

Chapter 8

Strengths and weaknesses of citation indices and impact factors

Alessandro Figà-Talamanca¹

Dipartimento di Matematica, Università di Roma “La Sapienza”, Piazzale Aldo Moro 5, 00185 Roma, Italy

Citations as the currency of science

According to Eugene Garfield², who had the remarkable ability to make a very lucrative business out of recording citations, the idea that citations are the currency through which scientists pay other scientists goes back to the Sociologist of Science Robert K. Merton, who expressed this opinion in a private conversation with Garfield, in 1962. In Garfield’s words, “the Mertonian description of normal science describes citations as the currency of science. Scientists make payments, in the form of citations, to their preceptors”.

For many years, however, it was not at all clear how this currency could be converted into dollars and cents (or pounds or euros). The conversion became possible when citation counts became a measure of scientific quality; indeed, a measure actually used to decide on promotion and hiring and to judge grant applications.

A convertible currency?

It is difficult to tell when and how this change occurred. But I do remember the context in which I heard for the first time that the number of citations could have an influence on someone’s career or salary. It was in 1968 or 1969, when I was a junior faculty member in the Department of Mathematics of the University of California at Berkeley. A mathema-

tician from an Eastern European country had just given a ‘colloquium talk’ and, during the party that followed the lecture, I heard him explicitly beg his colleagues to cite his work in their papers. He claimed that, in his country, the number of citations was used to determine the salary of scientists, and asked his Western colleagues the personal favour of citing his work. The implication was, of course, that in this Eastern European country, in the context of a totalitarian regime, closely controlled by a foreign power, promotions and salary levels were decided by people who did not have the competence to actually examine the scientific productions of scientists. The bureaucrats who decided on promotions were chosen for their loyalty to the regime and not for their scientific competence. When it came to considering scientific production (which, after all, was still one of the criteria for promotion), they could only rely on numbers: the number of papers and the number of citations. We thought at the time (I remember discussing the problem with American colleagues) that, in our ‘open society’, the scientists themselves could decide on promotions and hiring of other scientists, on the basis of the worth of the results, as judged by competent people. This made citation counts irrelevant. We were, of course, completely wrong.

In our defence, it must be said that we were still living in the magic (for Western science) world of the 1960s, a world that, for many people today, must be difficult to even imagine.

¹email: sandroft@mat.uniroma1.it

²Eugene Garfield was for many years the majority shareholder and the chief executive officer of the Institute of Scientific Information (ISI), a company which collects data on citations and sells access to these data to scientific institutions. Garfield is also a master of marketing for his products, whose virtues are extolled by pieces of writing which have the appearance of ‘scientific papers’ in the new science of ‘scientometrics’. An example of such writing is mentioned in the article *The Use of Journal Impact Factors and Citation Analysis for Evaluation of Science* [1], available at <http://www.garfield.library.upenn.edu>.

To give another example of how scientific communication worked at the time, I will quote a recollection, by Michael Taylor (now William R. Kenan Jr Professor of Mathematics at the University of North Carolina), of the advice received from his mentor, Tosio Kato, when he was a graduate student at Berkeley in the autumn of 1968. The recollection is published in *Notices of the American Mathematical Society* [2]: “He [Tosio Kato] told me I should learn interpolation theory. He put me on to the recent work of J.-L. Lions and E. Magenes. Their books, in French, on boundary problems were not out yet, and of their papers some were in French and some in Italian. But a \$2 paperback on Italian for beginners from Moe’s Bookstore helped me make them accessible”.

For the record, Enrico Magenes, now Emeritus Professor, was at the time Professor at the University of Pavia, while Jacques Lions was Professor at the University of Paris. They published in what would be called, nowadays, ‘local’ journals, writing in French or Italian, and, of course, they expected that an American graduate student would go through the trouble of learning the elements of Italian or French to read and understand their results. At the time, word-of-mouth citation of actual results, communicated by a mentor to his graduate students, was more important than citation counts. Senior scientists, at least in mathematics, were oblivious to the ‘prestige’ of the journals and published important results wherever it was most convenient, often in their own language.

Today’s different world

We live now in a different world. It is still difficult to imagine that citation counts would be used to hire, for instance, a permanent member of the Institute for Advanced Studies. If the choice is restricted to the best 20 scientists in the world, it is unlikely that greater or fewer citations, much less parameters such as the impact factor, would make a difference.

