

Eugene Garfield's Ideas and Legacy

Subjects: Information Science & Library Science

Contributor: Svetla Baykoucheva

Eugene Garfield introduced information systems that made the discovery of scientific information much more efficient. The founded by him Institute for Scientific Information (ISI) developed innovative information products and provided current scientific information to researchers all over the world. Garfield introduced the citation as a qualitative measure of academic impact and propelled the concepts of "citation indexing" and "citation linking", paving the way for today's search engines. He created the Science Citation Index (SCI), which provided a new way of retrieving, organizing, disseminating, and using scientific information; triggered the development of new disciplines (scientometrics, infometrics, webometrics); and became the foundation for building important new information products. The Journal Impact Factor, initially established to select journals for the SCI, became the most widely accepted tool for measuring academic impact. Garfield actively promoted English as the international language of science and became a powerful force in the globalization of research. His ideas revolutionized science and dramatically influenced the culture of research. This Encyclopedia entry is based on the following article:

Baykoucheva, S. Eugene Garfield's Ideas and Legacy and Their Impact on the Culture of Research. *Publications* 2019, 7, 43. (<https://www.mdpi.com/2304-6775/7/2/43>)

Keywords: information science ; Science Citation Index ; citations ; Web of Science ; Current Contents ; scientific information ; culture of research ; Journal Impact Factor



Eugene Garfield

I had always envisaged a time when scholars would become citation conscious, and to a large extent they have, for information retrieval, evaluation, and measuring impact...I did not imagine that the worldwide scholarly enterprise would grow to its present size, or that bibliometrics would become so widespread.

Eugene Garfield

Eugene Garfield advanced the theory and practice of information science, raised awareness about citations and academic impact, actively promoted English as the lingua franca of science, and became a powerful force in the globalization of research. He founded the Institute for Scientific Information (ISI) (Figure 1), which developed innovative information products and provided current scientific information to researchers all over the world. The creation of the Science Citation Index (SCI) is one of the most important events in modern science . Garfield's idea of using citations in articles to index

scientific literature offered a new and more efficient way of collecting, analyzing, disseminating, and discovering scientific information. The SCI laid the foundation for building new important information products such as Web of Science (one of the most widely used databases for finding scientific literature today), Essential Science Indicators, and Journal Citation Reports. It triggered the development of new disciplines (such as scientometrics, infometrics, and webometrics) and preceded the search engines, which use “citation linking”—a core concept of the SCI—to connect and rank documents. Some have called Garfield “The Grandfather of Google” . “Citation linking” was probably considered by Sergey Brin's and Larry Page for their article in which they mentioned Google for the first time. Many articles, books, conference presentations, and interviews discuss and analyze Garfield's ideas and legacy ^{[1][2][3][4][5][6][7][8][9][10][11][12][13][14][15]}.



Figure 1. Institute for Scientific Information (ISI) at 3501 Market Street, Philadelphia. Courtesy of Eugene Garfield's personal archive.

1. Revolutionizing Scientific Information

The citation becomes the subject ... It was a radical approach to retrieving information.

Eugene Garfield (Interviewed by Eric Ramsay) ^[13]

While working at the Welch Medical Library at Johns Hopkins University in the early 1950s, Garfield studied the role and importance of review articles and their linguistic structure. He concluded that each sentence in a review article could be an indexing statement and that there should be a way to structure it. About that time, he also became interested in the idea of creating an equivalent to Shepard's Citations (a legal index for case law) for science. Garfield described in an interview how these events led him to conceive the SCI ^[1]. The SCI is a multidisciplinary directory to scientific literature, which facilitates interdisciplinary research. It provided a new way of organizing, disseminating, and retrieving scientific information. Garfield introduced the SCI as a “new dimension in indexing”^[2]. Citation indexing uses the concept that the literature cited in an article represents the most significant research performed before. It stipulates that, when an article cites another article, there should be something in common between the two. In this case, each bibliographic citation acts as a descriptor or symbol that defines a question (treated from a certain aspect in the cited work). It is strange to imagine today that one of the most important resources for modern science was created in a converted chicken coop in rural New Jersey (Figure 2). The citation indexes produced by ISI were the most important sources of bibliometric information before Reed Elsevier launched Scopus in 2004. Most of the products associated with Garfield's legacy are now part of the Clarivate Analytics' platform Web of Science (WoS).



