Jointly published by Akadémiai Kiadó, Budapest
and Kluwer Academic Publishers, Dordrecht

Scientometrics,
Vol. 56, No. 2 (2003) 223–233

Can scientific impact be judged prospectively? A bibliometric test of Simonton's model of creative productivity

JOHN D. LEE,[1] KIM J. VICENTE,[2] ANDREA CASSANO,[2] ANNA SHEARER[1]

[1]*Cognitive Systems Laboratory, Department of Industrial Engineering, University of Iowa (USA)*
[2]*Cognitive Engineering Laboratory, Department of Mechanical & Industrial Engineering, University of Toronto (Canada)*

Simonton's (1997) model of creative productivity, based on a blind variation-selection process, predicts scientific impact can only be evaluated retrospectively, after recognition has been achieved. We test this hypothesis using bibliometric data from the *Human Factors* journal, which gives an award for the best paper published each year. If *Simonton*'s model is correct, award winning papers would not be cited much more frequently than non-award winning papers, showing that scientific success cannot be judged prospectively. The results generally confirm *Simonton*'s model. Receipt of the award increases the citation rate of articles, but accounts for only 0.8% to 1.2% of the variance in the citation rate. Consistent with *Simonton*'s model, the influence of the award on citation rate may reflect a selection process of an elite group of reviewers who are representative of the larger peer group that eventually determines the citation rate of the article. Consistent with *Simonton*'s model, author productivity accounts for far more variance in the authors' total citation rate (58.9%) and in the citation rate of the authors' most cited article (12.6%) than does award receipt.

Introduction

Is the impact of a scientific publication a foreseeable trait that can be envisioned prospectively at the time of publication, or is it an unpredictable trait that can only be identified retrospectively after recognition has been achieved? Most scientists would probably like to think that they can recognize a scientific "hit" when they see one. After all, the ability to judge discriminatingly should reflect the professional expertise that is a natural outgrowth of a career as a scientist – or so the argument goes. In contrast, *Simonton*'s (1997) model of creative productivity assumes that scientific creativity "… is to some significant degree blind or haphazard. This means that at some crucial level the individual has no a priori way of foreseeing which ideational combinations will prove most fruitful" (p. 67).

Received July 16, 2002.

Address for correspondence:
JOHN D. LEE
2130 Seamans Center, Department of Industrial Engineering,
University of Iowa, Iowa City, IA, 52242, USA
E-mail: jdlee@engineering.uiowa.edu

0138–9130/2003/US $ 20.00

This Darwinian approach views the scientific process as a variation-selection process (*Campbell*, 1960; *Hull*, 1988). Some ideas and articles will be useful and well-received but others will not, and the successful ones cannot usually be identified from the outset because the variation-selection process operates at three levels: the individual cognitive, the interpersonal, and the socio-cultural. Because selection occurs at multiple levels, selection at the individual cognitive level must be blind to its ultimate success at the interpersonal and socio-cultural levels. Likewise, selection at the interpersonal level must be blind to the socio-cultural factors affecting the ultimate impact of an article. Trends in technology, new scientific findings, and shifts in scientific values make it impossible for reviewers, acting at the interpersonal level, to judge the future socio-cultural selection of important articles (*Simonton*, 1997). This process of blind variation-selection also predicts that the ratio of highly cited works to total articles of an author should be constant and so the number of highly cited papers produced by an author will be a constant ratio of the total papers produced by that author.

