and puzzle about its significance. In later years, Garfield (1998b, p. 72) observed that, given the growth of scientific literature, the constant was remarkably stable over time, rising from 1.33 to 2.25 over the course of the 50-year period 1945–1995. Garfield's constant suggests that, as estimates of the average citation rates for most articles, impact factors will be extremely low. This was demonstrated by the impact factor statistics Garfield (1972a, p. 477) presented in his *Science* paper on the 1971 ISI project. He ranked the 152 journals with the highest impact factor, and, even over this select group, the impact factor dropped from 29.285 to 1.948—already close to Garfield's constant. What is most interesting but unstated in the *Science* article is how this ranking differed from the ranking of the 50 journals with the highest impact factors that Garfield (1972c, p. 8) presented in his *Current Contents* essay on the same project. There, the impact factor was calculated to only two decimal places, whereas, in the *Science* article, it was calculated to three—a practice that ISI continues to this day. Many years later, Garfield (2000, p. 10) publicly stated the reason for the calculation to three decimal places:

> I keep telling journal people that they should never mention JIF [journal impact factor] beyond the first decimal place. I mean, to quote a JIF like "12.345" is ridiculous. Its JIF is "12.3"; why do you need these two extra digits? It gives

> a false idea of precision. Now you say that ISI does report these numbers. We only do this to give an easy way to separate journals; otherwise we would have many journals with 12.3, and these journals would have to be listed in an alphabetical order. In order to solve this problem, ISI reports the numbers as exactly as they come out. That practice probably should be abandoned.

In a recent conference paper, Garfield (2005, p. 5) has declared, "I myself deplore the quotation of impact factors to three decimal places." In spite of the passion of these comments, Garfield actually understated the problem: Because of the skewed distribution of citations over journals, their impact factors necessarily concentrate at, or beneath, his constant, which, over the course of 50 years, rose merely from 1.33 to 2.25. Given Garfield's own assessment of the imprecision of the measure, this means that there is considerable randomness in impact factor rankings, particularly at the lower end of the range, where the measure cannot meaningfully distinguish one journal from another.

The Creation of the Journal Citation Reports

The Sociological Approach to Journals and Its Intellectual Origins

The 1971 ISI project laid the foundations for the creation of the *Journal Citation Reports*. In fact, the 1969 *SCI* data used in this project were published in-house in 1973 as a preliminary version of the *JCR* in the form of three bound volumes of computer printouts. The first official *Journal Citation Reports* (Garfield, 1976b), published as volume 9 of the 1975 *SCI*, consisted of a bibliometric analysis of references processed for the 1974 *SCI*. It marked an important evolution in Garfield's conceptualization of scientific journals, for, in his introduction to the volume, Garfield adopted a sociological approach toward the analysis of scientific journals. This can be attributed in part to the influence of Bernal, whose Marxism had led him to pioneer the consideration of science in relationship to other aspects of society. Garfield's (p. vii) dedication in the first *JCR* specifically thanks Bernal for the latter's "insight into the societal origins and impact of science." However, Bernal's influence in this respect was reinforced by that of two other men, whose names also appear in Garfield's dedication of the volume: Robert K. Merton and Derek J. de Solla Price, whose names Garfield added to those of Leake and Lederberg in the secondary dedication beneath the main dedication to Bernal. All four men in the secondary dedication are described as ones "whose acumen, criticism, and encouragement as scientists and friends envigored and guided the early research that led to the publication of the *Science Citation Index* and to its use not only for information retrieval but also for the social study of science" (p. vii). Indeed, in 1962, Garfield had sent samples of the *SCI* then under development to Merton and Price as well as to Bernal.

Merton is generally considered the founder of the sociology of science. He began to develop his ideas on the relationship of science to society in the 1930s contemporaneously with Bernal, whose work did not overly impress him. In his review of Bernal's *The Social Function of Science*, Merton (1941, p. 623) described the proposals outlined in the book as "avowedly an attempt at social engineering" and "the most complete and apparently well-grounded program that has appeared since the days of the founders of the Royal Society in the seventeenth century." However, he went on to offer the following assessment:

> The book contributes a great body of substantive materials in a field which has long needed cultivation. It would be ungracious to suggest that the physicist-author has failed to interpret these materials sociologically or has done so in an excessively simplified fashion—*ce n'est pas son métier*. This task may rather be conceived as a challenge to the sociologist of science who has all too often divorced theoretical speculation from empirical investigation. (p. 623)

Merton himself took up the challenge issued in this assessment. His student Jonathan R. Cole (2000, p. 283) credits him with demonstrating in his own work that "the social study of science, beyond the Marxian treatment of it, could yield important results."

Garfield developed an extraordinarily close relationship with Merton, with whom he shared both ethnic and institutional links. Merton was born Meyer R. Schkolnick in 1910 in South Philadelphia to Jewish immigrant parents from East Europe. For most of his career, he taught at Garfield's alma mater, Columbia University. In the mid-1960s Merton's interest in science as an exemplar of sociological theory caused him to establish a Columbia research program in the sociology of science, to which he recruited as its first students Jonathan Cole, Stephen Cole, and Harriet Zuckerman. By his own admission, Merton (1983) was slow to recognize the importance of the *Science Citation Index* for the sociology of science. This was probably because Merton was a qualitative, rather than a quantitative, sociologist. Indeed, Garfield (1977d, p. 6) has remarked, "I have always had the kind of reaction to much of Merton's writing that I associate with a great novelist, not a great scientist." Merton himself did no quantitative citation analyses, but his students were among the pioneers in the use of ISI data for the sociology of science. Employing citation and content analysis, Garfield (1980a) proved not only that Merton's influence extended far beyond his home discipline of sociology, being widely cited throughout the social and natural sciences, but also that he was cited much more for his unique sociological concepts than for his empirical findings. Garfield classified Merton's concepts in the sociology of science into five categories: priority disputes, the Matthew effect, the information structure of science, multiple discoveries, and norms of science.

Of Merton's concepts, the one most important for information science is the Matthew effect. Merton (1968) introduced it in a *Science* article on the psychosocial processes affecting the allocation of awards to scientists for their contributions. According to Merton, the scientific reward system is governed by the Matthew effect—a name he derived from the Gospel according to Matthew (13:12, 25:29), which states (in the King James translation he preferred): "For unto every one that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath." Merton (1968, p. 58) argued that the Matthew effect caused a complex pattern of misallocation of credit for scientific work:

> The Matthew effect consists in the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark.

Merton traced the consequences of the Matthew effect on the scientific communication system, using the highly stratified institutional structure of U.S. academic science to point out that the Matthew effect embodies the principle of cumulative advantage operative in many systems of social stratification and produces the same result: The rich get richer at a rate that makes the poor become relatively poorer. Garfield (1977d, p. 7) has observed, "I would have an opportunity to confirm the Mertonian 'law' called 'the Matthew effect,' by which scientific recognition is bestowed upon one who already has it." That he did so is certified by Merton (1988, pp. 611–612) himself, who, in a follow-up article on the Matthew effect, cited the inequality below as evidence:

> The distributions are even more skewed in the use of scientists' work by their peers, as that use is crudely indexed by the number of citations to it. Much the same distribution has been found in various data sets: typical is Garfield's finding that, for an aggregate of some nineteen million articles published in the physical and biological sciences between 1961 and 1980, 0.3 percent were cited more than one hundred times; another 2.7 percent between twenty-five and one hundred times; and, at the other extreme, some 58 percent of those that were cited at all were cited only once in that twenty-year period. This inequality, you will recognize, is steeper than Pareto-like distributions of income.

From this it can be seen that Merton's Matthew effect bears a strong resemblance to the Marxist doctrine of the concentration of the means of the production in the hands of the capitalist elite and the impoverishment of the masses.

Whereas Merton was the putative father of the sociology of science, Price played a major role in the development of the history of science. The two men differed in the nature of their contributions to their respective disciplines: Merton's were qualitative and conceptual, whereas Price made the greatest impact with his quantitative work. Garfield was related to Price both ethnically and intellectually. Price was born in 1922 in London to parents descended from early 19th-century Jewish immigrant families. Like Bernal, Price came from Sephardic stock, although not through the father but through the mother; around 1950 he adopted his mother's maiden name—de Solla—as his middle name to indicate his Sephardic roots. Price was educated in British state schools, received his first doctorate in experimental physics from the University of London, and obtained his second doctorate in the history of science from Cambridge University. He thus stemmed from the British intellectual tradition that had so influenced Garfield. Like Garfield, Price (1964) was an admirer of Bernal, contributing an article to the Festschrift celebrating the 25th anniversary of the publication of *The Social Function of Science*. It is thus fitting that Price was the first recipient of the John Desmond Bernal Award established by the Society for Social Studies of Science in collaboration with ISI.

As Price (1980) has recorded, he first met Garfield shortly after becoming a professor of the history of science at Yale University in 1959, where he spent the bulk of his academic career. He was serving on the NSF Science Information Council when Garfield applied for support to publish the *SCI*. Although the council rejected the proposal, Price (p. v) writes that he was "inoculated with Citation Fever." During this early period, Price (1963) published his most influential book, *Little Science, Big Science*, which originated as a series of lectures delivered in 1962 at the Brookhaven National Laboratory. The approach of this book was "to deal statistically ... with general problems of the shape and size of science and the ground rules governing growth and behavior of science-in-the-large" (p. viii). Price was particularly concerned with the exponential nature of science in both its growth and distributional patterns. In respect to the first, he analyzed the logistic character of scientific growth, comparing it to similar growth patterns in biology and human society. Concerning the second, he discussed the skewed distributions in the productivity of scientists posited by Lotka's Law and in scientific journal use discovered by D. J. Urquhart in his pioneer study of loans made by London's Science Museum Library in preparation for the establishment of the National Lending Library for Science and Technology. Price (p. 75) described these distributions as "the same Pareto curve as in the distributions of incomes or sizes of cities." *Little Science, Big Science* became an ISI Citation Classic. In commenting upon its evolution, Price (1983, July 18, p. 18) made the following interesting observations about his interests and the intellectual influences upon him:

> Although most of my time was then given to straight history of science, mainly in ancient astronomy and scientific instrumentation, the exponential growth business needled me a lot, and I began to pursue other quantitative researches about science, stimulated much by Robert Merton's writings in the sociology of science, by Eugene Garfield's new book on citation indexing, and by rereading Desmond Bernal's books which had prepared my mind for the initial sensitivity that led me to this field in the first place.

