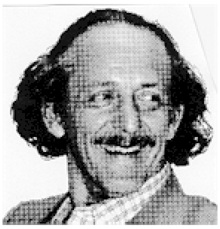
Reprinted in Essays of an Information Scientist, Vol:6, p.616-622, 1983



Current Contents: Its Impact on Scientific Communication

DR EUGENE GARFIELD

Institute for Scientific Information, Philadelphia, Pennsylvania, USA

Beginning in 1957, the literature explosion in science eliminated many personal subscriptions to journals and scientists with interdisciplinary interests demanded new approaches to current awareness. One of these, *Current Contents* now forms a major link in scientific communications. New 'Selective Dissemination of Information' services have helped to mechanize the process of scanning titles for keywords, and new tools for retrospective searching have facilitated historical studies of the scientific literature. The journal literature continues to evolve by means of a process analogous to natural selection.

One characteristic of science which has remained constant from the time of Isaac Newton to the present is the published word. It will remain so in the foreseeable future. Use of the published word is crucial because it forms a permanent, public, accessible record. The existence of this record allows the community of scientists to examine not only each others' results, but also the way in which the results were produced; it allows detailed and informed criticism. The use of the published word is what makes science a body of public knowledge.

Long before the advent of scientific journals, individual scientists communicated their findings to others by means of private letters. Until the 17th century, letters and books were the primary tools of scientific communication. But as scientists gathered together in academies and societies, the scientific journal, at first merely a printed compilation of various scientists' letters, began its development.

THE LITERATURE EXPLOSION

In the first half of this century, a simple model of the scientific information dissemination system could be constructed. It consisted essentially of a relatively small number of primary journals and a few abstracting journals. Chemists had personal subscriptions to several leading chemistry journals and *Chemical Abstracts*; physicists subscribed to several primary physics journals and the Physics Section of *Science Abstracts*, and biologists supplemented their biology journals with *Biological Abstracts*. Of course scientists also used several journals to which they did not subscribe personally. These usually could be found in a nearby departmental or institutional library.

Beginning about 1957, this simple model changed significantly. The change was due in large part to two important events: the advent of *Current Contents* and the decision by the American Chemical Society to eliminate personal subscriptions to *Chemical Abstracts*. At the same time, abetted by generously funded health and space research programs, the scientific literature's already rapid rate of growth accelerated.

The literature explosion caused the bulk of abstracting publications to increase until they became too formidable for even the most conscientious reader. These events combined to bring about the decline of the large abstracting journal as a current awareness tool. Indeed, both before and after 1957 most users of *Chemical Abstracts* had quite narrow areas of interest in applied chemistry, and they were often satisfied by its sectional approach. Current awareness was not very important to them.

THE BRADFORD DISTRIBUTION

The above model of scientific communication in the pre-literature explosion days is, of course, incomplete. A full description must also take into account the means by which scientists obtained primary documents, the original research reports that form the links between scientists all over the world.

As far back as 1934, Samuel Bradford[1] had described a phenomenon, first in electrical engineering and then in other scientific fields, which was to have important implications for individual scientists as well as libraries. Bradford observed that a small number of core journals accounted for a large percentage of the papers published on any given topic. An ideal Bradford distribution would be one in which 10 journals accounted for one-third of the references, 100 journals for the next one-third, and 1000 journals for the remaining one-third.

The Bradford formula may be used to predict the proportion of the literature subtended by a given number of journals for a given scientific field. However, since the distributions have been found to

CCC-0308-0188/79/0004-0318$03.00/1

vary considerably from field to field, it is necessary to specify the appropriate constant for the given field in order to use Bradford's formula. A vast literature on Bradford's law of scattering has been developed, especially during the past decade.