In this article, we provide an empirical test of *Simonton*'s (1997) model using bibliometric data from *Human Factors*, the North American flagship journal in the discipline of human factors engineering (or ergonomics). This analysis uses standard bibliometric methods that make some important assumptions. Most importantly, we assume that the citation rate of each article published in the journal can be used as an approximate, retrospective "gold standard" measure of its scientific impact (*Lawani*, 1986; *Zuckerman*, 1977). Although factors such as critique of previous work and social connections motivate citations, citations have been shown to be a robust indicator of scientific impact (*Vinkler*, 1998). Fortunately for our purposes, the *Human Factors* and *Ergonomics Society* has an annual Jerome H. Ely Human Factors Article Award for the most outstanding paper published in each volume of the journal. This award is given in the year following publication, so it can be used as an approximate measure of the prospective judgment of an article's scientific impact. All else being equal, if scientific impact can be assessed prospectively, the award process should be able to identify scientific "hits" reliably. Thus, articles that receive the award should garner more citations than non-award winning articles. However, if scientific impact can only be assessed retrospectively, as *Simonton*'s (1997) model assumes, then the award process will be "to some significant degree blind or haphazard." As a result, articles that receive the award will garner citations at roughly the same rate as non-award winning articles. An empirical test of these predictions thereby provides a rare and valuable opportunity to study scientific activity scientifically (*Hull*, 1998).

Method

The ISI Web of Science electronic database was used to compile the citation history of articles published in *Human Factors*. Data were collected for articles published between 1970 and 2000. A total of 1682 articles were included in the resulting citation database, which recorded: the title of each article, the authors, the date of publication, the number of citations received each year (as of May 2001), and the total number of citations received. This sample has several important features that limit the generality of our findings. First, the database includes only a single journal and so does not reflect the total productivity of many of the authors. Second, it is restricted to 30 years of publications and so it likely does not capture the lifetime contributions of many authors.

From this basic database, two additional databases were developed to support the citation analysis. The author database described the publication and citation history of each of the 2413 authors who contributed to *Human Factors* between 1970 and 2000, inclusive. It described the number of papers each author had written, the number of awards that he/she received, the mean rate of citations per year, and the total number of citations each author has received. This information was cataloged for articles where the author was the first through the seventh listed, as well as for all articles combined. The second database characterized the citation history of each article. It included the title, date of publication, number of citations per year since publication, the mean citation rate per year, whether the paper received an award, and the publication history of each contributing author as described in the author database (articles written, awards received, citation rate, total citations).

Results and discussion

Does the award predict success?

The simplest way to test if scientific impact can be judged prospectively is by comparing the mean annual citation rates for the 30 award-winning articles in our sample with those for the 1652 non-award-winning articles. The comparison showed that the former were cited 1.32 times per year (sd = 1.1), whereas the latter were cited only 0.59 times per year (sd = 0.80). This difference is statistically significant ($F(1, 1680) = 24.15$, $p < 0.001$). Moreover, if we examine the frequency with which award-winning articles are in the upper half of the most-cited articles published each year, we find that more award-winning articles appear in the top 50% (23) than in the bottom 50% (7), a statistically significant difference ($\chi^2(1)=11.20$, $p<0.01$). These two analyses show a modest effect of the award on subsequent citation rate, suggesting that the award process might be a valid prospective measure of scientific impact. If this is the case, then *Simonton*'s (1997) model may be incomplete.

However, there are several strong reasons that argue against adopting such a conclusion. First, the impact of the award accounts for only 1.4% of the variance in the citation data. Furthermore, of the 20 articles with the highest citation rate in our sample, only one – the 11th most frequently cited article – received the Jerome H. Ely Human Factors Article Award. In fact, only one of the 30 articles with the highest citation rates of all of the articles published in each volume of *Human Factors* in our sample received the award. Thus, the prospective judgments of scientific impact are clearly neither completely reliable nor precise.

Figure 1 reinforces this interpretation. It shows the cumulative distribution of the citation rates for the award-winning and non-award-winning articles. The award shifts the distribution slightly to the right, but it does not qualitatively change the distribution of citation rates. There is a substantial difference in the span of the two distributions.