But outside a few extreme cases, citations, in one way or another, have become a currency which is convertible into dollars and cents. Of course the ‘exchange rate’ is far from being fixed. Notwithstanding the diminishing purchasing value of the dollar, as times goes on, more and more citations are needed

to buy a decent salary. This is, of course, the result of inflation: more and more citations are produced each year, and their value against relatively stable currencies cannot but decrease.

What happened then in the last 40 years which changed the scientific environment so drastically?

To understand where we started from, we must go back for a moment to what I called the “magic world of the 1960s” which was already fading away, under our eyes, in 1968. Indeed, talking about the 1960s is imprecise. The magic decade for science in the U.S.A., and therefore in the Western World, was between 1957 and 1966, that is, between the launch of the first Sputnik and the explosion of the protest of university students (effectively and largely backed by the professors) against the war in Vietnam. During this magic decade politicians, but also the general public, became convinced that the development of science at all levels, from high school teaching to the top research institutions, was the only possible answer to the military, technological and political challenge represented by the Soviet bloc. For the brief span of about 10 years, the ‘military–industrial complex’ allied itself with the academic community to demand more investment in scientific research, including basic research, including even the training of scientifically competent secondary school teachers. This unnatural alliance between the military and the academic community in the U.S.A. dissolved at the end of the 1960s, over the differences caused by a war which was very unpopular in the campuses and sent many young students to die in the Vietnam marshes. A signal of the changed mood of society was a law, passed by a bipartisan initiative, motivated by opposite ideologies, which forbade the military to finance basic research, as they had been doing throughout the 1960s, by means of grants generously offered by ‘research offices’ of the Army, the Air Force and the Navy. By means of this law, the Right wanted to punish the academic community for their position on the war, while the Left, on the other hand, wanted to prevent the military from having an influence on the scientific community and on the campuses.

The end of the magic decade brought about the issue of ‘accountability’ of science and scientists, to society in general. In the last analysis, the same forces which were at work in the 1960s in Eastern

Europe, and made it necessary to seek 'objective' parameters to judge the quality of scientific work, became relevant in our society. We still live in an 'open society', but our society is no longer willing to sign blank cheques to scientists for their research, their salary and their promotions, without even asking what they are really doing. 'Accountability' has become a byword which is universally accepted. But one should not suppose that scientists are only accountable to experts of the same discipline. At some level in the hierarchy, science is supposed to be accountable to other 'stakeholders', who may not be, and generally will not be, competent in the specific research under scrutiny. Objective parameters, such as citation counts, are an easy way out for everybody concerned, because they can be used without understanding anything of the research under judgement. It is also an easy way out for scientists, because, as the recent explosion of impact factors in the scientific production of biomedical science in Italy proves, it is also much easier to improve on the impact factors than on the actual depth and relevance of the results.

Another important factor which was part of the change, and contributed to it, is the growth in the number of scientists, of scientific papers, and scientific journals. Forty years ago it was not difficult for a senior scientist to keep abreast of all the important novelties in his field. This made it easy to judge the new contributions to the field of a young person to be hired or to be promoted. Globalization and specialization have made this almost impossible.

Very few objections may be raised against the need of accountability to whoever supports scientific research and to stakeholders in general. It is also difficult to object to the need of supporting the discretionary judgement of experts with arguments which may be understood by non-experts. Citations, or better still a careful citation analysis, may be the right answer to this need.

The role of scientific publishers

Growth in the number of scientists, globalization and specialization, and accountability to stakeholders are not the only changes which occurred in the last 40 years. An invisible shift of

power took place, which was strictly connected to the other changes, but was also the consequence of the pursuit of legitimate commercial interests, with aggressive and well-planned marketing strategies. What happened is that the responsibility to judge on the quality of research which was formerly vested on the scientists themselves, was progressively entrusted to commercial publishers of scientific journals.

The scientific community effectively abdicated, in favour of commercial publishers, the duty and power to judge on the quality of scientific research. Independently of the use of 'bibliometric indicators', nowadays it is almost impossible to talk about the quality of the scientific production of a scientist without mentioning the place of publication of his work.