As Bradford showed, even in the preliterature explosion days, scientists were already using many journals outside their own personal subscriptions. In those days it was common for a researcher to go to a nearby library to examine articles in ancillary primary journals. After reading an article relevant to his work, the scientist would usually write a note to the author requesting a reprint of the article. This was almost a matter of courtesy. Among older and more traditional scientists, especially in the United Kingdom, it is still thought bad manners to request a reprint unless one has actually read the article being requested.

THE IMPACT OF *CURRENT CONTENTS*

With the acceptance of *Current Contents* many scientists became time conscious. They could no longer tolerate the excessive delays in gaining access to journals in libraries or by means of abstracts. It was not uncommon in those days for *Chemical Abstracts* or *Biological Abstracts* to list articles over a year old. I remember one article from *Chemische Berichte* which was abstracted three years late – and was not retrievable through the indexes to *Chemical Abstracts* for another two years.

Today, the performance of the leading abstracting services has improved enormously. But for some researchers, even a few months delay may be intolerable, especially in competitive fields like molecular biology, pharmaceuticals or polymers. In fact, it was industrial research that gave *Current Contents* its initial impetus. It began as an idea that merely expanded my experiences in undertaking research for the Army Medical Library, now the National Library of Medicine, on its *Current List of Medical Literature* (now *Index Medicus*).[2,3] Alan MacWatt had started a similar in-house service at Lederle Laboratories which was widely distributed to academics as well.[4]

The existence of *Current Contents* is now taken for granted by most scientists. It is difficult, therefore, for them to realize how gradual was the process by which it became a significant part of the structure of scientific communication. Although its influence increased imperceptibly over a period of almost 20 years, it now forms a major link in scientific communication, helping to launch new journals, affecting requests for reprints, and changing the publishing and citing habits of many scientists.

Even today, most of its regular readers also scan their own personal copies of the core journals in their fields. The increasing cost of subscriptions may have caused some cutbacks, but in most cases both membership and personal-use rates, as well as tax deductions, allow the individual scientist to continue to receive as many as a dozen personal subscriptions. In cases where the individual can no longer afford to buy his own subscription, the journal may be acquired cooperatively for departmental groups or libraries. It is difficult to imagine the typical biochemist buying personal subscriptions to the 10 leading journals in his field, as several yearly rates approach $1000. However, individual scientists continue to subscribe to the core journals. And this is why, contrary to the earlier fears of publishers, very few personal subscriptions were lost through the advent of *Current Contents*. On the contrary, as publishers discovered, it has in many cases promoted subscriptions, especially to new journals. And there have certainly been many new journals published during the past 20 years.

Subscription rates so high that they discourage personal subscriptions pose a profound dilemma. As long as we publish journals, it is usually more efficient to maximize personalized distribution. This should continue until it is no longer possible for paper to compete effectively with other means of disseminating information in bulk. I have, therefore, promoted the idea of the multicopy subscription to journals, and adopted it for ISI's *Index Chemicus*. Certainly, every library or department which subscribes to an important journal should order one or more copies for routing to the staff. It will also reduce the irritations over copyright and indiscriminate photocopying.

COPYRIGHT CLEARANCE CENTER

Since 1 January, 1978, there has been more than mere irritation over indiscriminate photocopying as on that day the revised United States copyright law came into effect and the newly formed Copyright Clearance Center (CCC) became operational. (See preliminary page of this copy of *Interdisciplinary Science Reviews*).

DR EUGENE GARFIELD is the president and founder of the Institute for Scientific Information. ISI produces a broad spectrum of publications, services, and systems designed to channel the flood of published scientific and technical information. His education includes a BS in Chemistry, a MS in Library Science from Columbia University, and a PhD from the University of Pennsylvania. His doctoral dissertation showed how the computer and modern linguistics could be combined in the automatic indexing of chemical information.
Address: Institute for Scientific Information, 3501 Market Street, University City Science Center, Philadelphia, Pennsylvania 19104, U.S.A.