In a bold attempt at a theoretical coup, Price (1976, 1978) attempted to unify all the empirical laws describing the skewed distributions governing science such as Bradford's and Lotka's laws together with the Pareto distribution by deriving a stochastic model he named the Cumulative Advantage Distribution (CAD) after the pioneering work of Merton and the Coles on the social stratification of science. He described Merton's Matthew effect as double-edged in that success increases the probability of success and failure decreases the probability of success. Price regarded Merton's Matthew effect as the stochastic model for the negative binomial distribution, which had been developed on the basis of industrial accidents and smallpox contagion. In contrast, he stated, his CAD was single-edged in that success increases the probability of further success but failure is a non-event that has no effect on subsequent probabilities. In Price's view, the CAD modeled the appropriate probabilistic theory for all the empirical results of citation frequency analysis.

The combined effect of the work of Bernal, Merton, Price, and Garfield was to open the way for analysis of the scientific journal system in terms of social stratification and its mechanisms—a field of study which Karl Marx can justifiably be said to have played a major role in pioneering. This is evident in Garfield's (1976b, p. ix) focus on the social aspects of scientific journals in the preface and introduction to the first *SCI JCR*:

> As, during the years, I and many others used the *SCI* for its planned and advertised purpose of information retrieval, I came to see that I had been advised not only to consider the meaning and usefulness of references and citations, but advised especially to consider their meaning *in a particular type of journal*. The data bank amassed over the years to produce the *SCI* gave me a unique and unprecedented opportunity to look at references and citations not just as tools for information retrieval, but to look at them also as characteristics of the journals they linked. Using the *SCI* data bank, I began to study journals as socio-scientific phenomena as well as communications media.

This conceptualization of scientific journals as sociological entities marked a distinct evolution in Garfield's thinking away from Bernal's idea of scientific journals as inefficient bundles of articles that needed to be broken down into more convenient packages for delivery to individual scientists. Garfield expressed the hope that the *JCR* would prove uniquely useful in exploring the relatively new field of the sociology of science.

Garfield (1976b) opened his introduction to the new *JCR* by elaborating upon the sociological aspects of scientific journals. He pointed out that a citation index is based upon the principle that there is some meaningful relationship between one paper and another that cites it and, thus, between the work of the two authors or two groups of authors who published the papers. Garfield (p. 1) then observed that an author's or a paper's frequency of citation has been found to correlate well with professional standing and argued that the same principle could be applied to the evaluation of journals: "The more frequently a journal's articles are cited, the more the world's scientific community implies that it finds the journal to be a carrier of useful information." He described the *JCR* as extending the use of citation analysis to examine the relationships among journals rather than among articles and their authors. With this conceptualization, Garfield made journals an integral part of the social stratification system of science.

Having thus defined the basic purpose of the *JCR*, Garfield (1976b) issued a number of caveats against its indiscriminate use in the evaluation of journals. He noted that citation frequency was biased both in favor of journals publishing original research findings and the indispensable summary of research findings provided by reviews, and against those published for other purposes, such as conveying scientific news. Echoing Merton's concerns about the misallocation of scientific credit, Garfield stated that citation frequency is sometimes a function of variables other than scientific merit. Among such variables, he listed first an author's reputation, followed by others such as the controversial character of subject matter, journal circulation and price, indexing coverage, society memberships, and library holdings. Garfield (p. 3) warned against the technical difficulties involved in utilizing the *JCR*:

> We have, thus, in compiling the *JCR* refrained from combining journal counts on the basis of "lineage," even when it is clearly definable. Except where a title change has been so minor ... that it neither affects the title's position in a catalog listing nor requires additional or different entries, the *JCR* does not combine counts for related journals (replacements, supersedents, continuations, descendants, etc.). Nor does it combine counts for "sections" of "the same journal." [The] *JCR* leaves it to the user to decide whether or not his purpose recommends that counts be combined in such cases.

Thus, utilization of the *JCR* requires intimate knowledge of the intricacies of serials cataloging—knowledge often absent in researchers and others who have drawn upon *JCR* data for their various purposes. Under the heading "Caution!," Garfield emphasized the need to be aware of the different citation patterns of different disciplines in comparing journals. Toward the end of the introduction, Garfield (p. 5) emphatically declared that "the *JCR* cannot be used alone in evaluating a journal's performance."

JCR Structure and Citation Measures: Final Formulation of the Impact Factor

The structure of the first *JCR* consisted of three packages. The first of these was the Journal Ranking Package, which began with an alphabetical list of journals cited in 1974 and then ranked these journals in descending order by the following measures: total citations in 1974, impact factor in 1974, immediacy index in 1974, source items published in 1974, and 1974 citations of their 1972 and 1973 articles. The Journal Ranking Package was followed by a Citing Journal Package, which listed citing journals alphabetically with subentry listings of the journals they cited in 1974, and a Cited Journal Package, which listed cited journals alphabetically with subentry listings of the journals that cited them in 1974. With certain important modifications, this structure was to remain essentially the same through the successive print and microform editions of the *JCR*s until electronic versions with new capabilities appeared in the mid-1990s on both CD-ROM and the World Wide Web. Prior to the appearance of the electronic versions, the rankings published in the *JCR*s were rigid and could be manipulated only with great difficulty.

In order to understand fully the citation frequency rankings of the *JCR*, it is necessary to have a clear grasp of the definitions of three key *JCR* terms. These (Garfield, 1976b) are quoted below. The first two definitions involve calculations; the *JCR* data for the *Journal of the American Chemical Society* as well as the method of calculation are placed beneath them to aid in their understanding.

> Immediacy Index. A measure of how quickly the "average cited article" in a particular journal is cited. A journal's immediacy index considers citations made during the year in which the cited items were published. Thus, the 1979 immediacy index of journal *X* would be calculated by dividing the number of all journals' 1979 citations of items it published in 1979 by the *total number* of source items it published in 1979. It should be obvious that an article published early in the year has a better chance of being cited than one published later in the year. As a result, journals published weekly and monthly will theoretically have an advantage, as regards

immediacy, over journals published quarterly and semi-annually. (p. 6)

CITATIONS IN 1974 TO 1974 SOURCE ITEMS	SOURCE ITEMS IN 1974	IMMEDIACY INDEX
J AM CHEM SOC		
1835	1432	1.281

[CALCULATION: 1835/1432 = 1.281]

Impact Factor. A measure of the frequency with which the "average cited article" in a journal has been cited in a particular year. The *JCR* impact factor is basically a ratio between citations and citable items published. Thus, the 1979 impact factor of journal *X* would be calculated by dividing the number of all the *SCI* source journals' 1979 citations of articles journal *X* published in 1977 and 1978 by the *total number* of source items it published in 1977 and 1978. There are other ways of calculating journal impact. ...

The impact factor is useful in evaluating the significance of absolute citation frequencies. It tends to discount the advantage of large journals over small ones, of frequently issued journals over less frequently issued ones (weeklies vs. quarterlies or annuals), of older journals over newer journals. In each such case the first is likely to produce or have produced a larger citable corpus than the second. All things being equal, the larger that corpus, the more often a journal will be cited. The impact factor allows some qualification of quantitative data. The qualification is algorithmic and objective, but nonetheless useful in journal evaluation. (pp. 6–7)

CITATIONS IN 1974 TO 1973 1972 72+73	SOURCE ITEMS IN 1973 1972 72+73	IMPACT FACTOR
J AM CHEM SOC		
7855 9233 17088	1776 2123 3899	4.383

[CALCULATION: 17088/3899 = 4.383]

Source Item. Called also source document or source article, a source item is an item published in one of the source journals processed for the *Science Citation Index* (*SCI*). Source items may be original substantive articles, editorials, letters, technical notes, correction notes, meeting reports, reviews, and so

> forth. From the references provided by a source item, citations are extracted to prepare the *Citation Index* of the *SCI*.
>
> Some types of source items (e.g., news items, non-scientific and non-technical correspondence) do not by their very nature invite citation in the references of scientific reports. Such source items are excluded from source-item counts in [the] compilation of the *JCR*. In the *JCR* only original articles, technical notes, and review articles are counted as source items, except in the case of [certain] journals, whose meeting abstracts are admitted as source items in impact-factor and immediacy-index calculation. (pp. 7–8)

From these definitions it can be seen that the *JCR* ranked journals by citation frequency in two different ways: the total citation frequency of journals and the citation frequency of the "average cited article" (p. 6). Each of these was, in turn, done in two versions. The total citation rankings were determined by calculations using both the total citations to all the issues of a given journal and the total citations only to the issues published in the two years preceding the processing year. The latter method—that of Martyn and Gilchrist (1968)—yielded the numerator of the impact factor. With respect to rankings by the "average cited article," the impact factor was calculated by dividing the citations of the processing year to the issues of the journals published in the two years preceding the processing year by the number of "source items" published during these two preceding years, whereas the immediacy index was calculated by adding the citations of the processing year to the issues of the processing year and dividing by the "source items" of the processing year (Garfield, 1976b, p. 7). This raises two contentious issues. First, both cases involved a dichotomy between the numerator and the denominator. The numerator was comprised of citations from all types of items published but the denominator included only the number of source items defined as "citable." This technically converted both the impact factor and the immediacy index from average citations per source item into ratios of citations to "citable" source item. Second, as a measure of the rapidity with which items are cited after publication, the immediacy index affected the impact factor due to the latter's two-year limitation. But scientific disciplines differ in this respect—a fact that severely complicates cross-disciplinary comparisons by means of the impact factor.