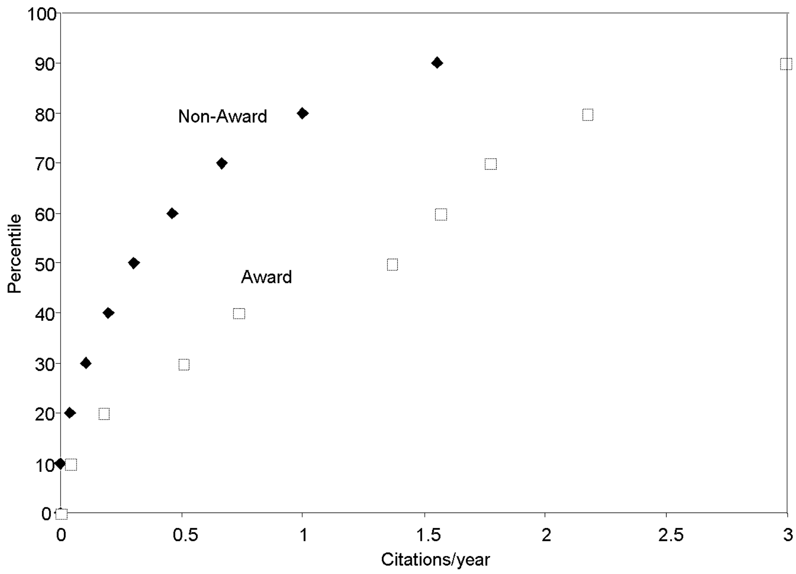

Figure 1. The cumulative distribution of the citation rate of award-winning and other articles

The papers that did not receive an award form a broader distribution. There are many non-award articles that are more highly cited than the award-winning articles – one article that did not receive an award has been cited 133 times and the most an award-winning article was cited is 109 times. There are also award-winning articles that are cited very infrequently – even 27 years after publication one award-winning article has yet to be cited.

So far, the analyses we have presented have compared the award- and non-award-winning papers in terms of citations or mean citation rates. It is also possible to examine the differences among these articles in terms of citation dynamics. Figure 2 shows the citation history of the award and the non-award articles. The award did affect citation rate for each year, as expected from the differences in mean rates identified earlier; however, the general citation dynamics remain very similar. The citation rate of both award and non-award papers increases over the first three or four years after publication, reaching a plateau before declining in the ninth or tenth year. Because only one award is given each year, the number of data points declines as a function of the years since publication. Only 15 award-winning papers have a 15-year publication history. The similarity of the citation dynamics suggests that the award process does not appear to have the foresight to identify papers that have a substantially faster rise to success or that enjoy substantially more enduring success. Figures 1 and 2 both show that the award does have a modest ability to differentiate between more and less influential papers.

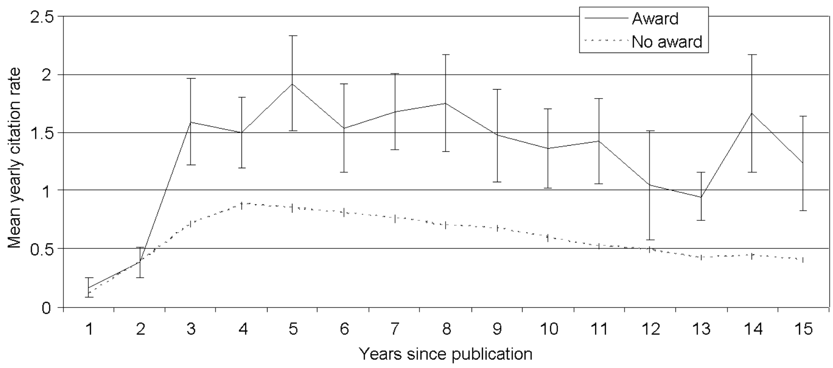

Figure 2. The citation history for award and non-award winning articles.
Bars show 95% confidence intervals

Does paper production of an author predict influence?

Is there a factor, other than award selection, that is a better correlate of scientific success? *Simonton*'s (1997) model predicts that the number of articles written by any given author would be a better predictor of a particular article's eventual citation rate than whether or not it receives an award. Because the selection process is blind, *Simonton* suggests that the quantity and quality of scientific output are highly correlated. Authors who write many articles are more likely to produce a highly influential article and will have more influence in general, as measured by citation rate, than authors who write few articles. Four linear statistical models of article citation rate were constructed to compare how well award receipt and paper production predict scientific influence.