In order to understand this phenomenon, and the forces which brought it about, we should start from the observation that scientific literature is a commodity, produced by scientists and consumed by scientists. Modern technology has, in fact, greatly reduced the cost of composition and printing. But scientists are employed by the same institutions which, at ever increasing prices, buy scientific publications. It follows that scientific institutions, such as universities, bear the cost of production of scientific literature, because they pay the salaries of scientists and the costs of their research, and then pay again to buy the journals from commercial publishers. It would seem that, under these circumstances, 'in-house' production of scientific literature would be more cost-effective and convenient. Even the threat of in-house production could be a deterrent against the unreasonable high price of scientific publications. But this threat cannot even be voiced.

A widespread opinion, fostered by scientific publishers, and 'proved' by indirect counts of citations such as the impact factor, has sentenced, now that in-house production of scientific literature is synonymous with low quality. This, of course, is a strong deterrent for anybody wishing to publish his own papers. As a consequence, whatever in-house production of scientific literature existed has quickly disappeared over the last few decades. In addition, commercial publishers were very successful in creating many new journals with prestigious 'boards of editors', by flattering the vanity, and the desire to be influential, of many senior scientists.

We see now that the effect of the progressive entrustment to the publishers of scientific journals of the authority to judge the quality of scientific research resulted in a commercial victory in favour of the multinational companies which control the market of scientific publications. This victory seems, at the moment, definitive, even though the growth of informal communication through electronic publishing may succeed in changing the picture in a few years. We are left to conclude that the same commercial interests which gained from this process were also at work in making this change possible.

It is in this purely commercial context that one should consider the role of the Institute for Scientific Information (ISI), whose interests merged with the interests of publishers and distributors of scientific literature. In particular, we should consider the role of the so-called impact factor which was advertised, and eventually imposed, as an indicator of the quality of a journal, a scientific article and, eventually, of an individual scientist. The impact factor of a journal is defined as the average number of citations that articles published by the journal receive for the following 2 years by articles published in a selected list of journals. Its direct relationship with minimum standards of quality of the journal is, at best, unproved. But impact factor is almost universally considered to be a measure of the quality or the prestige of a journal. This means that many authors attempt to publish their work in journals with high impact factors. As a consequence, the editors have more papers to choose from and the publishers may increase the number of pages and their profits. We see at work a self-feeding process of validation, which seems impossible to interrupt.

Of course, it could be argued that the power of judging the worth of a scientific paper and its publication in a prestigious journal is still exercised by the editors who are experts in the field and respected members of the scientific community. But publishers, who appoint editors, are interested in profit, and profit is enhanced by increasing the prestige of a journal. A prestigious journal cannot be ignored by a major scientific library, practically independently of its costs. As long as the impact factor is considered to be a measure of prestige, publishers are justified in demanding that editors try to increase

the impact factor of the journal. An editorial policy which aims to increase the impact factor may not be in the interest of science. As a rather extreme example, editors have been known to ask authors to cite recent articles in the same journal.

Thus the relevant question is not whether the impact factor actually reflects the quality of the papers published in a scientific journal. For the future of science, it is much more relevant to observe that relying on the impact factors to judge the scientific production of an individual or an institution effectively serves the purpose of creating a market distortion to the benefit of commercial publishers.

Strengths and weaknesses of the impact factor

We have come to the point in which we should draw some conclusion and, perhaps, respond to the title of this chapter, "Strengths and weaknesses of impact factors and citation indices".

The impact factor is undoubtedly very strong. It succeeded in imposing itself as a measure of quality, by simply asserting its value. This assertion cannot be disproved, and, as long as it is taken for granted, it produces effects. The strength of impact factor is proved beyond any doubt by the effects it produced in the scientific environment. Indeed, it is very hard to point to a single weakness of the impact factor. But perhaps a meeting on evaluation of research is not the right context for a discussion of the strength of the impact factor; it should be discussed in a meeting on business strategies. To appreciate the prodigy achieved by the impact factor, one should think of the benefits that the distributors of ready-to-eat food would gain if, through an allied 'independent' evaluator, they succeeded in 'proving' that homemade food is unsavoury, or unhealthy. The case of the impact factor is, of course, stronger, because it has a direct effect on both producers (scientists) and buyers (institutions) of scientific literature.

We conclude that the impact factor shows no weakness if its strength is related to the purpose for which it was created, which was to enhance the value of the databases owned by ISI and to create a market distortion to the benefit of the commercial

publishers of scientific journals. On the other hand, I believe that not a shred of evidence may be given to prove that the introduction of the impact factor as a measure of quality of journals, articles and individuals, brought any benefit to science.