Garfield (1976d) introduced the new *JCR* to the broader scientific community in an article published in *Nature*, in which he analyzed its data and compared the findings to those of the 1971 ISI project using 1969 *SCI* data. This analysis was instrumental in clarifying the structure of the impact factor. Garfield stated that the new *JCR* had two measures of journal significance—total citations and impact—and discussed their characteristics. In his treatment of total citations, he presented data in a table, which he divided into two sections (pp. 610–611). The first section arrayed the 206 journals ranking highest in 1974 total citations to all journal issues; the second included an additional 78 journals that ranked high in total citations to their 1972 and 1973 issues only—actually a ranking of journals by the numerator of the impact factor. The total citation rankings given in the *Nature* article exhibited significant deviations from those of the *JCR* because the table in *Nature* combined the counts for sections and retitled continuations. Analyzing these two total citation rankings, Garfield found that the section based upon citation counts to the 1972 and 1973 issues only had a large proportion of titles (63 percent) that had begun publication in the 1960s and 1970s. He declared that this second section was a necessary supplement to the first, based on counts to all issues, because "the journals have high current citation but lack historical mass to push them up into the top of a list ranked by total citations" (p. 609). Thus, this analysis revealed the effect of equalizing journals in terms of time. Comparing the journals that ranked highest in total citations to all issues in 1974 to the same category of journals in 1969, Garfield found what he considered a remarkable stability: Of the 206 journals most cited in 1969, 169 remained among the top 206 in 1974. Some 15 years later, Garfield (1991) confirmed this stability when he published two lists of journals: The 50 journals found to be most highly cited in his seminal study of 1969. *SCI* citations and the 50 titles most highly cited in the 1989 *SCI JCR*. Of the 50 titles on the 1969 list, 32 were also on the list in 1989.

Garfield (1976d, p. 613) also discussed the impact factor in his *Nature* article, using a table that ranked journals with impact factors greater than 2. He demonstrated the significance of these journals by pointing out that, out of 2,443 journals, only 150 had an impact factor above 3 and that the mean impact factor for all journals was 1.015. In presenting the impact factor data, Garfield (p. 613) issued the following warning about the denominator of the impact factor calculation:

> In using the data presented here, one should be aware that we revised our definition of "source items" used to calculate impacts. In 1969 we included as source items much material (editorials, non-scientific and non-technical correspondence, news notes, and so on) that does not by its very nature invite citation in scientific and technical reports. This policy worked to the disadvantage of some major journals. Our redefinition accounts in part for the changed impact in 1974 of journals like *Nature*, *Science*, *Lancet*, *Journal of the American Medical Association*, and *British Medical Journal*.

Garfield thus let it be known that ISI had begun to do what Martyn and Gilchrist (1968) found very difficult and Garfield (1972a, p. 479, n. 28) himself doubted the very possibility of doing—defining what a "citable item" is and constructing "an acceptable classification that would accommodate all of the different kinds of published material." Like the total citations table, the impact factor table in the *Nature* article ranked 284 journals and was divided into two sections (Garfield, 1976d, pp. 612–613). The first section listed 206 high-impact journals with the exclusion of review journals, whereas the second section ranked 78 high-impact review journals. Garfield found that the review journals had generally a higher impact factor, stating that this clearly showed the importance of review journals and confirmed previous ISI studies. He then made the following announcement concerning review journals: "Their extraordinary impact, along with a surge in the number of review-type articles and publications, led to ISI's decision to publish *Index to Scientific Reviews*" (p. 613).

Having analyzed the nature of total citations and impact factor as measures of journal significance, Garfield (1976d) went on to demonstrate the need for set definition in citation analysis. He did this with a table divided into three sections that listed the journals ranking highest in total citations and impact factor for three disciplines: astronomy/astrophysics, botany, and mathematics (p. 614). Garfield stated that the differences in average impact and citation between the three illustrative groups indicated that comparisons between journals in different subject areas might be invidious. He then summed up the causation of different citation patterns among disciplines and their effect on the impact factor:

> Variation from field to field is determined by the interplay of several factors. Perhaps the most important is the average number of references per paper in the field. In general, mathematicians cite less than half as many papers as do biochemists. Engineers on the other hand cite books as heavily as journals, as do social scientists. Furthermore, calculation of impact based on 1972 and 1973 publications is bound to affect the impact of journals in a field like mathematics, where citation of older literature is far more common than in others. Thus, the impact of mathematics journals would be higher if calculated on the basis of 1970 and 1971 publications. (p. 609)

In this way, Garfield pointed out the problem of selecting a fair time basis for calculating impact factor.

Modifications of the *JCR*s Affecting the Impact Factor

Over the next few years the *JCR*s underwent modifications that affected both the calculation of the impact factor and its evaluation. For the 1977 edition, with the introduction of *Social Sciences Citation Index JCR,* the database from which the *JCR*s were derived was expanded to include references processed for the *SCI* and the *SSCI*. The subject scope of the *JCR* database was further expanded for the 1979 edition by the inclusion of the references being processed for the newly established *Arts & Humanities Citation Index*. This added to the subject complexity of the *JCR*s, as Garfield (1980b, p. 8A) noted in the 1979 *SCI JCR*:

> The use of the combined data bases eliminates the often shadowy boundaries between the sciences and social sciences. Some journals are covered by *SCI* and *SSCI*; others are covered only by *SCI*, but may also be cited extensively by journals in the *SSCI* and *A&HCI* data base. To present a more accurate picture of these journals' citation rates and to include a complete list of the journals which cite them most often, citation data from *SSCI*-only and *A&HCI* journals are incorporated in the *SCI JCR*.

An equivalent announcement was put in the 1979 *SSCI JCR*. Given the interdisciplinary citation processes described by Garfield's Law of Concentration, the combination of the databases meant that the *JCR* citation measures were capturing the significance of journals not only within their specific fields but also for all branches of human knowledge. However, with the 1978 editions of both the *SCI* and *SSCI JCR*s, ISI began to compensate for the increased subject complexity of the *JCR* data by including a section listing the source journals within narrowly defined subject categories. Together with these subject changes, ISI introduced new measures and sections that aided in the clarification of citation patterns and their effect on the impact factor.

With the 1977 edition of the *SCI* and the newly created *SSCI JCR*s, ISI added a section called "Source Data Listing" that classified source items into non-review and review articles, giving the total number of items and references for each. These totals were used to calculate references-to-source-items ratios for each type and the two types combined. However, even this simple classification of source items into "non-review articles" and "review articles" proved difficult. Garfield (1982a, p. 5) later described the word "review" as "one of the more ambiguous terms in scholarship." Defining a review article in science as "an annotated summary or critical digest of the literature of a given topic," he declared that "even in science the line of demarcation is hazy." Garfield (1987b) discussed various methods of classifying review journals and articles, citing research that classified review journals into eight different types and showed that review articles could run the gamut from little more than bibliographies to highly subjective evaluations of material within a field in many bibliographic forms.

The 1978 *JCR* editions provided vastly improved measurement of the chronological patterns of citations with the introduction of a part entitled "Journal Half-Life Package," which was divided into two sections. The first cumulated the percentage of citations to each cited journal over a ten-year period. These cumulated percentages formed the basis for calculating for each cited journal a measure called "half-life," which Garfield (1979c, p. 10A) defined as "the number of journal publication years from the current year back whose articles have accounted for 50 percent of the total citations received in a given year." The second section of the "Journal Half-Life Package" ranked journals in ascending order by this measure.

The Enshrinement of the Impact Factor as ISI's Chief Measure of Journal Significance

The 1979 editions of the *JCR*s introduced a new feature that enshrined the impact factor as ISI's chief measure of journal significance. This new feature integrated many of the modifications already mentioned; Garfield (1980b, p. 1A) described it in these words:

> A refinement in the 1979 *JCR* enables researchers to access journal ranking information according to journals' subject fields. I've often stressed the importance of limiting comparisons between journals to those in the same field. The journal literature varies in importance as a means of disseminating information in different fields, and citation practices vary from field to field. Thus, a comparison of the citation data of a micro-biology journal and a journal of highway engineering would be meaningless. The researcher can avoid such "apples and oranges" comparisons by turning to Section 8 of the *JCR's Journal Ranking Package*. This new section shows journals grouped by subject category, then ranked by impact factor. Journal half-life is also provided.

No reasons were stated for the inclusion of journal half-life, but it does provide a means of comparing between journals and subject fields, whether or not the citations to journals were concentrated within the two-year limit on which the calculation of the impact factor was based. Until the appearance of the electronic versions of the *JCR*s in the mid-1990s, these impact factor rankings were the only easily and immediately accessible citation rankings of journals within narrowly defined subject fields. This must be considered the underlying reason the impact factor assumed such overwhelming importance in scientific evaluations.

