Communications of the Association for Information Systems

Volume 25 Article 22

7-1-2009

Journal Self-Citation XXII: On the Journal Impact Factor – A Historical Perspective

David E. Drew
Claremont Graduate University, David.Drew@cgu.edu

Communications of the Association for Information Systems

CAIS

Journal Self-Citation XXII: On the Journal Impact Factor – A Historical Perspective

David E. Drew

Claremont Graduate University

David.Drew@cgu.edu

Abstract:

This paper puts a historical perspective on the use of Impact Factors. It describes how Impact Factors were used in a U.S. National Research Council project 35 years ago to evaluate the improvements that resulted from a billion dollar University Science Development program funded by the National Science Foundation. Impact Factor rankings proved to have a remarkably high correlation to science departments ranking obtained from a different source. The paper considers additional policy aspects of impact factors.

Keywords: Impact Factors, department rankings, policy impacts

Volume 25, Article 22, pp. 177-182, July 2009

I. INTRODUCTION

This paper presents a perspective about the Thomson Reuter Scientific Impact Factor based upon personal policy evaluation experience 35 years ago. The objective is to put the quantitative findings in the first five papers (Holsapple [2009], Li [2009],Romano [2009] Palvia et al. [2009] and Straub and Anderson [2009] in the special set of CAIS articles on journal self-citation into historical perspective.

The context in which the Impact Factor evaluation was done was a policy evaluation of the use of National Science Foundation funds to improve the quality of scientific research in higher education. The idea was to find out whether there had been improvements over the previous 10 years and, if so, to what degree they could be traced to the NSF funding.

The evaluation involving Impact Factors was undertaken in a world where few, if any, tried to game the values of the Impact Factors and at a time when there were many fewer academic journals. What is intriguing is that the Impact Factor approach, which resembles analysis of social networks today, proved to be a remarkable predictor of how universities and university departments rank on quality.

I directed this large evaluation study at the National Research Council (NRC) [1] in 1973 - 1974, 35 years ago. This study was one of the earliest applications of this measure, if not the first, for public policy analysis [Drew, 1975].

The study was conducted in a unit called the National Board on Graduate Education (NBGE), housed within the NRC. We evaluated the major National Science Foundation (NSF) funding program of that era, Science Development, which often was referred to as the "Centers of Excellence" program. Between the mid-1960s and the mid-1970s, Science Development awarded 231 million dollars (approximately 1.25 billion dollars in 2009 dollars).

This paper is divided into five parts, including this introduction:

- The funding program
- The evaluation study
- Measuring the quality of scientific research using the Impact Factor
- Conclusions about journal self-citation

II. THE FUNDING PROGRAM

The purpose of this program was to increase the scientific productivity and excellence of institutions outside the top twenty. The existence of this program grew out of a perennial debate about the distribution of federal research funds, a debate that had first surfaced in the 1940s when the NSF was created. The project began with an assumption that everyone shared: the distribution of research funds should maximize scientific productivity and quality. Some argued strenuously that it was obvious that research funds should go to the best scientists and that, furthermore, those scientists were to be found exclusively in the top institutions. Others argued for geographical and institutional equity, also on the grounds of excellence. They contended that funding decisions betrayed an elitist bias towards the top universities (roughly half the federal funding in each scientific discipline was being awarded to the top 20 institutions) and that scientific talent was more widely dispersed. Furthermore, they contended that creative undergraduate and graduate students attending lower-ranked universities were less likely to enter scientific research careers unless they worked with professors who were active, funded researchers.

Perhaps the major impetus for the Science Development program was a report called *Science Progress, the Universities, and the Federal Government* issued by the President's Science Advisory Committee [1960]. As is often the case with government reports, this statement became known as the Seaborg report, after Glenn T. Seaborg, the

[1] The National Research Council (http://sites.nationalacademies.org/nrc/index.htm) is a part of the National Academy of Sciences of the United States.

chair of the committee. Among their recommendations was the following: "…increase support for rising centers of science. Over the next fifteen years the United States should seek to double the number of universities doing generally excellent work in basic research and graduate education." (p. 28)

This notion, first articulated when Eisenhower was still president, was subsequently endorsed and expanded by both the Kennedy and the Johnson administrations. The program took more definite form during Johnson's administration and was consistent with his philosophy of geographic diffusion of funds. A 1965 Executive Order on that subject was closely tied to the disbursement of NSF funds for university science development.