The new law, Public Law 94-553, supersedes the Copyright Act of 1908, which had remained substantially the same since its origin. The current revisions are major. They reflect the impact that technological advances of the last 68 years have had on unauthorized reproduction of copyrighted materials. However, the new law may be far from perfect and it may raise as many questions as it answers; but, there is no doubt that copyright infringement is now a criminal offense in the USA.

Whatever the final outcome of the CCC, ISI's *Original Article Tear Sheet* (*OATS*) service will continue unhampered in its role as a last resort source for the full text of the articles covered by our services. The Institute for Scientific Information has been paying royalties to publishers voluntarily for many years and will continue to do so through its own contractual arrangements. We believe the publisher should get fair compensation, but we also believe the public's right of access to knowledge should be protected.

In my opinion, there are several facets to that protection. One is that publishers should charge reasonable fees for the right to copy their articles. In addition, publishers should make it convenient to obtain information on copying charges and equally convenient to pay them. The CCC is a major step in that direction. But can a single center fully serve the varied needs of book and journal publishers?

Copyright considerations aside, it should be remembered that the main concern of the typical *Current Contents* reader is not just economic coverage of the core journals in his field. The core journals could be covered, albeit less conveniently, by timely sectional abstracting coverage or by visits to the library. The typical reader also wants coverage of those journals outside this core. And as Bradford's formula would predict, it is not uncommon for this list to exceed 100 journals. Certainly, it is unreasonable to expect an individual to subscribe to so many journals.

I have been concerned about the power attributed to *Current Contents*. Some publishers have stated that ISI's decision to list or not to list a journal can mean the difference between success and failure. *Current Contents* makes readers promptly aware not only of journal articles they might otherwise never have seen, but also of the journals themselves.

In fact, it may have altered the appearance of the Bradford distribution for many fields, because scientists are now citing a much larger repertoire of journals than they did formerly. A suitably controlled study might attempt to investigate this notion. This could be done by examining citation patterns within journals through ISI's *Journal Citation Reports*.[5,6] The citation habits of individual scientists could also be studied to see whether *Current Contents* has had a significant impact. In the *Science Citation Index* one can often observe the diversity of journals citing a particular author.

THE REPRINT EXPLOSION

It has now become typical of the *Current Contents* reader to write to colleagues requesting reprints before reading the requested article. These requests are usually material published outside the core journals to which they subscribe. At first there was some resentment, but this has diminished as authors realize that we all face the same problem; namely, without reprints we must photocopy. Reprint requests are a source of income for publishers and an important communication feedback mechanism for authors.

Assuming there are at least 200 000 readers of *Current Contents*, we estimate that they request as many as 10 million reprints per year. The average would be 50 per year per reader. Some readers request less than one reprint per week, while others request more than 50. That so many reprints are used is indicated by the number of reprints purchased by scientists. There is also an interesting correlation in terms of literature citations. Several studies have shown that more than half of all reprint requests are the result of *Current Contents* readership.

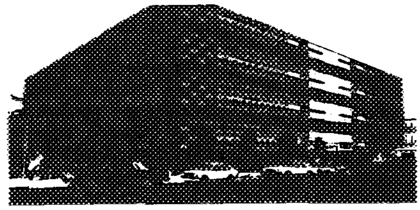
And in addition to reprints there is the growing use of photocopies. The British Lending Library has reported that over 20 percent of its requests for current photocopies are derived from *Current Contents*.[7] This is a substantial impact considering that this Library now supplies over one million photocopies per year, a legitimate source of concern to journal publishers.

THE MATTHEW EFFECT

Sometimes a primary purpose of publication is to achieve the prestige associated with acceptance in a large circulation journal. Ironically, many of these journals are so large that one's article may be buried among 50 to 100 separate contributions. The mere appearance of an article in the *Journal of the American Chemical Society* or the *Physical Review* may look impressive on a curriculum vitae. But that fact alone will not cause the article to be read or cited. Indeed, a smaller journal listed in *Current Contents* may receive more attention simply because its contents page is not as formidable to scan.