The question now arises as to why Garfield accorded such prominence to the impact factor as a measure of journal significance. There are manifold reasons—some of them traceable to Bernal. First, it is possible to hypothesize that, as a youth, Garfield was indoctrinated into leftist political thinking, which tends to regard equality almost as a moral imperative. The impact factor levels the playing field, putting the measurement of all journals—big and small, old and new—on the same basis: Thus, the ostensible fairness of the impact factor may have played a role in its widespread adoption as the standard measure of journal significance. Second, quite apart from politics, Garfield was heavily influenced by Bernal's idea that the article, not the journal as a whole, is the prime vehicle of scientific communication. The impact factor is definitely an attempt to measure the significance of a journal not globally but by the importance of its individual articles. For this reason, Garfield (1972b, pp. 5–6) considered the impact factor to be helpful to scientists in selecting a journal in which to publish, observing that "[p]ublishing in a high-prestige journal which has a large circulation may have less impact on the scientific community than publication in a journal with smaller circulation but high impact." Third, for various reasons—some of which will become clearer later—Garfield was focused on measuring current, not historical, significance. As has been discussed, one of the reasons Garfield developed the impact factor was to discount the effect of older papers that had become citation classics. From the perspective of current significance, such citation classics can be highly distortional. For example, Garfield (1996) pointed out that Lowry's classic 1951 protein determination paper, which fascinated Garfield over his entire career, alone accounted for about 7,000 (3 percent) of the 265,000 citations in 1994 to the *Journal of Biological Chemistry*, which was the journal most highly ranked in total citations that year. Fourth, when journals are measured by impact factor, review journals are the most highly ranked ones. Review articles had always held a special meaning for Garfield. It has already been shown that Bernal stressed the importance of review literature in *The Social Function of Science*, which so influenced Garfield in his youth, and had uncovered the importance of review materials for scientists in the research he presented to the 1948 Royal Society Scientific Information Conference. Moreover, under Leake's guidance, Garfield had developed the entire concept of the *Science Citation Index* as a result of his analysis of review articles. Throughout his career, Garfield (1974a, 1974b, 1977c, 1982a, 1987b, 1987c) maintained his high esteem for scientific reviews, even suggesting that there be a counterpart for science of the law reviews published by American law schools and that reviewing be established as a scientific profession. In describing his attitude toward review literature, Garfield (1987b, p. 5) wrote:

> The "culture" of reviewing the literature is so fundamental to my own professional life that I too may forget that in comparison with research discoveries one reads about in the press, and for which Nobel Prizes are awarded, reviewing may seem to the uninitiated to be a relatively humdrum topic.
>
> But it is precisely this mistaken notion that I want to dispel. It is not an accident that so many of our greatest scientists have used, created, and contributed to the review literature. Like an important opinion rendered by the chief

justice of the Supreme Court, reviews can have great value and influence.

Review literature was the subject of a number of analyses by researchers at ISI, which joined with Annual Reviews Inc. in 1979 to sponsor the National Academy of Sciences Award for Excellence in Scientific Reviewing. However, Garfield (1987b) presented evidence that his opinion on the importance of review literature was at variance with that of many scientists, reporting that many authors of ISI Citation Classics believed that review articles should not be automatically granted this award and that some felt that their review articles should not be judged on the same criteria as their articles reporting original research. And, finally—a reason never stated by Garfield but obvious to anybody with a knowledge of serials cataloging who has ever worked with *JCR* data—the impact factor's restriction of the field of observation to the two most recent years minimizes the technical problems and subjective judgments occasioned by the bibliographic instability of journals, which have a tendency to change titles, continue or supersede each other, and divide into parts that subsequently do or do not recombine.

Utilization of Citation Measures in Determining ISI Journal Coverage

Development of Selection Criteria

Garfield (1972d, 1979b, 1985a) has reported that, in its early days, ISI had no objective criteria or formal policies for selecting journals to cover. Journal selection was by necessity highly subjective for most disciplines and so Garfield compiled a basic list in the same manner as any experienced scientific publisher or librarian. Although there were classic studies of scientific serials and library surveys, ISI did not need them to know that publications such as *Science*, *Nature*, *New England Journal of Medicine*, *Proceedings of the National Academy of Sciences-US*, were among the most important scientific journals. Garfield (1985a, p. 6) has stated that there was no question that journals of this caliber should be covered because these intuitive choices were backed up by all sorts of objective criteria. He has also pointed out that this same basic group of hard-core journals continued to maintain high quality year after year and had an uncanny way of surviving and growing. Journal selection did not become a major problem until economics permitted ISI to extend its coverage beyond this overwhelmingly important core.

With the growth of *SCI* journal coverage during the 1960s from about 600 to over 2,000 titles, ISI developed the capability of experimenting with citation measures of journal significance. One of the first of these experiments was with the impact factor. In his report Garfield (1970b) cited Martyn and Gilchrist's (1968) evaluation of British scientific journals as a study that quantified the impact measure that Garfield and

Sher (1963) had initially proposed. Defining the impact factor as "the mean number of citations to a journal's articles by papers subsequently published," he stated that, "though perhaps a somewhat crude measure, [it] does reveal some interesting characteristics of scientists, as well as of journals" (Garfield, 1970b, p. 5). Garfield (p. 5) urged caution in using measures such as the impact factor for comparative purposes, specifically warning about "the 'immediacy factor'—the 'bunching,' or more frequent citation, of recent papers relative to earlier ones." For this early experiment, ISI constructed a prototype of impact factor by selecting the 1965 *SCI* references to articles published in 1963, sorting the references by cited journal, and then ranking the journals by "impact factor." Garfield found the results revealing. The average impact factor was only 1.9, but the relatively new *Journal of Molecular Biology*, which began publication in 1961, had an impact factor of 7, whereas "long established 'significant' journals like the *Journal of Biological Chemistry*" (p. 5) had an impact factor of about 3. This caused Garfield (p. 5) to observe:

> It is generally recognized that molecular biology is a highly focussed [*sic*], "hot" field; so, the question arises whether the rapidity of developments in such a field tends to distort the impact factor. Rapid citation of one paper by others in a fast-moving field presents a picture of impact different from that of journals publishing articles that are cited in later years not only frequently but in a wide range of journals.

He then declared that "whether a journal is cited because it is in a rapidly developing field, or because it publishes articles of long-term impact, the journal is significant" (p. 6). At the end of the report, Garfield (p. 6) observed that most journals containing a large number of source items also prove to be significant, thus making the connection between size and significance, and concluded with the following recommendation:

> Hopefully, some of the less significant journals would take steps to improve their quality or to merge with other small journals to form larger ones, which ... tend to acquire a special significance, due possibly to greater exposure.

The 1971 project analyzing 1969 *SCI* data was pivotal in ISI's development of criteria for selecting which journals to cover with its indexes, and Garfield (1972d, 1973a, 1973b, 1973c) devoted a number of "Current Comments" essays to this aspect of the project. The fact that, in these essays, he initially referred to the *Journal Citation Reports* as the *Journal Citation Index* serves as a sign that they were being written as the discoveries were made. In the earliest of these essays Garfield (1972d, p. 5) made a very interesting distinction between the roles of total citations—or "absolute citation frequency" in his terminology—and the impact factor in selecting journals for coverage. With respect to the former, Garfield (p. 5) observed that "absolute citation frequency is not sufficient for the task of journal selection except perhaps to establish 'core' journal collections." He then noted the difficulty of new and small journals to make it into the top ranks of the most-cited journals, especially given that "an almost immutable 'constant' citation rate will obtain and that the average article in *SCI* is cited about 1.67 times a year" (p. 5). Garfield then went on to state that, because absolute citation frequency does not tell the entire story, ISI had developed the impact factor to discount the advantage that large, established journals have in absolute citation counts. Here we see in embryonic form the relationship of total citations and impact factor to Garfield's Law of Concentration. Total citations identify the core journals of each discipline that form the small multidisciplinary core dominating science as a whole, whereas the impact factor is a measure enabling one to select in a more rational fashion which journals should be selected for coverage from the long tails of the Bradfordian citation comets.

Garfield (1973a) joyously hailed the publication of his 1972 *Science* article on the 1971 ISI project with a "Current Comments" essay in which he summarized some of the project's key findings. Of greatest interest in this essay are his conclusions on the relationship of size to significance and his justification of the impact factor. As regards the former, Garfield (p. 5) wrote:

> The [*Science*] paper deals primarily with the use of our *Journal Citation Index* data bank to determine the frequency with which scientific and technical journals are cited in the journal literature. It shows that a "large" journal that publishes many articles is, as a rule, more frequently cited than a journal that publishes fewer articles. In addition, however, through development of "impact factors," it shows that articles in about half of these most-cited journals are cited less frequently than articles in smaller, less-cited journals.

Concerning the latter, Garfield stated that the project provided justification for ISI's coverage of review journals, despite the extreme expense required to process their many references, because it found that they ranked high in impact factor even though they were not often among the most-cited titles. He also emphasized that objective criteria alone do not solve the journal selection problem. In a follow-up essay, Garfield (1973b) declared that, although citation studies enable the identification of the obviously important journals, ISI also had to take into account other factors in evaluating the less-important titles. He pointed out that, because biochemists composed a major segment of ISI readership, a new journal in biochemistry or molecular biology would have higher priority than a journal in horticulture or a journal of a local medical society.

At the beginning of his most comprehensive essay on the implications of the 1971 project for ISI journal coverage, Garfield (1973c) announced that the company was using citation analysis to evaluate journals but then immediately reiterated his caution that neither men nor journals can be judged on the basis of citation analysis alone. He stated that citation analysis adds objectivity to the evaluation process and listed three different ways of using citation data to evaluate journals: (1) the frequency with which the journal is cited, (2) the frequency with which the average article in the journal is cited, in other words, the impact factor, and (3) the frequency with which a given journal publishes articles that become citation superstars. Garfield concentrated his attention in this essay on the second and third methods. To illustrate the third method, he used Lowry's paper on protein determination as an example of a citation superstar, noting that in 1969 only about 150 journals were cited as frequently as this single paper. Garfield then reported that he had recently examined a list of the 1,000 papers most frequently cited during the past decade. He had found that only about 200 journals had accounted for these 1,000 articles, of which half had been published in only fifteen journals. In Garfield's view, although a citation blockbuster like Lowry's paper was atypical of most scientific papers, it was not so atypical of other papers published in the *Journal of Biological Chemistry*. Summing up his main point on this matter, Garfield (p. 6) wrote that "it is remarkable that of the 1000 or more most heavily cited articles in the literature, not one appeared in an 'obscure' journal." This concentration of high-quality articles in a few high-prestige and high-visibility journals was also found by Merton's student, Stephen Cole (2000), who regarded it as a sign of the effectiveness of the scientific journal system in concentrating attention on key research and therefore making this journal system capable of playing a role in scientific evaluation. However, Cole did note that, due to the inherent difficulty of predicting quality, most of the articles even in the most prestigious journals were also of minor significance, thus hampering the journal system as a tool for scientific evaluation.