Figure 2. Eugene Garfield in front of the converted chicken coop in Thorofare, NJ, where he started the Science Citation Index. With permission from Eugene Garfield's personal archive.

Some critical developments prepared the ground for the emergence of a tool such as the SCI. In the years after World War II, there was a significant increase in funding for research, which led to a rapid expansion of science and literature growth. There was widespread dissatisfaction with the traditional indexing and abstracting services, because they were discipline-focused and took a long time to reach users. Researchers wanted recognition, and there was a need for quantifiable tools to evaluate journals and individuals' work. The Journal Impact Factor (IF) has become the most accepted tool for measuring the quality of journals. A by-product from the SCI, it was originally designed as a tool to select journals for coverage in the SCI. It has been used later [largely inappropriately] for measuring the quality of individual researchers' work, as well as in science policy and research funding. As researchers are evaluated, hired, promoted, and funded based on the impact of their work, the importance of publishing in high-impact journals has created an extremely competitive environment, where the "publish or perish" philosophy dominates academic life. In his book, "Citation Indexing—Its Theory and Application in Science, Technology, and Humanities", Garfield presented the core concepts on which SCI was built ^[16] (Figure 3). Garfield will be remembered mostly for the Current Contents and the Science Citation Index (SCI), but other important information products are also part of his legacy (Table 1).

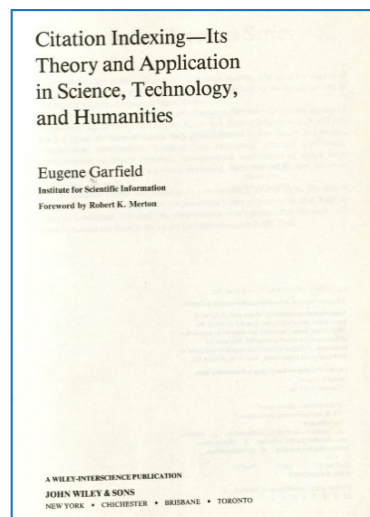


Figure 3. Garfield's book on citation indexing that laid out the core concepts of the SCI.

Table 1. Information products associated with the legacy of Eugene Garfield.

Product	Features
Science Citation Index (SCI) *	A citation database launched by ISI in 1964. Science Citation Index Expanded covers over 8500 major journals, across 150 disciplines, from 1900 to the present.
Social Sciences Citation Index (SSCI) *	A citation database, covering over 3200 of the world's leading academic journals in the social sciences across more than 55 disciplines, as well as selected items from 3500 of the world's leading scientific and technical journals, from 1900 to present.
Arts & Humanities Citation Index (A&HCI)*	A citation database covering over 1700 arts and humanities fully indexed journals, as well as selected items from over 250 scientific and social sciences journals, from 1975 to present.
Index Chemicus *	A text- and substructure-searchable database, offering full graphical summaries, important reaction diagrams, and complete bibliographic information from over 100 of the world's leading organic chemistry journals. Included in Clarivate Analytics' Web of Science platform.
Current Chemical Reactions *	A database containing single- and multi-step new synthetic methods. The overall reaction flow is provided for each method, along with a detailed and accurate graphical representation of each reaction step. Included in Clarivate Analytics' Web of Science platform.
<i>Current Contents</i>	A weekly journal (first published in print) reproducing the table of contents from journal issues of major peer-reviewed scientific journals published only a few weeks ago—a shorter time lag than any service then available. Contained an author index and a keyword subject index. Author addresses were provided so readers could send reprint requests for copies of the actual articles. Current status: Still published in print, it is available as one of the databases included in Clarivate Analytics' platform.
Essays of an Information Scientist	Eugene Garfield, <i>Essays of an Information Scientist</i> , Volumes 1–15, Philadelphia: ISI Press. A 15-volume series, which includes the original essays of Eugene Garfield published in the <i>Current Contents</i> , as well as some of his articles published elsewhere.
Web of Science Core Collection *	Includes the Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, Emerging Sources Citation Index, Book Citation Index, and Conference Proceedings Citation Index.
Journal Citation Reports (JCR) *	Publishes the annual journal impact factors of scientific journals and provides metrics and analysis for them. As of 1 June 2019, it covered 11,500+ indexed journals, 230+ disciplines, 80 countries/regions, and 2.2 million articles, reviews, and other source items.
InCites Essential Science Indicators (ESI) *	A compilation of science performance statistics and science trends data based on journal article publication counts and citation data from Clarivate Analytics' databases, which shows the influential individuals, institutions, papers, publications, and countries in their field of study, as well as emerging research areas that can impact their work.
<i>The Scientist</i> *	A newspaper for scientists.