The most significant research tends to be submitted to the highest prestige journals. This so-called 'Matthew' effect means that the best journals receive the best manuscripts.[8] But occasionally extremely high quality work bypasses the prestige journals in favor of new or less prestigious journals. For example, a group of scientists may break away from the pattern described by the Matthew effect and a new high prestige journal will emerge. This was the case with *Tetrahedron*, where the established chemical journals did not exhibit enough flexibility in meeting the needs of even highly esteemed scientists.

In other cases, a large backlog of mediocre manuscripts can stand in the way of a significant breakthrough. Since priority in scientific discovery is still important, the breakthrough paper may be submitted to a newer or less prestigious journal for quicker publication. In a democratically functioning editorial system, new journals will arise when older journals are too slow or too inefficient to meet authors' needs. That is why, perhaps, there is so much discussion today of the so-called synopsis journal. One such journal is the *IRCS Journal of Medical Science*, published in the UK. Another is the *Journal of Chemical Research* published by the Chemical Societies of the United Kingdom, Germany and France.



The Institute's new headquarters, in Philadelphia's University City Science Center, began to be used by 400 ISI employees towards the end of October 1979; the building was architectured by Venturi and Rauch.

SELECTIVE DISSEMINATION OF INFORMATION

Current Contents Life Sciences covers about 4000 items per week, published in about 1000 journals – but the total worthwhile scientific literature is closer to 10 000 items per week appearing in several thousand journals. Most of these 10 000 items will not be of interest to life scientists, but there is always the occasional item for each reader which must be discovered by other means. Some may even have to scan the chemical and physical editions.

That is why many readers also use a Selective Dissemination of Information (SDI) service that can provide a back-up to regular reading and scanning. In addition, subject indexes can be used to pick up articles from journals in that long tail of the Bradford curve – those 1000 or more journals which only occasionally contain relevant articles. This is provided through the *Weekly Subject Index*. Other scientists may do this by using *Chemical Titles*, *Bioresearch Index* (BIOSIS) or *Current Physical Titles*.

The use of subject index can be augmented by various SDI services, such as ISI's *Automatic Subject Citation Alert* (*ASCA*). With this, it is possible for the researcher to learn of any current paper which has cited any of his previously published articles or books. One can also mechanize this process of scanning titles for keywords. Scientists in numerous countries like Canada, Israel, Sweden and Mexico have SDI services which are based partly on the use of the same magnetic tapes that help produce the *Science Citation Index* and the *Social Sciences Citation Index*. They also involve use of tapes provided by the many members of the *National Federation of Abstracting and Indexing Services* and the *Information Industry Association*.

All of these methods for alerting scientists to what is relevant are today supported by formidable retrospective search capabilities. Not only can the scientist himself search such printed indexes as *Chemical Abstracts* and *Science Citation Index*, both of which publish five year cumulations, but it is also possible to do so-called on-line searches at computer terminals located throughout the United States, Europe or elsewhere. Undoubtedly, over the next decade, this network of terminals will be expanded, making retrospective access more convenient and less costly than at present. And in many instances, these terminals will be used for current SDI or alert purposes.

On-line service is now relatively expensive, but most of its cost is based on that of communication and storage hardware. These costs are expected to maintain a downward trend. Eventually, on-line systems must compete with relatively inexpensive mini-computers available to individual research groups. Such mini-computers can store large masses of data, and thereby eliminate communication problems altogether. For instance, it is still necessary to waste time waiting for access to on-line facilities. This is caused by queuing problems and physical difficulties with communication lines. Apart from this, on-line and mini-computer systems may have to compete with microfilm and other micro-storage technologies as they improve.

CITATIONS, PREMATURITY AND POST-PUBLICATION IMPACT

Even after publication, scientists are vitally concerned about the utilization of their research. By citing other authors, we acknowledge that these authors have had some impact on our work. The cumulative assessment of the impact of previous contributions is indicated by citation analyses. But one never knows for certain what causes a particular paper to be cited. Surely, in order for a paper to be cited, the citing author must somehow encounter the paper. This encounter may occur by scanning primary journals, at meetings, through the use of current awareness tools, by retrospective search, or through a referee. A significant percentage of papers are encountered through the references cited in articles we read.