Garfield (1976c) stated that one of the most surprising discoveries of the 1971 ISI project was the relatively low impact of articles published in most journals, including journals that seem almost universally accepted as preeminent. As proof of this, he observed that there were only about 150 journals with impact factors greater than 2 and fewer than 500 journals with impact factors above 1. He then went on to translate the statement in his 1972 *Science* article that developing a fair way of calculating impact factors was "a formidable challenge to statistical analysis" (Garfield, 1972a, p. 476) in the following manner: "In less stilted prose, it offers ample opportunity for statistical carping" (Garfield, 1973c, p. 5). He also cautioned that in dealing with journals of small size and low impact—in other words, with the vast majority of journals—considerations other than citation counts had to be taken into account in justifying coverage.

In this and in two subsequent essays dedicated to coverage by *Current Contents*, Garfield (1973c, 1979b, 1985a) discussed these other considerations in detail. The most controversial ones were those he categorized under the heading "Geopolitical Representation," which consisted of two elements: language and geographical representation. Concerning language, he stated that, given two equivalent journals, ISI would choose the one that published articles in English on the grounds that most of its readership could handle English, but few could deal with Slavic or Asian languages. Failing this, preference would be given those journals that contained informative abstracts and summaries in English as it would be absurd for scientists to expect colleagues abroad to be able to read all of the exotic languages in which original data could be reported. Garfield (1976c, 1977a, 1977b) was fairly adamant on this point and he caused a scandal in France by publishing there an article on the provinciality of French science, in which he used citation analysis to prove that French science was declining in influence because of the refusal of French scientists to recognize that French was no longer a significant international language. Garfield declared English to be the true international language of science, recommending that French scientists publish in English and that French-language journals also be published in an English edition or at least have contents pages and summaries in English. This "*nouveau défi américain*" appears to have had an effect, for in a follow-up study Garfield (1988) found that French researchers were publishing much more in English and citing English (international) literature more extensively. Garfield (2002) offered much the same advice to the Germans.

In outlining ISI criteria for selecting which foreign journals to cover, Garfield described language as being closely interconnected with geographic representation. Thus, given a choice of two journals in the same subject area, ISI would choose the one with international representation. Garfield held that the fundamental issue for scientists in small countries like Finland was that research of international significance should be submitted to international journals. He reported that his own research showed that, of the 23 most-cited articles by Third World authors, none had been published in Third World journals, 13 had been published in U.S. journals, seven in the U.K., two in the Netherlands, and one in Germany (Garfield, 1983, p. 120). This research also revealed that most of the citations to these articles came from scientists in developed countries. In a paper on the implications of the quantitative analysis of scientific literature for science policymaking in Latin America and the Caribbean, Garfield (1995, p. 88) gave the following characterization of *SCI* journal coverage:

> Currently, ISI indexes about 3,300 journals for the *SCI*, all peer-reviewed and internationally influential. This selective

> coverage is not merely a matter of economics; it reflects a virtual law of nature with regard to the use of the journal literature. Just a handful of journals in any field account for the lion's share of the really important, frequently read, and frequently cited journals.
>
> Thus, what the ISI database represents is the set of journals that constitute the internationally influential literature. It does not represent the science of any given country or region as a whole, but it does represent the portion of research that is published within and cited within the internationally elite literature. Beyond that, it generally represents the best science performed in any nation.

Garfield (1997a) returned to the question of the internationality of ISI journal coverage in an article examining whether the *SCI* discriminated against Third World journals. This article is most interesting not only for Garfield's explanation of what he meant by *SCI*'s selective journal coverage being reflective of "a virtual law of nature" (Garfield, 1995, p. 88), but also for his description of the role of impact factor in ISI's determination of which journal to cover. With respect to the first point, Garfield identified the "virtual law of nature" with Bradford's Law of Scattering and his own Law of Concentration, pointing out that such laws are operative in other areas of human endeavor. He then commented on the alleged bias in the ISI databases:

> All such discussions are essentially concerned with the tail of a long hyperbolic curve. Once the core journals are selected, the remainder of one's effort is spent selecting from thousands of relatively small and low-impact journals published, both in the advanced as well as in the developing countries. (Garfield, 1997a, p. 640)

Garfield declared that ISI had developed the journal impact factor as just one method of supplementing the subjective appraisal of small journals by objective, unobtrusive means, and noted that citation impact was just one of many criteria used to select such journals. Of these criteria, he first mentioned that English had become the *lingua franca* of science, declaring: "Any journal which claims international significance will at minimum include English titles and abstracts" (p. 641). Garfield also pointed out that ISI coverage of articles by Third World scientists had substantially increased because they were increasingly publishing in the international journals. Returning to the relationship of size to significance, he stated that many Third World countries were suffering by publishing marginal journals and urged them to combine the best material into larger regional journals to achieve a critical mass, thereby following the precedent of the numerous European journals that had made many national journals essentially obsolete. In a later interview, Garfield (1999b, p. 69) characterized ISI bias in journal selection, if any, not as an “American” but as an “English-language” bias.

Quantitative versus Qualitative Criteria

Garfield’s (1990) most cogent description of ISI journal selection policies is a “Current Comments” essay adapted from a talk he gave in Taipei at the Symposium on Science Journal Evaluation sponsored by the National Science Council of Taiwan Science and Technology Information Center. The policies and considerations he set forth in this essay are still being largely followed by ISI today (Testa, 2004), long after his retirement as head of the company.

Garfield (1990) stated that ISI took into account three types of information, ranging from the quantitative to the qualitative, when evaluating journals for coverage: citation data, journal standards, and expert judgment. He gave the following overview of these three sources for evaluative criteria:

> Citation data are a source of quantitative indicators that can be used to evaluate existing journals with established track records. … But selection of new journals often relies on other, more qualitative considerations. Journal standards are an example. A journal’s ability to meet its declared schedule and frequency is perhaps the most basic expectation. Standards can also include editorial requirements for abstracts titles, and references set by professional associations of publishers and editors. Peer review of submissions, editorial board membership, and the reputation of the publisher or sponsoring society are other indicators of journal quality.
>
> Finally, journal selection also relies on the subjective judgment of experts in a particular field—subscribers, editors and publishers, and ISI’s many editorial advisory board members and staff specialists. (pp. 5–6)

Need for Subject Set Definition

Garfield (1990) began his discussion of the utilization of citation measures in ISI journal selection policies with his oft-reiterated warning on the need for a set definition before performing any citation analysis, stating in this respect:

> It should always be stressed that citation data must be carefully interpreted—and their limitations clearly understood—when they are used for evaluating anything. … For example, the number of authors and journals varies greatly

> between and within disciplines, as do their citation levels and rates. Smaller fields like botany or mathematics do not generate as many articles or citations as, say, biotechnology or genetics. Also, in certain fields it may take 10 or more years for an article to attract a meaningful number of citations, while in other research areas citations can typically peak after only a few years. (p. 7)

Total Citations, the Law of Concentration, and the Multidisciplinary Core

Following this mandatory warning, Garfield (1990) introduced the citation measures ISI used in determining journal coverage. He presented a table listing, in descending rank order, the 25 journals with the most total citations in the 1988 *SCI*. His comments on these journals were instructive because they dealt with the characteristics of the extreme upper stratum of those journals that, according to his Law of Concentration, form the relatively small, multidisciplinary core dominating all science:

> The list of the 25 most-cited journals in the 1988 *SCI* (Table 1) probably agrees closely with most readers' mental list of the most important scientific journals. Hardly anyone would dispute the inclusion of the *Journal of Biological Chemistry*, *Nature*, the *Proceedings of the National Academy of Sciences of the USA*, the *Journal of the American Chemical Society*, *Science*, or any others. The same basic group of journals tends to be most cited year after year. A few may gradually decline or be replaced by successful newcomers like *Cell* as editors and audiences change. But most successful journals survive and prosper for decades. Not surprisingly, the list is dominated by larger journals and the big life-sciences specialties. Fourteen also ranked among the top 25 by number of articles published. (pp. 7–8)

What is to be particularly noted here is that, according to Garfield, these most highly cited journals conform to scientists' conceptions of what an important scientific journal is, that they are among the largest in terms of number of articles published, and that they maintain their dominance over time, giving the journal social stratification system a high degree of stability. It can be hypothesized that scientists deem them important because they are old, largely bibliographically stable, and historically significant in terms of the number of citation classics they have published over the years.

Impact Factor, the Bradfordian Citation Comet Tail, and Review versus Research Journals

After discussing the characteristics of the journals comprising the upper stratum of the multidisciplinary core of journals posited by his Law of Concentration as dominating all science, Garfield (1990) turned to the impact factor—the citation measure used by ISI to select journals from among those journals that the law theorized as forming the long tails of the Bradfordian citation comets of each discipline. He began by stating that the main purpose of the impact factor—or "the average number of citations per article"—was to compensate for the "putative size advantage" of the journals comprising the multidisciplinary core (p. 8). Garfield's approach to the use of the impact factor in journal selection was twofold and he presented two tables to demonstrate his method: One listed the 25 journals ranking highest in impact factor in the 1988 *SCI* and the other enumerated the 25 journals publishing at least 100 articles that ranked highest in impact factor in the 1988 *SCI*. Garfield gave the following analysis of these tables:

> In Table 2 impact is calculated as follows: the number of articles published by a journal in 1986 and 1987 is divided into the number of citations they received in 1988. For example, the *Annual Review of Biochemistry* published 67 articles in 1986 and 1987. They received a total of 3,237 citations from ISI-covered journals in 1988. Thus, its impact factor is 48.3.
>
> The list is obviously dominated by review journals, which tend to publish fewer contributions than original research journals, but these are cited much more frequently. Table 3 presents another impact ranking, showing only journals that published at least 100 articles, which effectively excludes most review journals.
>
> Eighteen life-sciences journals are listed, compared to two each for chemistry and physics. Again, although impact compensates somewhat for the size of a journal or literature, it tends to favor research areas that more heavily cite recent research published in the last two years. As we found several years ago, the average number of references cited per article is perhaps the most significant contributing factor. This may or may not be a reflection of the field size. (p. 8)

Garfield's approach was interesting from two perspectives. First, both in the passage just quoted and in the tables, the impact factor was calculated only to the first decimal place. However, Garfield was dealing only with the extreme high end of a skewed distribution, where ties in rank from such a calculation would not be as overwhelmingly common as at the lower end of the distribution, reducing the impact factor's ability to discriminate. Second, Garfield used the unadulterated impact factor to identify his all-important review journals, but, for the evaluation of research journals, he reintroduced the element of size (which his research had found to be causal in journal significance) by restricting the set to those journals publishing 100 or more articles. The results supported Garfield's theory that high-impact research articles tend to concentrate in the multidisciplinary core dominating all science: Of the 25 journals that he identified as being highest in total citations, 12 were also on his list of the 25 journals highest in impact factor but publishing 100 or more articles per year.