Science Development involved three subprograms.

1. The major component was University Science Development (USD), under which 31 institutions received 177 million dollars;

2. Special Science Development, under which 11 universities that had been denied USD support received limited funding (about a million dollars each) for a few promising departments; and

3. The Departmental Science Development subprogram, which awarded grants (about $600,000 each) to individual departments that showed promise, but were housed in less highly -ranked universities, often urban universities.

The central philosophy was embodied in the USD subprogram. The emphasis was on systemic institutional support, not project support.

I should note that there often is a discontinuity between Washington politics and evaluation rigor. The top 20 research universities were excluded from funding under this program for obvious reasons. But the NSF administrators were quite cautious and never actually named the excluded universities. So there was the hilarious spectacle of leading research universities searching for some measure that would define them as outside the top 20, so that they could then compete for these (then) massive funds. The most successful of those leading institutions was the University of Washington, which, by my calculations, then ranked about eighth in the nation, and was funded under USD. The largest Departmental Science Development grant went to Yale! And so it goes in Washington.

This debate about the concentration of federal science funding has not gone away. Last month the American Association of Universities, which consists largely of leading research institutions, issued a statement urging that, in these economically stressed times, federal research funding should be focused on the leading universities [Basken, 2009]. Otherwise, they contended, the nation's research standing might suffer.

III. THE EVALUATION STUDY

My staff and I at the National Research Council were tasked with conducting a comprehensive evaluation of Science Development. This evaluation included:

- field visit case studies of both funded and non-funded institutions, and
- statistical analyses of quantitative data.

For the field visits, we assembled teams of scientists. Recruiting scientists to assist the National Academy of Sciences is not difficult (for example, our team included national leaders like John Bardeen--who twice won the Nobel Prize for physics and was co-inventor of the transistor--policy experts, institutional leaders, and evaluation specialists). We interviewed university administrators, faculty and staff, and, sometimes, students at each university.

We amassed a data base containing longitudinal indicators about graduate departments in physics, chemistry, and mathematics (the main disciplines that were funded, given the era), and a control field, history. We collected information from primary sources and from existing data bases on a number of indicators of science department structure, functioning and success. The data spanned a 15-year period which included years before, during, and after Science Development funding.

We built multivariate models to assess the unique impact of Science Development funding on important outcomes, while controlling for baseline differences between institutions. With the help of several experts on the project advisory panel (i.e., Donald Campbell, one of the founders of the discipline of evaluation, and Lincoln Moses, president of the American Statistical Association and graduate dean at Stanford), we developed a special analysis and presentation technique for comparing group time series residuals [Drew 1983].

The project report appeared as a book, *Science Development: An Evaluation Study*, along with an accompanying National BNBGE policy report [National Board on Graduate Education, 1975]. While we were objective and included criticisms (e.g., there was virtually no concern expressed for undergraduate education and the scientific pipeline at the funded institutions), we reported dramatic changes and quantum leaps in research quality at a number of institutions. (For example, the program funded three universities in close proximity in North Carolina; this support catalyzed the development of the North Carolina Research Triangle. The Astronomy department of the University of Arizona jumped to a top position in national rankings, in great part because of USD funding.) These changes were reflected in significant impacts we found in the national data when we conducted the multivariate analyses.

IV. MEASURING THE QUALITY OF SCIENTIFIC RESEARCH USING THE IMPACT FACTOR

We struggled with how to measure the quality of scholarly productivity. We considered publication counts, citation counts, various weighting schemes for publications (e.g., a book was worth four articles), and the like. We developed a measure of departmental rate of publication in the top 20 journals in the field. To do so, we defined the top 20 journals by their Impact Factor. Then we simply counted how many publications per year appeared in these journals written by scientists at our sample universities.[2] We summed the counts at the departmental level, i.e., the graduate department was the unit of analysis.