The first model considers whether the number of articles written by an author influences the citation rate of a particular article. *Simonton*'s (1997) model would suggest that the influence of any one particular article would be blind to the number of articles an author has generated in the past; the chance of surviving the selection process is independent of authorship. The second model provides a rough test of *Simonton*'s (1997) contention that "a scientist's total output is the single best predictor of the total number of citations he or she receives in the technical literature" (p. 76). The total number of articles each author published in *Human Factors* between 1970 and 2000 was compared with the total number of citations received for articles published in this journal over the same period. The third model examines the same relationship on a per-article basis. *Simonton*'s model suggests that the number of quality articles divided by the total number of articles is a constant, and so the number of articles written by an author should not affect the average citation rate per article. The fourth model examines the citation rate of each author's most highly cited article. Because *Simonton* argues that the number of quality articles is proportional to the number of articles produced, the total number of articles should relate to the citation rate of each author's most-cited article. In total, these four models explore some of the underlying assumptions of *Simonton*'s model and address how the award and the broader socio-cultural selection process affect an article's influence.

Each of the four linear statistical models was constructed to test first the predictive power of author production (total papers produced by the authors of a paper) and then the effect of the award. The effect of the award was evaluated by examining whether the incremental improvement in the model predictions was statistically significant. This effect was tested using the following equation:

$$F(1, N\text{-}2) = (R_P\text{-}R_{PA})((N\text{-}2)/R_{PA})$$

R_P is the residual sum of squares of the model that includes only author productivity, R_{PA} is the residual sum of squares of the model that includes publications and award, and N is the number of articles or authors in the data set.

The first analysis considered whether the number of papers written by the authors in our sample influences the citation rate of a particular paper. Contrary to *Simonton*'s model, the number of papers written does have a statistically significant effect on the citation rate for individual papers ($F(1, 1680) = 23.61$, $p < 0.001$). However, the size of this effect is very small, with the number of articles written accounting for only 1.3% of the variance. The effect of the award accounts for an additional 1.2% of the variance, which is also statistically significant ($F(2, 1679) = 9.58$, $p < 0.01$). This result suggests that although the citation rate is not completely blind to the award or the authors' past production, these effects are not particularly strong predictors of the citation rate.

The second analysis, where the citation data were considered from the perspective of individual authors, provides a more direct test of *Simonton*'s model. Rather than predicting the citation rate of particular articles, this model predicts the citation rate of individual authors, testing the hypothesis that the best predictor of an author's total citation rate is the number of papers he or she has written. The total number of articles written by an author in our sample has a very strong effect on the overall citation rate for that author ($F(1, 2411) = 3462.0$, $p < 0.001$), accounting for 58.9% of the variance in the total citation rate of each author. Including the effect of the award accounts for a marginally greater proportion of the variance, for a total of 59.1%. Although quite small, the incremental increase is also statistically significant ($F(1, 2410) = 4.28$, $p < 0.05$)). This result is quite consistent with *Simonton*'s model, showing that the total citation rate of any author depends on the number of papers written and that most highly cited authors are also highly productive. Award receipt does not have a major incremental impact on the total citation rate of authors. The negligible contribution of the award is also consistent with *Simonton*'s model, suggesting that peer judgement does not predict citation rate very well. The relationship between papers written and overall citation rate is also consistent with the finding that paper production is strongly related to peer judgment of scientific performance over a career (*Sonnert*, 1995).

The third analysis goes beyond this simple evaluation of total citations and evaluates the effect of articles written and award receipt on the average citation rate (the total citation rate divided by the total number of articles written by the author). In this analysis, like the previous one, the total number of articles written by an author accounts for a statistically significant portion of the variance in citation rate ($F(1, 2411) = 3.06$, $p < 0.05$). The number of articles accounts for only 0.2% of the variance in the average citation rate of each author. The effect of the award accounts for a slightly greater proportion of the variance, for a total of 1.3%, and produces a statistically significant effect on the average citation rate ($F(1, 2410) = 14.47$, $p < 0.001$). These results are consistent with *Simonton*'s model, which argues that the ratio of high-quality articles to the total number of articles is constant. This relationship is consistent with our finding that the average citation rate does not depend on the number of articles written by an author.

