1. Citation Classics: The Idea and the Collection

ALEX C. MICHALOS

After about three years in graduate school studying the history of religions, my interests turned to logic and the philosophy of science. My doctoral thesis was on a dispute between Karl Popper and Rudolf Carnap over the construction of quantitative measures of the acceptability of scientific theories (Michalos, 1971). For most of the 40 years that I was a university teacher, I taught courses in the philosophy of science, always including issues related to the evaluation of the activities and products of social and natural scientists. Since there are no generally accepted definitions of such key concepts as science, scientific explanation, scientific laws, scientific theories and scientific acceptability, but there are particular groups of researchers that tend to accept certain concepts while rejecting others (Michalos 1980), I take a fairly pragmatic approach to the construction of my own scientific vocabulary. For present purposes, the only part of that vocabulary that requires explanation concerns the term ‘citation classic’.

I first encountered the term some years ago reading something by Eugene Garfield, the founder of the Institute for Scientific Information (ISI), the Science Citation Index, Social Sciences Citation Index and several other important works. According to Garfield (1985), “By definition, a Citation Classic is a paper or book that has been highly cited in its field” (p.404). From 1977 to 1993 the ISI publication Current Contents regularly featured an article called Citation Classics Commentary, which was subsequently replaced by In-Cites. Because different fields are characterized by different institutional arrangements (e.g., numbers and kinds of communication media, practitioners, standardized practices and rules of procedure), different numerical thresholds are used to identify classics in each field. In an earlier paper, Garfield (1976) wrote that

“...less than 25% of all papers will be cited ten times in all eternity! ... any paper cited ten times in one year is ipso facto significant. Occasionally there is an anomaly. But a paper cited ten times in each of two successive years is well on its way to citation stardom. Whether the author is on the way to immortality depends on how well he or she does in other papers” (p.419).

An average paper in the Science Citation Index is cited about 1.7 times per year (Garfield, 1972). For the years 1955-1987, Garfield (1989) claimed that for the whole Index database,

“...more than 56 percent of the source items are uncited – not even self-cited. (Many of these source items are abstracts, letters, and editorials, of limited interest; nevertheless, a huge number of papers go uncited.)” (p.7).

Alex C. Michalos (ed.), Citation Classics from Social Indicators Research, 1–56.
In 1973 Garfield distinguished three kinds of “uncitedness”, namely, “the incitedness of the mediocre, the unintelligible, the irrelevant”, then that “of the meritorious but undiscovered or forgotten”, and finally that

“of the distinction that comes to those whose work has become so well known (and presumably been previously so heavily cited) that one finds it at first tedious, then unnecessary, and finally actually gauche to cite such men at all” (p.413).

Hamilton (1990) reported that ISI data revealed that about 55% of the “papers published between 1981 and 1985 in journals indexed by the institute received no citations at all in the 5 years after they were published” (p.1331). In response, Pendlebury (1991) wrote that the precise figures were “47.4% uncited for the sciences, 74.7% for the social sciences, and 98.0% for the arts and humanities”. However, more importantly, he explained that

“These statistics represent every type of article that appears in journals indexed by the Institute for Scientific Information (ISI) in its Science Citation Index, Social Sciences Citation Index, and Arts & Humanities Citation Index. The journals’ ISI indexes contain not only articles, reviews, and notes, but also meeting abstracts, editorials, obituaries, letters like this one, and other marginalia, which one might expect to be largely un-cited. In 1984, the year of the data quoted by Hamilton, about 27% of the items indexed in the Science Citation Index were such marginalia. The comparable figures for the social sciences and arts and humanities were 48% and 69%, respectively.

If one analyzes the data more narrowly and examines the extent of uncited articles alone (this information was not yet available when Hamilton wrote his articles), the figures shrink, some more than others: 22.4% of 1984 science articles remained uncited by the end of 1988, as did 48.0% of social sciences articles and 93.1% of articles in arts and humanities journals. Only 14.7% of 1984 science articles by U.S. authors were left uncited by the end of 1988. Articles published in the highest impact journals like Science are almost never left uncited” (pp.1410-1411).

Exhibit 1 illustrates a few thresholds for citation classics from different fields, ranging from papers with 50 or more citations covering the fields of geography and marine biology to 500 or more citations covering all fields listed in the Science Citation Index. Plomp (1990) used a threshold of 25 citations to identify “highly cited papers”, which is a more modest label than ‘citation classic’ and perhaps not quite the same idea. His rationale for the figure was interesting.

“Considering the (average) number of references in a paper as its ‘input’ and the number of citations achieved by that paper as its ‘output’, the ratio citations/references may be interpreted as the ‘gain factor’ of the paper; it sounds reasonable that a gain factor of 1 is a sort of watershed between papers recognized by the scientific community as important and papers not recognized as important. As the average number of references in scientific articles is about 20 (according to the 1986 SCI), I consider