We must say, however, that it is impossible to turn the clock back. The 'good old times' will never come back. We should look forwards instead. The phenomenon which may put an end to the market distortion fostered by the impact factor is the development of electronic publications. We may foresee a future in which scientific papers become generally available through the internet site of scientific institutions, or rather a special site in which only articles which are guaranteed by the institution to be relevant and correct are posted. This type of in-house production and posting of 'guaranteed preprints', freely available also to scientists working in institutions which cannot afford to meet the rising costs of subscription to all prestigious journals, may eventually be able to compete in the market of scientific publications, and oblige the publishers to reduce their prices. The guarantee of fame afforded by an institution may eventually be stronger than the stamp of approval of a prestigious editorial board and a high impact factor. In any case, the article would be available for free. To make such a development possible, it is imperative that scientific institutions retain the right of electronic publication of the work of the scientists they employ. The 'lobby' of publishers is working in the opposite direction, under the guise of protecting the interests of the authors. We should all be aware of where the real interests of scientific authors and of science lie.

What about citation indices?

We have talked up to now of the impact factor which is a derivative product of actual citations. But citations may refer not only to a journal, as in the case of the impact factor, but also to a single paper, to an author, to small or large groups of authors. Citation counts have been used to assess the scientific production of entire countries.

There is no doubt that access to a good database of scientific articles and citations may be of great help in judging the quality of the scientific

production of a researcher. Citations, at least those citations which are not mere inclusions in an endless list of contributors to the same area or problem, may provide a measure of how the results achieved by an author were received by other scientists. To verify the impact of the author's ideas, one should go back to the citing papers and distinguish, at the very least, between favourable citations, unfavourable citations and mere inclusion in a list of contributors. Unfortunately, impact factors and bare citation counts have made it more difficult to use appropriately a database of citations. Only a fraction of actual citations of a paper say something about the paper. In many cases, it is evident that the work of the cited author was not even read and that the citation was added by pasting a list copied from somewhere. Still real citations, which say something about the results of the cited paper, by a scientist who does not appear to be in the same group of cronies as the author cited may be very useful to evaluate the results of a paper and the authors who contributed to these results. All other citations, which are by far the majority, are practically of no use, except to testify that the author cited belongs to a group of researchers working on the same topics.

But the mere counting of citations is unreliable for another reason. The most popular databases of citations, owned by the ISI, is very unreliable in terms of correct identification of the authors and their institutions. There is no database of authors and institutions. The object that is recorded is only a name, with no attempt to relate the name to an individual. People who are familiar with a database which identifies authors and not just their names, like the databases of the American Mathematical Society, can appreciate the difference. Even abbreviations in the names of journals are often misinterpreted, a phenomenon which produced gigantic variations in the impact factor of an astronomical journal which shared the most common abbreviation with journals of other disciplines.

Citation counts have also been used to assess the scientific production of scientists in a given country. Here, we must beware of the fact that international comparisons are very difficult. The example of evaluation of scientific production in Italy is very interesting. For many years, we heard

high cries that scientific production in Italy was very low, in comparison with other developed countries. The usual measure was the ratio between papers published in 'international journals' (that is journals in the ISI databases) and the resident population of Italy. Sometimes the number of citations was used as the numerator with the population as the denominator. At some point someone, to wit Professor Carlo Rizzuto of Genoa, observed that the denominator should not include all the people living in a given country, but only those people who were employed in research. An application of this observation was enough to move Italy from the bottom score to a very good position, better than other comparable industrial countries. The choice between one denominator or the other is actually arbitrary. In Italy, there is no legislation which makes it convenient for a company to classify some

of its activity as research and some of its employees as researchers. As a consequence, there is very little investment in research, and practically no researcher is employed in the private sector. A change in tax legislation would instantly produce claims that certain activities are indeed research and certain employees are researchers. This exemplifies the intrinsic weakness of international comparisons, independently of the bias of the sources and the inadequacies of the databases.

References

- 1 Garfield, E. (1998) *The Use of Journal Impact Factors and Citation Analysis for Evaluation of Science*, 41st Annual Meeting of the Council of Biology Editors, Salt Lake City, UT, U.S.A., 4 May 1998
- 2 Cordes, H., Jensen, A., Kuroda, S.T., Ponce, G., Simon, B. and Taylor, M. (2000) Tosio Kato (1917–1999). *Notices of the American Mathematical Society* 47(6), 650–657