One day I asked Ron Karpf, a member of my staff and a crack computer programmer, to run a multiple regression analysis. We had peer review ratings of the science departments from 1970 done by KD Roose and CJ Anderson of the American Council on Education [1970]. I wanted to see which of our many departmental variables might predict them. The results were simply astounding. We did not successfully estimate a *multiple* regression equation because only one variable was significant—and it was powerful! It was our publication variable based in part on the Impact Factor. The multiple correlation coefficient, which, of course, was equal to the Pearson correlation coefficient, was .87. We noticed that the data showed an inverse exponential curve. So we performed a transformation, after which the linear correlation coefficient increased to .91.

As you know, one rarely obtains a correlation that high in social science research. Ron and I continued this work after I left the NRC and joined the Rand Corporation. We published the results in a Rand Corporation paper [Drew and Karpf, 1975] and in a peer-reviewed journal article [Drew and Karpf, 1981]. Since five years had passed since the 1970 national peer review ranking of academic departments by Roose and Anderson, and since our Impact Factor measure correlated almost perfectly with their previously published departmental rankings, we modestly included tables showing what the new rankings would be, based on the latest available publication data. I received an angry letter from the head of the national association of applied mathematicians complaining that my top 20 list of mathematics journals unfairly favored theoretical mathematics. I replied that we relied objectively on the ISI impact factor and that it appeared clear that articles in theoretical journals were cited more frequently.

II. CONCLUSIONS ABOUT IMPACT FACTORS

This set of articles deals with coercive use of self-citations by journals. But we should note that there are other forms of coercion and citation inflation. Some authors extensively self-cite. Some senior professors coerce junior colleagues into citing their work. Increasingly, in an economically and technologically flat world, we will see comparisons of scientific productivity across nations. Will national scientific and academic leaders coerce scholars into excessive citations of work from their own country?

The move by Thomson Reuters Scientific to publish Impact Factors with and without credit for journal self-citation is welcome. A similar index that removes from the Impact Factor self-citation by authors of their work in other journals would also be ethically desirable.

We should differentiate healthy or legitimate journal self-citation from coerced journal self-citation. Straub and Anderson [2009] observe that there may be good reasons for non-coerced higher rates of citation of a leading journal as that is where published work has "made a splash." As Romano [2009] notes, this is a classic example of Merton's Matthew Effect [Merton, 1968] , the theory of accumulated advantage; the rich get richer. But, unlike coerced citation, it is an ethical and legitimate process.

[2] It was a different era; I hired research assistants who spent all day reading through paper journals and counting institutional affiliations to determine the department publication counts.

Having said that, we still need to consider that, throughout society, people find ways to beat the system. If a perfect policing system renders journal self-citation fruitless, two journals could collaborate to increase their impact factors by coercing citations to one another.

We know that not all citations reflect quality. For example, there are negative citations, e.g., “Jones totally misunderstood my theory.” I have a colleague who is one of the best in his field. He was one of the early adopters of a new multivariate statistical method and he made a mistake while using it. For years after that, authors warned their readers not to make that mistake and cited my colleague’s article. Although he was embarrassed by it, he enjoyed a high citation rate.

We also know, as Holsapple [2009] points out, that impact factors are averages based on citation rates that vary considerably by journal, by discipline, and, of course, by author. Thus individual articles contribute differently to the impact factor. Consider the raw data from the five articles discussed here, plus my own, on impact factors. The number of references listed in these six articles is : 2,5,9,15,24,75. The mean of these numbers is 21.7 and the median is 12. Neither of these measures captures the complexity of even this small distribution, as indicated by its range of 2 to 75.

Palvia, Palvia, and Baqir [2009] report that journal self-citation among the top three IS journals is frequent and exceeds self-citation by non-top-tier journals. They note that ISR had no citations of *Decision Science*, *Information and Management*, or *Journal of the Association for Information Systems*. They ask, “Is the set of top tier journals an elite club?”