Garfield (1990) also dealt with three other aspects of impact factor: calculating the impact factor of journals not covered by ISI, rankings by five-year impact factor, and item-by-item impact. These topics will not be covered in this chapter, as they are exceedingly complex, are crucial elements in the validation of impact factor, and play an important role in the use of impact factor for evaluative purposes; they therefore require the introduction of an entire set of new considerations for their explanation. However, it is worth noting that, in discussing the first of these topics, Garfield observed à propos of the complaints of Third World editors about their journals having a higher impact factor than journals covered by ISI:

> This reflects another common misconception—that impact factors are the sole or single most important criterion for coverage. In fact, journal impact is only one of several quantitative and qualitative factors described in this essay that we take into account. (p. 9)

Garfield has largely himself to blame for this misconception, for he has continually stressed the need for subject set definition in citation analysis, then ranked journals within defined subject groups only by impact factor in the *JCR*s, and calculated impact factor to three decimal places to assist the ranking process, thereby giving a false impression of the measure's precision.

Qualitative Criteria

At the end of the essay, Garfield (1990) discussed the other considerations ISI took into account in the selection of journals. First among these was the "internationality" of a journal, defined by the nationality of items it publishes and the nationality of the articles that cite it (p. 10). This consideration was followed by others such as the ability of the journal to meet publication schedules, whether it followed international editorial conventions, whether its articles were written in English or at least had informative English summaries, whether its manuscripts were subject to peer review, the track record of its editorial board and contributors, the reputation of its publisher, and the judgment of ISI and outside experts.

Summary and Conclusions

This chapter represents an attempt to analyze the impact factor from the perspective of Garfield's own intellectual development. Garfield was found to be a complex and sometimes contradictory thinker, with many of his contradictions stemming from the complexity of the phenomena he was examining. His intellectual roots were traced back to the scientific revolution in Britain that laid the foundations for modern information science and established the theoretical framework within which his ideas developed. This revolution had both a reformist and a revolutionary wing. The ideas of the reformist wing were set forth by S. C. Bradford, who sought to improve the existing methods of scientific communication through the established system of journals and libraries. He justified capping this system with a central document delivery library by his Law of Scattering. Bradford's Law stemmed from the unity of science and described the skewed distribution of articles on a given topic over journals that made it impossible for any special library to have complete journal holdings on the subject of interest to it. The revolutionary wing was led by J. D. Bernal, who regarded articles as the prime vehicle of scientific communication and journals as inefficient bundles of articles. Being a Communist, Bernal sought to overthrow the established journal system and nationalize scientific communication by means of a central distribution system that would group articles into convenient packages meeting the needs of individual scientists. However, Bernal himself was forced to recognize the inapplicability of his system due to Bradford's Law, which made it virtually impossible to define such convenient packages. Bernal emphasized the importance of scientific review literature in his theoretical writings and validated this importance in his empirical research.

Due to family influences, Garfield was introduced to the British literature through the work of Bernal. As a result, he had a tendency to emphasize the importance of the article over the journal. Moreover, Bernal's stress on the importance of scientific review literature was reinforced in Garfield's thought by Chauncey D. Leake, who urged him to study the importance and function of the review article. Garfield combined the structure of the review article with that of the legal citator, to which he was introduced by William C. Adair, to create the *Science Citation Index*. As a result of these influences, Garfield came to esteem review articles as the epitome of scientific writing, whose function was to serve as arbiters in scientific controversies and direct the course of scientific research. As a further sign of Bernal's influence, one of the primary motivations in Garfield's creation of the *Science Citation Index* was to solve the problem posed by Bradford's Law of Scattering for defining convenient packages of articles of interest to individual scientists.

However, Garfield's attempt to solve the problem that stymied Bernal ultimately resulted only in confirming Bernal's conclusion. By describing the inherently interdisciplinary nature of science with his Law of Concentration, Garfield may have solved the problem posed by Bradford's Law for the comprehensive indexing of scientific journals, but he also demonstrated that Bernal's problem is perhaps impossible to solve.

Garfield's intellectual development was also heavily influenced by Bernal in another important respect. As a Marxist, Bernal pioneered the study of the societal aspects of science. His influence on Garfield in this respect was reinforced by the latter's association with the founder of the sociology of science, Robert K. Merton, and the historian of science, Derek J. de Solla Price. Together, these three men formed an invisible college. Partly as a result of these influences, Garfield came to change his concept of scientific journals from Bernal's inefficient bundles of articles to sociological entities, whose scientific and social significance could be measured by citation frequency. One of the most important contributions of Merton and Price to Garfield's thinking was to help shape his understanding of the skewed distributions underlying scientific phenomena. Merton supplied a causal theory of such distributions with his Matthew effect, describing a cumulative advantage or success-breeds-success process. For his part, Price clarified the form of these distributions with his analyses of the exponential character of science and the similarity of scientific distributions to ones in biology and society. It was with this changed concept of scientific journals that Garfield launched a series of statistical analyses of the citation structure of the scientific journal system that led to his formulation of the Law of Concentration and the creation of the *Journal Citation Reports*.

Garfield's Law of Concentration was a reformulation of Bradford's Law of Scattering, according to which articles on a given subject concentrate primarily in a small core of journals and then spread over zones of other journals that have to increase exponentially in numbers of titles to contain the same number of articles on the topic as the journals in the core. Garfield found the same pattern in the distribution of citations over the journals of a given scientific discipline and he likened the pattern to a comet, with the nucleus symbolizing the core of journals receiving the bulk of the citations and the tail symbolizing zones of the lesser journals that have to increase exponentially in numbers to receive the same number of citations as the journals in the core. Garfield's main innovation was to raise Bradford's Law from the level of individual disciplines to that of science as a whole and show that the tail journals of one discipline consist largely of core journals of other disciplines, resulting in a small multidisciplinary core of journals dominating all of science. Garfield's analyses of the journals of this multidisciplinary core found them to be large in terms of numbers of articles published and old in the sense that they maintain their dominance over decades: They are the journals in which the most significant articles in terms of citation counts concentrate. Garfield found that these core journals are also those journals that scientists, librarians, scientific publishers, and he himself have tended intuitively to identify as the most significant journals of science.

Garfield's understanding and use of the impact factor must be viewed from the standpoint of his evaluation of the importance of review articles and his Law of Concentration. Garfield found that the journals of the dominant multidisciplinary core are easily identified by total citations. Once this is done, he claimed, the main problem is to select journals from the long tail of the citation comet: This was the function of the impact factor. This measure counteracted the age and size advantage of the journals of the multidisciplinary core by restricting the observation period to the two most recent years and calculating the ratio of citations to citable items to estimate the average citation rate per article. Doing this provides a different perspective on journal significance in two respects. First, it brings review journals to the top, allowing an easy way to identify these all-important journals. Second, it allows one to estimate a journal's current, as against its historical, significance by evaluating only its most recent materials. Garfield was well aware of the impact factor's propensity to rank review journals at the top and therefore used the impact factor in two different ways. To identify review journals, he used the impact factor in its unadulterated form, but, to evaluate research journals, he reintroduced size by limiting the set under evaluation to those journals publishing 100 or more articles. The result of the latter move was a tendency to bring to the top the journals of the multidisciplinary core, because these are the journals in which the high-impact research articles tend to concentrate. Garfield was well aware of another aspect of the impact factor—the extremely low citation frequency of the vast bulk of scientific articles. This finding was expressed by his invention of "Garfield's constant"—the ratio of citations to cited items—which rose from only 1.33 to merely 2.25 over the 50-year period 1945–1995. Given the skewed distributions of journals by citations, this meant that journal impact factors had concentrated for the most part in an extremely short range of around 2 or below. As a result, the impact factor could not be used by itself to discriminate among journals at the lower range, where most journals were. Therefore, Garfield always urged that the impact factor should always be used in conjunction with other, subjective considerations.

This brings us to what seems to be a logical contradiction between the structure of Garfield's thought as it pertains to the impact factor and the structure of the *JCR*s. Garfield always mandated that citation analysis could be performed only within carefully defined subject sets. Yet in the print and microform editions of the *JCR*s, where the rankings were immutable, the only ranking of journals within defined subject sets was by impact factor. Given the identification by total citations of the titles in the multidisciplinary journal core posited by the Law of Concentration as dominating all science, it would seem to have been more logical to rank journals within defined subject sets by this measure instead of by impact factor. Moreover, the impact factor rankings were given a false sense of precision by the calculation of the impact factor to three decimal places instead of the one decimal place Garfield himself regarded as more realistic. This was done to avoid the overwhelming number of ties in ranks at the lower range that would inevitably have resulted if the impact factor would have been calculated to only one decimal place. The consequence was to enshrine the impact factor as ISI's main measure of journal significance for the outside world, spawning what can only be termed the impact factor industry. One can only speculate on the reasons for this: (1) the impact factor is an egalitarian measure and Garfield was indoctrinated into leftist political thinking as a youth, (2) the impact factor is a measure of article significance and Garfield regarded the article instead of the journal as the main vehicle of scientific communication, (3) the impact factor ranks review journals at the top and Garfield regarded the review article as the epitome of scientific writing, (4) the impact factor minimizes the problems resulting from the bibliographic instability of journals by limiting the time period to only two years, and (5) the impact factor is a measure of current, in contrast to historical, significance. This last consideration may have been of crucial importance for Garfield, because, to ensure that ISI journal coverage was current with the constant shifts in scientific importance, he needed a measure that helped him detect new rising stars and changes in journal significance.