* Published by Clarivate Analytics.

2. *Current Contents*: A Little Weekly Journal with a Global Reach

At a time when there were no online databases and internet, Garfield came up with the simple idea of making current scientific information available to researchers by compiling the tables of contents (TOCs) of key scientific journals and publishing them in a weekly journal called *Current Contents* (CC). Printed on thin paper (to reduce mailing costs), this pocket-sized journal ran series for agricultural, biological, and environmental sciences; arts and humanities; clinical practice; engineering, technology, and applied sciences; life science; physical chemical and earth sciences; and social and behavioral sciences. Instead of browsing the actual print issues, readers could look up quickly the journals' TOCs. For many years, CC was the most important source for current scientific information for researchers all over the world. CC included authors' addresses, which allowed readers to send requests for article reprints.

Almost every issue of *Current Contents* included an essay by Garfield, accompanied by his picture (some of these pictures are shown in Figure 4). Garfield was a prolific writer; he wrote more than 1 500 essays, which were later collected in a 15-volume series under the title "*Essays of an Information Scientist*"(Figure 5). While the essays reflected his own interests, they also touched on topics that were of interest to a large audience. He wrote about medical issues, jazz, psychedelic art, ice cream, and windsurfing, but he also discussed citation indexing, new information technologies, Nobel

Prizes, scientific publishing, and cultural institutions. In these essays, Garfield was able to convey his ideas in such way that even people who were not familiar with the topic could understand what he was saying.

Some have blamed Garfield for being too personal and that the essays were autobiographical—"a journey to himself". He defended his style and the topics he wrote about by saying that it was for these essays that he received the most letters and requests for reprints. This is what Garfield said in an interview about his essays: "... there never was a shortage of topics to cover in CC. We had meetings every week and put out ideas on topics to cover and then we would assign them to one of my many assistants. As the years went by the scope became larger and larger. One series of essays we did took over two years to complete, since we had to generate a database for each part. Some of the more personal topics just came up in the course of everyday events. Eventually I realized that our readers were happy to have a distraction from the scientific topics and found personal subjects more interesting. Topics like jazz and family appealed to a lot of people. The CC essays were very popular in East Europe because CC was not censored. I wrote in such a way that it gave those scientists a window on the west, but did not offend their bosses". Reading and discussing Garfield's essays became a favorite pastime for many researchers who impatiently waited to see the next issue of CC—not only to find out what was published in the newest issues of the scientific journals, but also to read his essay. Garfield's writings had a huge educational impact on readers, as they allowed them to see beyond their own discipline and narrow area of research.



Figure 4. Pictures of Eugene Garfield, some of which accompanied his essays in *Current Contents*. With permission from Eugene Garfield's personal archive.

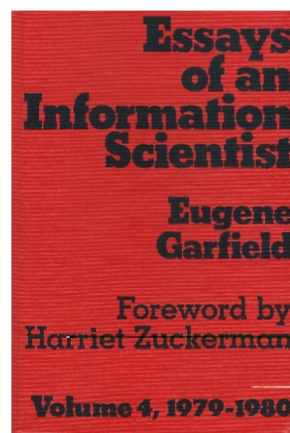


Figure 5. Eugene Garfield wrote more than 1500 essays, which were later collected in a 15-volume series published by ISI Press.

| 3. The Making of a Citation-Conscious [or Citation-Obsessed] Society

There are many reasons why the Journal Impact Factor (IF) has attracted so much attention. The impact of researchers' work is measured by the quality of the journals in which they have published. Having articles in high-impact journals is of critical importance for researchers' promotion and research funding. The term "Impact Factor" has gradually changed to include both journal and author impact. The IF of the journals, in which the articles are published, and the number of times they have been cited are taken into account when evaluating researchers' work. The IF measures the journal's performance by the frequency with which it was cited and is used to compare and rank journals—those with higher IF are

often considered to be of higher quality than journals with lower IF within their subject areas. Journal Citation Reports (JCR) publishes annually the IFs of journals and their relative ranking in different disciplines.