The fourth analysis more precisely tests *Simonton*'s model by evaluating the citation rate of each author's most-cited article. According to *Simonton*, the citation rate of the authors' most-cited work should be sensitive to the total number of papers written because a blind selection process will produce highly cited articles in proportion to the total number of articles written. This prediction was confirmed, with the number of articles written in our sample accounting for 12.6% of the variance in citation rate of the authors' most cited article in our sample ($F(1, 1168) = 169.12$, $p <0.001$). Including the award accounts for a slightly greater proportion of the variance, for a total of 13.4% ($F(1, 1167) = 14.05$, $p < 0.001$). This analysis provides further support for *Simonton*'s model by showing that the citation rate of each author's most highly cited work depends heavily on the number of articles the author has written and only slightly on the receipt of the award.

Table 1 shows that, in combination, the analyses of citation data from 30 years of *Human Factors* is generally consistent with *Simonton*'s model. Specifically, the citation rate of each article was relatively unaffected by the productivity of the authors. The article influence is generally blind to the history of the author, and the small influence of past productivity may reflect self-citations. As predicted, the total citation rate of an author is highly dependent on the total number of papers written. The average citation rate per article of each author is independent of the productivity of the author, with the ratio of influential articles to total articles being relatively constant. Finally, the citation rate of authors' most-cited work is highly dependent on the total articles written by the author.

In all of the analyses, award receipt had a small but statistically significant effect on the citation rate. Award-winning authors and award-winning papers are cited more often. This conflicts with the idea that determination of article success is a completely blind process. One way to reconcile this apparent conflict is to consider the award as one-level in the variation-selection process. The award selection process represents the consensus of a small elite peer group, and it is not too surprising that the consensus of such an elite group would be somewhat correlated with that of the larger peer group. The award process may not identify the most highly cited papers, but it tends to differentiate good papers that are likely to be cited from the worst of a given year, which are unlikely to be cited. In this way the award acts as an additional selection process that operates at the interpersonal level of *Simonton*'s three-level description of the variation and selection model. Articles that survive the interpersonal level may not succeed at the socio-cultural level. Trends in technology, new scientific findings, and shifts in scientific values make it impossible for reviewers, even elite reviewers who identify the award-winning papers, to anticipate the factors influencing the socio-cultural selection.

Table 1. Summary of analysis to evaluate the effect of author productivity and award receipt on citation rate, with the proportion of variance accounted by each factor shown in parentheses

Predicted citation rate	Productivity	Award
Citation rate of article	Total articles written by all contributing authors (1.3%) $F(1, 1680) = 23.61$, $p < 0.001$	(1.2%) $F(2, 1679) = 9.58$, $p < 0.01$
Citation rate of author	Articles written by author (58.9%) $F(1, 2411) = 3462.0$, $p < 0.001$	(1.2%) $F(1, 2410) = 4.28$, $p < 0.05$)
Average citation rate of author	Articles written by author (0.2%) $F(1, 2411) = 3.06$, $p < 0.05$	(1.1%) $F(1, 2410) = 14.47$, $p < 0.001$
Citation rate of the author's most cited work	Articles written by author (12.6%) $F(1, 1168) = 169.12$, $p < 0.001$	(0.8%) $F(1, 1167) = 14.05$, $p < 0.001$

Conclusions

Can scientific impact be judged prospectively? The results from our research suggest that the answer is "not very well." Prospective judgments of scientific success seem to be, at best, weak and inconsistent, accounting for only 0.8% to 1.2% of the variance in the citation rate of articles. Instead, the number of papers produced by an author seem to provide a much stronger correlate of scientific success, accounting for 58.9% of the variance in the authors' total citation rate and 12.6% of the variance in the citation rate of the authors' most-cited article. These results suggest that the award process may not be able to select articles that will be the most highly cited, but that it can help differentiate them from those that are unlikely to ever be cited. This interpretation is consistent with the results of *Lawani* (1986), who found that prospective peer assessment of quality can significantly discriminate good papers from average papers, but not the best from good. Collectively, these results are generally consistent with *Simonton*'s model of creative productivity.