Endnote

1. A comprehensive collection of Garfield's writings—scientific, (auto)biographical, and occasional—is available on his home page (www.garfield.library.upenn.edu).

References

Adair, W. C. (1955). Citation indexes for scientific literature? *American Documentation*, *6*(1), 31–32.

Baker, J. (1939). Counterblast to Bernalism. *The New Statesman and Nation*, *18*(440), New Series, 174–175.

Bernal, J. D. (1931). Science and society. *Spectator*, *147*(5376), 43–44.

Bernal, J. D. (1940). *The social function of science* (2nd corrected ed.). London: Routledge.

Bernal, J. D. (1948a). The form and distribution of scientific papers. In *The Royal Society Empire Scientific Conference, June-July 1946: Report* (Vol. 1, pp. 698–699). London: The Society.

Bernal, J. D. (1948b). Preliminary analysis of pilot questionnaire on the use of scientific literature. In *The Royal Society Scientific Information Conference, 21 June-2 July 1948: Report and papers submitted* (pp. 589–637). London: The Society.

Bernal, J. D. (1948c). Provisional scheme for central distribution of scientific publications. In *The Royal Society Scientific Information Conference, 21 June-2 July 1948: Report and papers submitted* (pp. 253–258). London: The Society.

Bernal, J. D. (1965). Science Citation Index. *Science Progress*, *53*, 455–459.

Bradford, S. C. (1934, January 26). Sources of information on specific subjects. *Engineering*, *137*, 85–86.

Bradford, S. C. (1948a). Complete documentation. In *The Royal Society Empire Scientific Conference, June-July 1946: Report* (Vol. 1, pp. 729–748). London: The Society.

Bradford, S. C. (1948b). *Documentation*. London: Crosby Lockwood.

Bukharin, N. I. (1931). Theory and practice from the standpoint of dialectical materialism. In *Science at the cross roads: Papers presented to the International Congress of the History of Science and Technology held in London from June 29th to July 3rd, 1931, by the delegates of the U.S.S.R.* London: Kniga.

Cole, J. R. (2000). A short history of the use of citations as a measure of the impact of scientific and scholarly work. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge: A Festschrift in honor of Eugene Garfield* (pp. 281–300). Medford, NJ: Information Today.

Cole, S. (2000). The role of journals in the growth of scientific knowledge. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge: A Festschrift in honor of Eugene Garfield* (pp. 109–142). Medford, NJ: Information Today.

Davis, W. (1940). Project for scientific publication and bibliography. In J. D. Bernal, *The social function of science* (2nd corrected ed., pp. 449–457). London: Routledge.

East, H. (1998). Professor Bernal's "insidious and cavalier proposals": The Royal Society Scientific Information Conference, 1948. *Journal of Documentation*, *54*, 293–302.

Garfield, E. (1955). Citation indexes for science: A new dimension through association of ideas. *Science*, *122*, 108–111.

Garfield, E. (1956). Citation indexes—new paths to scientific knowledge. *Chemical Bulletin*, *43*(4), 11.

Garfield, E. (1959). A unified index to science. *Proceedings of the International Conference on Scientific Information, Washington, D.C., November 16–21, 1958* (Vol. 1, pp. 461–469). Washington, D.C.: National Academy of Science, National Research Council.

Garfield, E. (1963). Citation indexes in sociological and historical research. *American Documentation*, *14*, 289–291.

Garfield, E. (1970a, April 22). Calling attention to Chauncey D. Leake—Renaissance scholar extraordinaire. *Current Contents, No. 16*, 5–6.

Garfield, E. (1970b, May 6). What is a significant journal? *Current Contents, No. 18*, 5–6.

Garfield, E. (1971, Aug. 4). The mystery of transposed journal lists—wherein Bradford's Law of Scattering is generalized according to Garfield's Law of Concentration. *Current Contents, No. 17*, 5–6.

Garfield, E. (1972a). Citation analysis as a tool in journal evaluation. *Science*, *178*, 471–479.

Garfield, E. (1972b, February 16). Citation statistics may help scientists choose journals in which to publish. *Current Contents, No. 6*, 5–6.

Garfield, E. (1972c, February 23). Citations-to divided by items-published gives journal impact factor; ISI lists the top fifty high-impact journals of science. *Current Contents, No. 7*, 5–8.

Garfield, E. (1972d, April 5). Is citation frequency a valid criterion for selecting journals? *Current Contents, No. 13*, 5–6.

Garfield, E. (1973a, February 7). Citation frequency and citation impact; and the role they play in journal selection for *Current Contents* and other ISI services. *Current Contents, No. 6*, 5–6.

Garfield, E. (1973b, August 22). The economics and Realpolitik of exponential information growth; or, Journal selection ain't easy! *Current Contents, No. 34*, 5–6.

Garfield, E. (1973c, September 26). Which journals attract the most frequently cited articles? Here's a list of the top fifteen. *Current Contents, No. 39*, 5–6.

Garfield, E. (1974a, November 13). So why don't we have *science* reviews? *Current Contents, No. 46*, 5–6.

Garfield, E. (1974b, October 30). So you wanted more review articles—ISI's new *Index to Scientific Reviews* (*ISR*) will help you find them. *Current Contents, No. 44*, 5–6.

Garfield, E. (1975, January 6). The who and why of ISI. *Current Contents, No. 1*, 5–6.

Garfield, E. (1976a, February 9). Is the ratio between number of citations and publications cited a true constant? *Current Contents, No. 6*, 5–7.

Garfield, E. (Ed.). (1976b). *Journal citation reports: A bibliometric analysis of references processed for the 1974 Science Citation Index.* (Science Citation Index 1975 Annual, v. 5). Philadelphia: Institute for Scientific Information.

Garfield, E. (1976c). La science française est-elle trop provinciale? [Is French science too provincial?] *La Recherche, 7*, 757–760.

Garfield, E. (1976d). Significant journals of science. *Nature, 264*, 609–615.

Garfield, E. (1977a, April 11). Le nouveau défi américain, I. *Current Contents, No. 15*, 5–10.

Garfield, E. (1977b, April 18). Le nouveau défi américain, II. *Current Contents, No. 16*, 5–12.

Garfield, E. (1977c, April 4). Proposal for a new profession: Scientific reviewer. *Current Contents, No. 14*, 5–8.

Garfield, E. (1977d, July 11). Robert K. Merton: Among the giants. *Current Contents, No. 28*, 5–8.

Garfield, E. (1978, February 13). To remember Chauncey D. Leake. *Current Contents, No. 7*, 5–15.

Garfield, E. (1979a). *Citation indexing: Its theory and application in science, technology, and humanities.* Philadelphia: ISI Press.

Garfield, E. (1979b, November 5). How do we select journals for *Current Contents*? *Current Contents, No. 45*, 5–8.

Garfield, E. (Ed.). (1979c). *SCI journal citation reports: A bibliometric analysis of science journals in the ISI data base.* (Science Citation Index 1978 Annual, v. 13). Philadelphia: Institute for Scientific Information.

Garfield, E. (1980a). Citation measures of the influence of Robert K. Merton. In T. F. Gieryn (Ed.), *Science and social structure: A Festschrift for Robert K. Merton* (Transactions of the New York Academy of Sciences, series 2, vol. 39, pp. 61–74). New York: New York Academy of Sciences.

Garfield, E. (Ed.). (1980b). *SCI journal citation reports: A bibliometric analysis of science journals in the ISI data base.* (Science Citation Index 1979 Annual, v. 14). Philadelphia: Institute for Scientific Information.

Garfield, E. (1982a, September 27). ISI's "new" *Index to Scientific Reviews* (*ISR*): Applying research front specialty searching to the retrieval of the review literature. *Current Contents, No. 39*, 5–12.

Garfield, E. (1982b, May 10). J. D. Bernal—The sage of Cambridge: 4S award memorializes his contributions to the social studies of science. *Current Contents, No. 19*, 5–17.

Garfield, E. (1983). Mapping science in the Third World. *Science and Public Policy, 10*, 112–127.

Garfield, E. (1985a, March 18). Journal selection for *Current Contents*: Editorial merit vs. political pressure. *Current Contents, No. 11*, 3–11.

Garfield, E. (1985b, August 26). Origins of *Current Contents*, ISI, and computer-aided information retrieval. How it all began at the Welch Medical Library Indexing Project. *Current Contents, No. 34*, 3–9.

Garfield, E. (1987a). *Oral history: Transcript of an interview conducted by A. Thackray and J. Sturchio at the Institute for Scientific Information, Philadelphia, PA, on 16 November 1987*. Philadelphia: Beckman Center for the History of Chemistry. Retrieved January 20, 2006, from www.garfield.library.upenn.edu/oralhistory/interview.html

Garfield, E. (1987b, May 4). Reviewing review literature: Part 1, Definitions and uses of reviews. *Current Contents, No. 18*, 5–8.

Garfield, E. (1987c, May 11). Reviewing the review literature: Part 2, The place of reviews in the scientific literature. *Current Contents, No. 19*, 5–10.

Garfield, E. (1988, June 6). French research: Citation analysis indicates trends are more than just a slip of the tongue. *Current Contents, No. 23*, 3–11.