Scientists believe that exposure of an article in a large circulation journal, like locating a gas station at a busy intersection, is going to increase the chances of citation. However, we cannot now assess the effect of a journal's circulation on the citation of an article as opposed to the effect of the journal's coverage by abstracting and indexing services. In many fields like physics and chemistry most scientists prefer to publish in the leading journal of their discipline. In other fields some prefer a prestige journal such as *Nature* or *Lancet*. Nevertheless, there is no evidence that circulation or prestige alone increases the chances of citation.

In fact, a large percentage of the literature, both refereed and un-refereed, is rarely or never cited. Even the most prestigious journals publish articles that are never cited, and some produce a staggering number. One possibility is that some of these uncited papers contain premature discoveries.

Do many scientists, like Gregor Mendel, produce significant work which is largely ignored by their contemporaries because it is premature? In order to determine whether any of these uncited papers are premature, we need a system of retrospectively reviewing the literature. This might include a statement by the author explaining why a paper deemed worthy for publication by his peers was so systematically ignored.

Such post-publication impact statements might be valuable not only for authors, but also might help editors and referees establish better criteria for selecting articles. It might also be interesting to learn from authors whether they felt that papers subsequently published by others ought to have cited their work. This might tell us something about the effectiveness of our total information retrieval mechanisms and, in particular, how much of the relevant literature is retrieved by citation indexing.

If it is found that a significant number of papers omit pertinent references, we would have cause for concern. Furthermore, one might expect unwitting duplicate research to increase in the era of information overload, but there is no current evidence to support this expectation. Unwitting duplication may have been more characteristic of earlier decades when we had tight little islands of research everywhere and modern information systems were unavailable.

In fact, historians will particularly want to evaluate the literature published in the first half of this century. And for this reason one may hope that the *Science Citation Index* covering the years 1900–1960 will be constructed eventually. This needs to be done, not only to evaluate the uncited work, but also to resolve the controversies surrounding so many discoveries. For example, Lederberg has commented[9] on the need for citation indexes to help determine the impact of Avery's 1944 work on DNA.[10] Using computerized methodologies, such a large scale data bank will foster numerous historical studies necessary to understand better the changing role of scientific literature.

NATURAL SELECTION IN SCIENTIFIC COMMUNICATION

A few years ago, a group of eminent chemists signed a letter deploring the proliferation of journals.[11] They advocated a system of controlling new journals that, if put into effect, would prove disastrous. Fortunately, this letter has been ignored and the journal system has so far remained essentially free and competitive. The authors of this letter failed to take into account many of the points noted above. As new areas of science emerge, and some of them do so quite rapidly today, the journal production system should remain flexible. In spite of rising paper and postal costs, the journal is still an efficient means for distributing primary scientific information. It will remain so if we do not force existing journals to accept articles which have limited readership.

Consider what would have happened if, by legislation or otherwise, a cartel of biochemical journals had prevented publication of *Prostaglandins* or some similar specialized biochemical journals. Subscribers to the core biochemical journals would have been forced to pay for numerous prostaglandin papers that could not otherwise get published. The editors of core journals would have been inundated with prostaglandin papers that now logically go to this new journal. The natural process whereby journals are born, or die, as science evolves, would have been frustrated.