Their question brings my own reflections full circle. I would note the irony that it was a parallel observation that top funded universities formed an elite club that was at the heart of the Seaborg Report [President's Scientific Advisory Board, 1960] and was the impetus for the NSF Science Development program (Section II) that my team evaluated using Impact Factor data (Section IV).

I believe that it is dangerous for important resources in research, whether funding or citations, to be limited to an elite club. As Thomas Kuhn pointed out in *The Structure of Scientific Revolutions* [Kuhn, 1962], science progresses when creative scholars question the fundamental assumptions of an existing paradigm. Such creative, free-spirited inquiry is less likely to occur among the true believers who publish only in the leading journals.

REFERENCES

Basken, P. (2009). "U.S. May Need to Prune Number of Research Universities, Lobby Group Says," *The Chronicle of Higher Education*, June 26.

Drew, D.E. (1975). *Science Development: An Evaluation Study*, Washington, DC, National Academy of Sciences Press.

Drew, D.E. (1983). "Advanced Ocularmetrics: Graphing Multiple Time Series Residual Data," *Educational Evaluation and Policy Analysis* (5)1, Spring, pp. 97-105.

Drew, D.E. and R. Karpf (1975). *Evaluating Science Departments*: *A New Index*, Santa Monica, CA: The Rand Corporation, P-5521, October, 37.

Drew, D.E. and R. Karpf (1981). "Ranking Academic Departments: Empirical Findings and a Theoretical Perspective," *Research in Higher Education* (14(4).

Holsapple, C. W. (2009). "Journal Self-Citation II: The Quest for High Impact – Truth or Consequences," *Communications of the Association for Information Systems* 25(1), Article 2, pp. 11-20.

Kuhn T. (1962). *The Structure of Scientific Revolutions,* Chicago, University of Chicago Press.

Li, E. (2009). "Journal Self-Citation III: Exploring the Self-Citation Patterns in MIS Journals, *Communications of the Association for Information Systems* 25(1), Article 3, pp. 21-32.

Merton, R.K. (1968). "The Mathew Effect in Science" *Science* 159 (3810) January 10, pp. 56-63.

National Board on Graduate Education (1975). *Science Development, University Development and the Federal Government*, Washington, DC: National Academy of Sciences, 48 pp.

Palvia, P., S. Palvia, and M.N. Baqir (2009). "Journal Self-Citation IV: Citations Analysis of IS Journals – Separating Fact from Fiction," *Communications of the Association for Information Systems* 25(1), Article 4, pp. 33-40.

President's Science Advisory Committee (1960). *Science Progress, the Universities, and the Federal Government* Washington, DC: The White House.

Romano, N. (2009). "Journal Self-Citation V: Coercive Journal Self-Citation: Manipulations to Increase Impact Factors May Do More Harm than Good in the Long Run," *Communications of the Association for Information Systems* 25(1), Article 5, pp. 41-56.

Roose, K. D. and C.J. Anderson (1970). *A Rating of Graduate Programs*, Washington, DC: American Council on Education.

Straub, D. W. and C. Anderson (2009). "Journal Self-Citation VI: Forced Journal Self-Citation – Common, Appropriate, Ethical?" *Communications of the Association for Information Systems* 25(1), Article 6, pp. 57-66.

LIST OF ABBREVIATIONS

IF Impact Factor

NBGE National Board on Graduate Education

NSF National Science Foundation

NRC National Research Council

USD University Scientific Development project

ABOUT THE AUTHOR

David E. Drew holds the Joseph B. Platt chair in the management of technology at the Claremont Graduate University. His main appointment is in Education; he also holds appointments in Management and Mathematical Sciences. He has assessed faculty scholarly productivity as a researcher, as a member of the CGU appointment, promotion and tenure committee, and as a dean. Early in his career, he designed information systems as the head applications programmer at the Harvard Computing Center.

Communications of the Association for Information Systems

ISSN: 1529-3181