Garfield, E. (1990, May 28). How ISI selects journals for coverage: Quantitative and qualitative considerations. *Current Contents, No. 22*, 5–13.

Garfield, E. (1991, September 2). In truth, the 'flood' of scientific literature is only a myth. *The Scientist*, *5*(17), 11.

Garfield, E. (1995). Quantitative analysis of the scientific literature and its implications for science policymaking in Latin America and the Caribbean. *Bulletin of the Pan American Health Organization*, *29*(1), 87–95.

Garfield, E. (1996, September 2). The significant scientific literature appears in a small core of journals. *The Scientist*, *10*(17), 14–16.

Garfield, E. (1997a). A statistically valid definition of bias is needed to determine whether the *Science Citation Index* discriminates against Third World journals. *Current Science*, *73*, 639–641.

Garfield, E. (1997b). *Transcript of an interview conducted by R. V. Williams at Philadelphia, Pennsylvania, on 29 July 1997 (with subsequent corrections and additions)*. Philadelphia: Chemical Heritage Foundation. Retrieved January 20, 2006, from http://garfield.library.upenn.edu/papers/oralhistorybywilliams.pdf

Garfield, E. (1998a). From citation indexes to informetrics: Is the tail now wagging the dog? *Libri*, *48*, 67–80.

Garfield, E. (1998b). Random thoughts on citationology: Its theory and practice. *Scientometrics*, *43*, 69–76.

Garfield, E. (1999a, October). Deeds and dreams of Eugene Garfield: Interview by Istvan Hargittai. *Chemical Intelligencer*, *5*(4), pp. 26–31.

Garfield, E. (1999b). An interview with Eugene Garfield: Conducted by Beatrice L. Caraway. *Serials Review*, *25*(3), 67–80.

Garfield, E. (1999c). On the shoulders of giants. In M. E. Bowden, T .B. Hahn, & R. V. Williams (Eds.), *Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems* (pp. 237–251). Medford, NJ: Published for the American

Garfield, E. (1982b, May 10). J. D. Bernal—The sage of Cambridge: 4S award memorializes his contributions to the social studies of science. *Current Contents, No. 19*, 5–17.

Garfield, E. (1983). Mapping science in the Third World. *Science and Public Policy, 10*, 112–127.

Garfield, E. (1985a, March 18). Journal selection for *Current Contents*: Editorial merit vs. political pressure. *Current Contents, No. 11*, 3–11.

Garfield, E. (1985b, August 26). Origins of *Current Contents*, ISI, and computer-aided information retrieval. How it all began at the Welch Medical Library Indexing Project. *Current Contents, No. 34*, 3–9.

Garfield, E. (1987a). *Oral history: Transcript of an interview conducted by A. Thackray and J. Sturchio at the Institute for Scientific Information, Philadelphia, PA, on 16 November 1987*. Philadelphia: Beckman Center for the History of Chemistry. Retrieved January 20, 2006, from www.garfield.library.upenn.edu/oralhistory/interview.html

Garfield, E. (1987b, May 4). Reviewing review literature: Part 1, Definitions and uses of reviews. *Current Contents, No. 18*, 5–8.

Garfield, E. (1987c, May 11). Reviewing the review literature: Part 2, The place of reviews in the scientific literature. *Current Contents, No. 19*, 5–10.

Garfield, E. (1988, June 6). French research: Citation analysis indicates trends are more than just a slip of the tongue. *Current Contents, No. 23*, 3–11.

Garfield, E. (1990, May 28). How ISI selects journals for coverage: Quantitative and qualitative considerations. *Current Contents, No. 22*, 5–13.

Garfield, E. (1991, September 2). In truth, the 'flood' of scientific literature is only a myth. *The Scientist, 5*(17), 11.

Garfield, E. (1995). Quantitative analysis of the scientific literature and its implications for science policymaking in Latin America and the Caribbean. *Bulletin of the Pan American Health Organization, 29*(1), 87–95.

Garfield, E. (1996, September 2). The significant scientific literature appears in a small core of journals. *The Scientist, 10*(17), 14–16.

Garfield, E. (1997a). A statistically valid definition of bias is needed to determine whether the *Science Citation Index* discriminates against Third World journals. *Current Science, 73*, 639–641.

Garfield, E. (1997b). *Transcript of an interview conducted by R. V. Williams at Philadelphia, Pennsylvania, on 29 July 1997 (with subsequent corrections and additions)*. Philadelphia: Chemical Heritage Foundation. Retrieved January 20, 2006, from http://garfield.library.upenn.edu/papers/oralhistorybywilliams.pdf

Garfield, E. (1998a). From citation indexes to informetrics: Is the tail now wagging the dog? *Libri, 48*, 67–80.

Garfield, E. (1998b). Random thoughts on citationology: Its theory and practice. *Scientometrics, 43*, 69–76.

Garfield, E. (1999a, October). Deeds and dreams of Eugene Garfield: Interview by Istvan Hargittai. *Chemical Intelligencer, 5*(4), pp. 26–31.

Garfield, E. (1999b). An interview with Eugene Garfield: Conducted by Beatrice L. Caraway. *Serials Review, 25*(3), 67–80.

Garfield, E. (1999c). On the shoulders of giants. In M. E. Bowden, T .B. Hahn, & R. V. Williams (Eds.), *Proceedings of the 1998 Conference on the History and Heritage of Science Information Systems* (pp. 237–251). Medford, NJ: Published for the American

Society for Information Science and the Chemical Heritage Foundation by Information Today.

Garfield, E. (2000). In-depth interview: Interview with Eugene Garfield, Ph.D., by Chungji Kim. *Medical Writing, 8*(1). Retrieved July 20, 2005, from www.garfield.library.upenn.edu/papers/medicalwritingv8(1)1999.html

Garfield, E. (2002, June). Citation consciousness: The origins of citation indexing in science: Interview with Eugene Garfield, Chairman Emeritus of ISI, Philadelphia, U.S.A. *Password, Nr.* 6, 22–25.

Garfield, E. (2003). The meaning of impact factor. *Revista Internacional de Psicología Clínica y de la Salud, 3*, 363–369.

Garfield, E. (2005). The agony and the ecstasy—the history and meaning of the journal impact factor. Presented at the International Conference on Peer Review and Biomedical Publication. Retrieved November 18, 2005 from http://garfield.library.upenn.edu/papers/jifchicago2005.pdf

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation, 14*(3), 195–201.

Goldsmith, M., & Mackay, A. (1964). Introduction. In M. Goldsmith & A. Mackay (Eds.), *The science of science: Society in the technological age* (pp. 9–17). London: Souvenir Press.

Hodgkin, D. M. C. (1980). John Desmond Bernal, 10 May 1901–15 September 1971. *Biographical Memoirs of Fellows of the Royal Society, 26*, 17–84.

Lederberg, J. (2000). How the *Science Citation Index* got started. In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge: A Festschrift in honor of Eugene Garfield* (pp. 25–64). Medford, NJ: Information Today.

Martyn, J., & Gilchrist, A. (1968). *An evaluation of British scientific journals* (Aslib Occasional Publication, no. 1). London: Aslib.

Memorandum on Section I dealing with the central distribution of scientific publications. (1948). In *The Royal Society Scientific Information Conference, 21 June-2 July 1948: Report and papers submitted* (pp. 516–519). London: The Society.

Merton, R. K. (1941). [Review of *The social function of science*, by J. D. Bernal]. *American Journal of Sociology, 46*, 622–623.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*, 56–63.

Merton, R. K. (1983). Foreword. In E. Garfield, *Essays of an Information Scientist: Vol. 5, 1981–1982* (pp. xv–xix). Philadelphia: ISI Press.

Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis, 79*, 606–623.

Muddiman, D. (2003). Red information science: The information career of J. D. Bernal. *Journal of Documentation, 59*, 387–409.

Muddiman, D. (2004). Red information science: J. D. Bernal and the nationalization of scientific information in Britain from 1930 to 1949. In W. B. Rayward & M. E. Bowden (Eds.), *The History and Heritage of Scientific and Technological Information Systems: Proceedings of the 2002 conference* (pp. 258–266). Medford, NJ: Information Today.

Price, D. J. D. (1963). *Little science, big science*. New York: Columbia University Press.

Price, D. J. D. (1964). The science of science. In M. Goldsmith & A. Mackay (Eds.), *The science of science: Society in the technological age* (pp. 195–208). London: Souvenir Press.

Price, D. J. D. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science, 27*, 292–306.

Price, D. J. D. (1978). Cumulative advantage urn games explained: A reply to Kantor. *Journal of the American Society for Information Science*, *29*, 204–206.

Price, D. J. D. (1980). Foreword. In E. Garfield, *Essays of an information scientist: Vol. 3, 1977–1978* (pp. v–ix). Philadelphia: ISI Press.

Price, D. J. D. (1983, July 18). The week's citation classic: Price D. J. D. *Little science, big science* (New York: Columbia University Press, 1963). *Current Contents, No. 29*, 18.

Recommendations. (1948). In *The Royal Society Scientific Information Conference, 21 June–2 July 1948: Report and papers submitted* (pp. 195–208). London: The Society.

Robinson, R. (1948). [Opening address]. In *The Royal Society Scientific Information Conference, 21 June-2 July 1948: Report and papers submitted* (pp. 15–17). London: The Society.

Sandison, A. (1971). Library optimum. *Nature*, *234*, 368–369.

Sher, I. H., & Garfield, E. (1966). New tools for improving and evaluating the effectiveness of research. In M. C. Yovits, D. M. Gilford, R. H. Wilcox, E. Staveley, & H. D. Lerner (Eds.), *Research program effectiveness: Proceedings of the conference sponsored by the Office of Naval Research, Washington, D.C., July 27–29, 1965* (pp. 135–146). New York: Gordon and Breach.

Testa, J. (2004). *The ISI database: The journal selection process*. Retrieved September 27, 2005, from http://scientific.thomson.com/knowtrend/essays/selectionofmaterial/journal selection