Unfortunately, except under the most extreme conditions, we do not let journals die often enough. Euthanasia is as relevant for journals as it is for people. Even though the amount of journal literature increases, we should never discourage the death or transformation of journals which cannot satisfy modern needs. In particular, the death sentence ought to be passed on dozens of journals that never meet a publication deadline. This process of journal evolution, if not tampered with by government or others, will produce an efficient system of distribution of printed information. Eventually, depending upon the cost and supply of paper, electronic or other media may become more relevant. This presumes that the cost of electronic storage will continue to decrease.[12]

Scientific communication, unlike other forms of human communication, requires the careful, deliberate examination of ideas written down for perusal by peers. The timing is different from that which takes place in a television interview or even at a scientific meeting. Undoubtedly, oral exchange of ideas through discussion speeds up the development of new ideas. But in order to remain paramount in developing new scientific ideas and theories, the individual scientist must commit his ideas to written form, so that peers can examine them in exquisite, deliberate detail. This distinguishes the scientific process from the adversary process of debate.[13]

If ever the scientific communication system degenerates to nothing more than oral intercourse, then, I believe, we will not recognize it any longer as science. Motivation for the individual will also disappear. Whether this kind of oral think-tank science is as satisfying to the participants as is the present system remains to be seen. It may be suitable to mission oriented research but not to basic science.

CONCLUSIONS

The two simple models of scientific communication described here have evolved, one from the other, in just two decades. It is reasonable to expect that equally significant changes may take place in the next two decades. While the more dramatic scientific breakthroughs may be reported even in the popular press, the nuts and bolts of scientific discovery will continue to be reported in the journal literature. What form this literature takes – whether the print and paper of the Gutenberg era or the electronic circuitry of the McLuhan era – remains to be seen.

In either case, I question whether it will be the medium of communication that determines the scientific message. Rather the medium affects the speed of transmission and access and thereby may alter the quality and timeliness. All of these changes in scientific journals will improve our ability to understand nature through our uniquely human ability to communicate through language.

But no matter what form develops for journals there will continue to be a need for the functions now performed by *Current Contents*. So it may not be too outlandish to suggest that a daily, electronic version may be what scientists will need to keep abreast in the future.

LITERATURE CITED

1. S. C. Bradford, *Documentation*, Public Affairs Press, Washington (1950).
2. S. V. Larkey, the Welch Medical Library Indexing Project. *Bull. Med. Libr. Assoc.* **41**, 32–40 (1953).
3. H. G. Field, W. A. Himwich, E. Garfield, J. M. Wittock and S. V. Larkey, Final report on subject headings and on subject indexing. Welch Medical Library Indexing Project, sponsored by the Armed Forces Medical Library, p. 226. The Johns Hopkins University, Baltimore (1955).
4. J. A. MacWatt, Improving scientific communication; reprints directly available from publisher at a reasonable fee could supplement today's journals. *Science* **134**, 313 (1961).
5. E. Garfield, Citation analysis as a tool in journal evaluation. *Science* **178**, 471–479 (1972).
6. E. Garfield, Significant journals of science. *Nature (London)* **264**, 609–615 (1976).
7. D. N. Wood and C. A. Bower, The use of biomedical literature at the National Lending Library for Science and Technology. *Methods Inform. Med.* **9**, 46–53 (1970).
8. R. K. Merton, The Matthew effect in science; the reward and communication systems of science. *Science* **199**, 55–63 (1968).
9. J. Lederberg, Letter to the editor in reply to H. V. Wyatt. *Nature (London)* **239**, 234 (1972). (H. V. Wyatt. When does information become knowledge? *Nature (London)* **235**, 86 (1972).)
10. O. T. Avery, C. McLeod and M. McCarty, Studies on the chemical nature of the substance inducing transformation of pneumococcal types; induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. *J. Exp. Med.* **79**, 137–158 (1944).
11. C. J. Ballhausen, F. A. Cotton *et al.*, Too many chemistry journals. *Chem. Eng. News* **51**(50), 43–44 (1973).
12. E. Garfield, Is there a future for the scientific journal? *Sci-Tech News* **29**, 42–44 (1975).
13. A. Strick, What's wrong with the adversary system: paranoia, hatred, and suspicion. *The Washington Monthly* **8**, 19–28 (1977).

The manuscript was received 20 March 1978