Several important caveats should temper these conclusions. Most importantly, the sample of citation history included in this analysis is quite limited. It includes a single journal and covers only 30 years of publications. Therefore, it is a small sample of the population of authors' publication and citation patterns. The database does not span the entire career of many authors and does not include the entire output of many authors. The analyses assume that this limited sample is a representative of the publication rate and citation history of scientific publications. *Human Factors* is the flagship journal for the profession and so may not be representative of the full range of journals in which the authors might publish. In addition, the ISI database is incomplete and tracks a subset of journals, which may represent only 50% of the scientific output those citing the *Human Factors* articles (*Brourke* and *Butler*, 1996). However, ISI provides a useful surrogate for the total citation history, with ISI citation patterns generally correlating with non-ISI citation patterns at a level in excess of 0.90 (*Brourke* and *Butler*, 1996). The rigorous peer review may lead to self-selection, in which prolific authors are over-represented. This bias may explain why a paper written by prolific authors tends to be more frequently cited; however, the proportion of self-citation tends to decrease with high-quality papers (*Lawani*, 1986). It is also possible that receiving the award could have influenced the citation rate of the papers rather than anticipating their impact. This seems unlikely because a much higher profile award, such as the Nobel prize, does not effect the total citation rate of an author (*Cole* and *Cole*, 1967; *Garfield* and *Welljams-Dorof*, 1992), but does have a slight effect on the citation rate for individual papers (*Garfield* and *Welljams-Dorof*, 1992). In comparison, receiving Jerome H. Ely Human Factors Article Award is unlikely to change the way citation rate of an article or an author.

Even with these caveats, the research presented here shows that it is possible to study science scientifically (*Hull*, 1998). Doing so can provide important insights into the factors that contribute to scientific success, a relationship that is of practical interest and relevance to scientists in all disciplines.

*

This research was sponsored, in part, by a research grant from the Natural Sciences and Engineering Research Council of Canada. We would like to thank Dean *Simonton*, David *Hull*, and Renee *Chow* for their comments.

References

BOURKE, P., BUTLER, L. (1996), Publication types, citation rates and evaluation, *Scientometrics*, 37 : 473–494.

CAMPBELL, D. T. (1960), Blind variation and selective retention in creative thought as in other knowledge processes, *Psychological Review*, 67 : 380–400.

COLE, S., COLE, J. R. (1967), Scientific output and recognition: a study in the operation of the reward system in science, *American Sociological Review*, 32 : 377–390.

GARFIELD, E., WELLJAMS-DOROF, A. (1992), Of Nobel class: An overview of ISI studies on highly cited authors and Nobel laureates, *Theoretical Medicine*, 13 : 117–135.

HULL, D. L. (1988), *Science as a process: An Evolutionary Account of the Social and Conceptual Development of Science*, Chicago: University of Chicago Press.

HULL, D. L. (1998), Studying the study of science scientifically, *Perspectives on Science,* 6 : 209–231.

LAWANI, S. M. (1986), Some bibliometric correlates of quality in scientific research, *Scientometrics*, 9 : 13–25.

SIMONTON, D. K. (1997), Creative productivity: A predictive and explanatory model of career trajectories and landmarks, *Psychological Review*, 104 : 66–89.

SONNERT, G. (1995), What makes a good scientist? Determinants of peer evaluation among biologists, *Social Studies of Science*, 25 : 35–55.

VINKLER, P. (1998), Comparative investigation of frequency and strength of motives toward referencing: The reference threshold model, *Scientometrics*, 43 : 107–127.

ZUCKERMAN, H. (1977), *Scientific Elite,* Free Press: New York.