Many articles have discussed the use [but also the abuse] of citations and IFs ^{[17][18][19][20]}. Satyanarayana pointed out the limitations and deficiencies of using citation-based evaluation systems. Some of his criticisms are about the non-transparent policy of journal selection for inclusion in the citation databases and the awkward calculation of the IF. There are many possible scenarios for manipulating citations and IFs. For example, publishing longer papers, controversial articles, and more reviews (which receive more citations than original research papers) will boost a journal's IF. Changing journal content in favor of fields with higher citation rates (such as molecular biology), dropping sections that bring few citations; avoiding publishing articles that are less likely to be cited; and such that do not fall in the editors' view of "novelty" will also increase the IF of a journal ^{[17][18][19][20]}.

Citations and IFs heavily depend on the citation behavior of individual scientists and vary from discipline to discipline [1,11]. Some of the intentional misuse that can skew citations and IFs include avoiding citing competitors, extensive self-citing, "citation padding" (scholars citing each other to boost their citation numbers), citation bias (for example, not citing articles by authors from some countries), or copying (plagiarizing) citations from other articles. As people tend to read papers from higher-ranking journals, such papers are more likely to be cited, which will also increase the IFs of these journals. While SCI makes it easy for researchers to find literature on past research by looking at the citations included in articles, it allows doing it without making a big effort. Mike Kessler, at ISI, looked at the similarities between individual papers in terms of what these papers were citing ^[21]. He used the term "bibliographic coupling" to measure similarities between individual papers. For example, if Paper A cites a group of papers in common with the papers cited in Paper B, papers A and B have something in common. If the two papers have 100% bibliographic coupling (i.e., both articles cite the same citations), this could suggest a case of plagiarism. Copying citations from other articles could lead to the marginalization of some papers in the field in favor of others. Biased citations, if copied by others, could perpetuate a distorted view of past research. Other factors that could influence citation counts are the online availability of articles and continued citing of articles retracted from journals.

SCI and IFs have caught the attention of those who direct the course of scientific research. Academic administrators closely monitor bibliometric and citation patterns to make strategic and funding decisions. Garfield has often argued against using SCI in science policy, because SCI was designed only for information retrieval. Some governments even provide monetary incentives to researchers who publish in highly ranked international journals, particularly in *Science* and *Nature*. As Garfield has always pointed out, it is inappropriate to apply the IF of a journal to an individual article published in it, because the article may have never been cited. What matters, he said, is the citation count for an individual article and not the IF of the journal.

As it takes time to accumulate citations and have papers published in high-impact journals, the existing criteria for getting positions and funding are unfavorable to researchers at an early stage of their careers. Kun shows how the advantage of heritage, social status, and good education are likely to secure successful careers. He is also concerned that the "publish or perish" principle could lead to "... an excessive focus on the publication of groundbreaking results in prestigious journals. But science cannot only be groundbreaking, as there is a lot of important digging to do after new discoveries—but there is not enough credit in the system for this work, and it may remain unpublished because researchers prioritize their time on the eye-catching papers, hurriedly put together". The pressure to publish quickly in high-impact journals could tempt researchers to engage in unethical practices, most common of which are plagiarism and the fabrication of results. As pointed out by many authors, the number of incidents of scientific fraud is increasing ^{[22][23][24]}. Retraction Watch, a web site monitoring the retraction of articles from scientific journals, shows that even reputable journals had to retract articles because of authors' misconduct ^[25].

4. Facilitating the Globalization of Science

Through his writings and the selection of journals for the SCI, Garfield actively promoted English as the *lingua franca* of science and became a powerful force in the globalization of research. Many of his essays, as well as his articles in *The*

Scientist were devoted to convincing the world that science needs a common language and that this language should be English (Figure 6). The selection of journals for the SCI and SSCI favors English-language journals, which has worked against content published in other languages.



