Scientometrics (2012) 92:391–393
DOI 10.1007/s11192-012-0678-9

Decades of progress, or the progress of decades?

Francis Narin

Received: 2 February 2012 / Published online: 21 February 2012

Abstract In the almost 40 years since we wrote Evaluative bibliometrics enormous advances have been made in data availability and analytic technique. The journal impact factor of the 1960s has clearly not kept up with the state of the art. However, for both old and new indicators, basic validity and relevance issues remain, such as by what standard can we validate our results, and what external use can appropriately be made of them? As funding support becomes more difficult, we should not lose sight of the necessity to again demonstrate the importance of our research, and must keep in mind that it is the relevance of our results that count, not the elegance of our mathematics.

Keywords Journal impact factor · Science policy · Validation

At the suggestion of Professor Braun I am writing this essay to take an overview of the progress that has been made, and the challenges that remain, in our field. In particular in the almost 40 years since we wrote Evaluative bibliometrics (Narin 1976) enormous advances have been made in data availability and analytic technique. However, basic evaluation issues remain, such as by what standard should we measure our results, and what external use can appropriately be made of them? Both of these issues are crucial to any analysis of the journal impact factor, or any other bibliometric measure. The recent paper by Vanclay (2012) shows in great detail the technical limitations of the impact factor, and the considerable challenge of upgrading it to a useful, valid indicator.

Clearly, one of the prime motivations for the development of our field has been the need for funding agencies and governments to measure the quality and productivity of the scientists they support. An acceptable indicator should be simple enough that it can be explained to our non-bibliometric colleagues.

Early in the formative years of bibliometrics we were, of course, acutely aware of these problems, both the mechanical problems and the conceptual ones. In the 1970s we dealt exclusively with the Science Citation Index and were aware of the limitations of the impact

F. Narin (✉)
2075 Mountain City Street, Henderson, NV 89052, USA
e-mail: narinf@cox.net

factor as it was defined then, and developed the influence (eigenvector) methodology to overcome some of the limitations.

The available data processing power has increased by tens of thousands over the last 40 years. The analysis techniques today are far more mathematically sophisticated. The citation data are far more widely available. But arriving at a solution to the fundamental problem does not seem to be any easier today. To quote from Evaluative bibliometrics,

> "*Older measures of influence all suffer from some defect which limits their use as evaluative measures.*
> *The total number of publications of an individual, school or country is a measure of total activity only; no inferences concerning importance may be drawn.*
> *The total number of citations to set of publications, while incorporating a measure of peer group recognition, depends on the size of the set involved and has no meaning on an absolute scale.*
> *The journal "impact factor" introduced by Garfield* (1972) *is a size-independent measure, since it is defined as the ratio of the number of citations the journal receives to the number of publications in a specified earlier time. This measure, like the total number of citations has no meaning on an absolute scale. In addition, the impact factor suffers from three more significant limitations. Although the size of the journal, as reflected number of publications, is corrected for, the average length of individual papers appearing in the journal is not. Thus, journals, which publish longer papers, namely review journals, tend to have higher impact that. In fact the nine highest impact factors attained by Garfield* (1972) *were for review journals. This measure can therefore not be used to establish a "pecking order" for journal prestige.*
> *The second limitation is that citations are un-weighted, all citations being counted with equal weight regardless of the citing journal. It seems more reasonable to give a higher weight to a citation from a prestigious journal than to a citation from a peripheral one. The idea of counting a reference from a more prestigious journal more heavily has also been suggested by Kochen* (1974).
> *A third limitation is that there is no normalization for the different referencing characteristics of different segments of the literature: a citation received by a biochemistry journal, in a field noted for its large number of references and short citation times may be quite different in value from a citation in astronomy, where the overall citation density is much lower and citation time lag much longer.*"

We then went on to describe the influence methodology, which we first applied to the literature of physics (Pinski and Narin 1976).

Almost all of the above points are still being addressed and discussed today, without any ultimate solution in sight.

There are some analogies here to the problem of modeling financial behavior discussed in Derman's (2011) recent book "*Models Behaving Badly*", in which he contrasts the application of mathematical modeling techniques to physics and to finance. In the application to physics the mathematical models represent what is essentially an immutable physical world, to which their outcome can be quantitatively and precisely compared. In contrast, the same mathematical techniques applied to the financial world assume as immutable a very mutable world, governed by human behavior. On the jacket of his book is the statement "*why confusing illusion with reality can lead to disaster, on wall street and in life.*"

In bibliometrics the systems we are modeling are also governed by human behavior, and sometimes also small enough to be gamed by the participants: for example, by altering journal policies to increase the apparent impact factor of a journal, or by creating formal or informal self-citation groups.

Nevertheless, the question of modeling the scientific community is important, and research and modeling should be encouraged, while keeping in mind that the target was and is an elusive and changing one.

Almost 50 years ago Weinberg (1963) anticipated the problems that the growth of "big science" would provide for the policy analyst. He states that:

> "*As science grows its demands on our society's resources grow. It seems inevitable that science's demand will eventually be limited by what society they can allocate to it. We shall have to make choices. These choices are of two kinds. We shall have to choose among different often incommensurable fields of science—between, for example, high-energy physics and oceanography or between molecular biology and science of metals. We shall also have to choose among the different institutions that receive support for science from the government—among universities, governmental laboratories and industry. The first choice I call scientific choice: the second institutional choice. My purpose is to suggest criteria for making scientific choices—to formulate a scale of values which might help establish priorities among scientific fields whose only common characteristics is that they all derive support from the government.*"

The strength of publication and citation analysis lies in its flexibility to meet the small-scale demands of the bibliometric community, as well as its ability to encompass the much larger scale needs of the science analysts.

From the point of view of the world of 2012, almost all of the concerns and questions addressed in the 1970s, are still being considered, analyzed and addressed. The techniques today are far more precise, and the available data for more vast. But the underlying questions are still there. What is quality science? How do you measure it? How should it be supported?

These questions will become even more important in the difficult economic times facing the West. We should just not expect absolute answers to emerge, and must remember that it is the relevance of our results that count, not the elegance of our mathematics.

References

Derman, E. (2011). *Models behaving badly*. New York: Free Press.

Garfield, E. (1972, Nov, 3). Citation analysis as a tool in journal evaluation. *Science, 178,* 472–479.

Kochen, M. (1974). *Principals of information retrieval*. New York: Wiley.

Narin, F. (1976). *Evaluative bibliometrics*. New Jersey: Computer Horizons, Inc.

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: theory, with applications to the literature of physics. *Information Processing and Management, 12*, 297.

Vanclay, J. K. (in press). Impact factor: Outdated artifact or stepping stone to journal certification? *Scientometrics*.

Weinberg, A. M. (1963). Criteria for scientific choice. *Minerva, 1*, 159–171.