Figure 6. An article by Eugene Garfield promoting English as the international language of science published in *The Scientist* magazine. Reproduced with permission from *The Scientist*.

Garfield's efforts to promote English have not always been welcomed by other countries. In an article that he published in the French journal *La Recherche*, he attributed the low citation rate of French articles to the fact that French researchers publish their work in French (Figure 7a). This conclusion and the provocative title of the article ("Is French Science Too Provincial?") caused a storm in the academic and political circles in France in the 1970s. A former prime minister of France, Michel Debré, angrily responded to Garfield's challenge and called it a "linguistic imperialism". The French reacted by launching a new multi-language chemistry journal (Figure 7b).



Figure 7. Garfield published an article in the French journal *La Recherche* under the provocative title "Is French Science Too Provincial?"

(a) A picture of the translation of this article in English. (b) The French reacted by launching a new multi-language chemistry journal.

Having English as the dominant language of science has its pros and cons. We have one language in which we can easily communicate with peers from other countries. This gives, though, native speakers of English the benefit of writing and speaking in their own language that they learned when they were kids, while all others have to make an enormous effort to learn it. In an environment where there is a fierce competition for publishing in top-tier journals, non-native speakers

have to spend time learning how to write in English, while native speakers could use that time to do research. Many articles are rejected from scholarly journals just because of their poor English. Gingras and Khelifaoui ^{[26][27]} assessed “the significance of a parameter that is seldom taken into account in evaluation studies: the existence of a USA comparative citation (visibility) advantage built in the database [SCI] and thus affecting countries that collaborate more with the USA than with other countries”. They analyzed “... how this USA citation advantage affects the measure of the scientific impact (usually measured through citations received) of major countries ...”. They concluded that, “... given the strong presence of the USA in the WoS database, the comparative rankings tend, by construction, to give a citation advantage to countries having the closest relation to that country.”

In his book, *Scientific Babel* ^[28], Michael Gordin discussed the evolution of scientific languages and how English became the dominant language of science. Until the early 20th century, scientific writing was evenly split between English, French, and German. When America emerged as a global scientific power after World War II, English gradually became the dominant language of scientific communication ^{[28][29]}. Garfield contributed to the wider use of English as the international language of science. The top scientific journals are now published in English, and the official language of most conferences is English. Citation networks based on SCI connect researchers and allow them to find new partnerships.

Garfield's vision for the future of science, scientific communication, and information technology has become a reality, but he could not have foreseen some of the unexpected ramifications of his ideas and their impact on the global scholarly ecosystem, where fierce competition for publishing in high-impact journals and obsession with citations and impact factors would permeate so deeply the culture of research.

References

1. Baykoucheva, S. From the Science Citation Index to the Journal Impact Factor and Web of Science: Interview with Eugene Garfield. In *Managing Scientific Information and Research Data*; Baykoucheva, S., Ed.; Chandos Publishing: Sawston, CA, USA, 2015; pp. 115–121. <http://hdl.handle.net/1903/19169>
2. Garfield, E. “Science Citation Index”—A New Dimension in Indexing. *Science* 1964, 144, 649–654, doi:10.1126/science.144.3619.649.
3. Prathap, G. Eugene Garfield: From the metrics of science to the science of metrics. *Scientometrics* 2018, 114, 637–650, doi:10.1007/s11192-017-2525-5.
4. Baynes, G. Scientometrics, bibliometrics, altmetrics: Some introductory advice for the lost and bemused. *Insights* 2012, 25, 311–315, doi:10.1629/2048-7754.25.3.311.
5. Brin, S.; Page, L. The Anatomy of a Search Engine. Available online: <http://infolab.stanford.edu/~backrub/google.html> (accessed on 12 April 2019).
6. Satyanarayana, K. Impact factor and other indices to assess science, scientists and scientific journals. *Indian J. Physiol. Pharmacol.* 2011, 54, 197–212.
7. Brody, S. Impact factor: Imperfect but not yet replaceable. *Scientometrics* 2013, 96, 255–257, doi:10.1007/s11192-012-0863-x.
8. Eyre-Walker, A.; Stoletzki, N. The Assessment of Science: The Relative Merits of Post-Publication Review, the Impact Factor, and the Number of Citations. *PLoS Biol.* 2013, 11, e1001675, doi:10.1371/journal.pbio.1001675.
9. Kun, Á. Publish and Who Should Perish: You or Science? *Publications* 2018, 6, 18. <https://www.mdpi.com/2304-6775/6/2/18>
10. Moed, H.F.; van Leeuwen, T.N. Impact factors can mislead. *Nature* 1996, 381, 186.
11. Baykoucheva, S. Interview with Eugene Garfield. *Chem. Inf. Bull.* 2006, 58, 7–9. <http://hdl.handle.net/1903/11412>
12. Baykoucheva, S. Eugene Garfield's legacy and its impact on the culture of research. In *Abstracts of Papers; 255th ACS National Meeting & Exposition*, New Orleans, LA, USA, March 18–22, 2018.
13. Rumsey, E. Eugene Garfield: Librarian* & Grandfather of Google — Seeing the Picture. Available online: <https://blog.lib.uiowa.edu/hardinmd/2010/07/12/eugene-garfield-librarian-grandfather-of-google> (accessed on 22 May 2019).
14. Lawlor, B. The Institute for Scientific Information: A Brief History. In *The Future of the History of Chemical Information*; McEwan, L.R., Buntrock, R.E., Eds.; American Chemical Society: Washington, DC, USA, 2014; Volume 1164, pp. 109–

15. Baykoucheva, S. What it looked like to work at the Institute for Scientific Information (ISI): Interview with Bonnie Lawlor. *Chem. Inf. Bull.* 2010, 62, 1, 17-23. <http://hdl.handle.net/1903/13955>
16. Garfield, E. *Citation Indexing—Its Theory and Application in Science, Technology, and Humanities*; Wiley: New York, NY, USA, 1979.
17. Huang, P.C.C. Citation Indexes: Uses and Misuses. *China* 2018, 44, 559–590, doi:10.1177/0097700418796778.
18. Elsaie, M.L.; Kammer, J. Impactitis: The impact factor myth syndrome. *Indian J. Dermatol.* 2009, 54, 83–85.
19. Pendlebury, D.A. The use and misuse of journal metrics and other citation indicators. *Archivum Immunologiae Et Therapiae Experimentalis* 2009, 57, 1–11, doi:10.1007/s00005-009-0008-y.
20. Vanclay, J.K. Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics* 2012, 92, 211–238.
21. Kessler, M.M. Bibliographic coupling between scientific papers. *Doc.* 1963, 14, 10–25, doi:doi:10.1002/asi.5090140103.
22. Steen, R.G. Retractions in the scientific literature: Do authors deliberately commit research fraud? *Med. Ethics* 2011, 37, 113–117.
23. Baykoucheva, S. Ethics in scientific publishing. In *Managing Scientific Information and Research Data*; Baykoucheva, S., Ed.; Chandos Publishing: Sawston, CA, USA, 2015; pp. 19–28.
24. Van Gunsteren, W.F. The Seven Sins in Academic Behavior in the Natural Sciences. *Chem. Int. Ed.* 2013, 52, 118–122, doi:10.1002/anie.201204076.
25. Retraction Watch. <http://retractionwatch.com>
26. Kirchherr, J. A PhD should be about improving society, not chasing academic kudos. *The Guardian*, International Edition, 9 August 2018. Available online: <https://www.theguardian.com/higher-education-network/2018/aug/09/a-phd-should-be-about-improving-society-not-chasing-academic-kudos> (accessed on 5 February 2019).
27. Gingras, Y.; Khelifaoui, M. Assessing the effect of the United States' "citation advantage" on other countries' scientific impact as measured in the Web of Science (WoS) database. *Scientometrics* 2018, 114, 517–532, doi:10.1007/s11192-017-2593-6.
28. Gordin, M.D. *Scientific Babel : How Science Was Done before and after Global English*; The University of Chicago Press: Chicago, IL, USA, 2015.
29. Ammon, U. Linguistic inequality and its effects on participation in scientific discourse and on global knowledge accumulation—With a closer look at the problems of the second-rank language communities. *Linguist. Rev.* 2012, 3, 333.