

Does Counting Publications Provide Any Useful Information about Academic Performance?

By Michael Skolnik

At the heart of the Publish or Perish Syndrome is the practice of valuing quantity of publication over quality. This practice is so pervasive that it extends even to the selection process for positions in which one would not expect publication counts to be a significant indicator of a person's qualifications to do the job.¹

This was brought home to me several years ago when I served on a selection committee for an academic administrator position. The committee agreed early on that "academic credibility" should be one of several criteria which the successful candidate ought to meet. Members of the committee then looked at the Curriculum Vitae of potential candidates and screened for academic credibility simply and solely on the basis of the numbers of the candidates' publications. In addition, the selection

Michael Skolnik is a professor of higher education at the Ontario Institute for Studies in Education of the University of Toronto, Toronto, Ontario, Canada.

committee received input from faculty-at-large which urged them to reject any candidates whose Curriculum Vitae did not show a substantial list of publications. There was no suggestion at all that quality or impact of candidates' publications should be a criterion, and no discussion in the committee of the content or significance of candidates' publications.

The practice of judging a professor's worth by counting his or her publications is common throughout the academy, whether the professor's primary

Counting Publications

responsibilities are administration, teaching, research or service. Moreover, the fixation on quantity is common across all disciplines. Although this is therefore a general problem in the university, it should be of particular concern for professional fields, like Education, for at least four reasons.

First, it is important for faculty in professional schools to maintain a substantial relationship with the field of practice, and this requires a considerable amount of time. Not only does this demand draw time away from work on preparation of manuscripts for publication, but frequently interaction with the field does not directly foster scholarly publication.

Second, with respect to students the main job of a professional school is preparing them for practice—not necessarily for scholarship, although a minority of students may be headed in that direction. This is in contrast to the academic disciplines where preparation for scholarship takes on a relatively greater role. Consequently, the complementarity between a professor's work of "training" students and his or her own scholarship is much less in professions such as Education than it is typically for academic disciplines. Again, the implication is that, relative to other parts of the academy, in Education, adequately discharging the professor's core professional responsibilities provides less time and opportunity for adding to one's publication counts.

Many observers have pointed out the conflicting nature of the demands on professors of Education, trying on the one hand to meet the needs of the professional field, and on the other to conform to the performance norms of the academy, of which number of publications is of paramount importance (e.g., Judge, 1982; Knowles & Cole, 1998). How readily the former can be sacrificed for the latter in Education and other professional programs is suggested by an incident which I related in a previous publication. In that study I cited a letter I had received from a faculty member in a clinically oriented graduate program who reported that when his program reduced the amount and quality of clinical supervision of its students in order to free up more time for faculty to apply for research grants and prepare publications, its quality rating in the provincial review of graduate programs went up (Skolnik, 1989, p. 631). The professor expressed shock that "a graduate program which purported to be providing clinical training could be rated so highly when it was failing so badly to achieve its stated instructional goals" (p. 632).

Third, the fact that Education is at or near the bottom of the prestige hierarchy of disciplines and fields within the university means that it is particularly vulnerable to charges that it is not meeting the conventional performance norms of the academy. Because of its marginality in this regard, Education professors often feel pressure to overcompensate with respect to the quantity and nature of publication in order to justify their place in academe. Gary Knowles and Ardra Cole have suggested that the situation of Schools of Education in the university community militates against the very reforms in teacher education which are needed in order to improve the quality of education in school systems (Knowles & Cole, 1998).

A fourth reason why the fixation on quantity of publication takes on special significance within Education is that, as will be elaborated subsequently, the practice appears to have disproportionately negative consequences for female academics. Education is one of the fields that has been least resistant to hiring women, but the maintenance of a reward system which penalizes women means that getting in the door will not necessarily be followed by moving up the career escalator.

“Worst-Practice” Evaluation and Some Criticisms

It is something of a paradox that a profession which appears to believe so strongly that surface impressions of things should not be taken at face value, and which prides itself for its appreciation of complexity and subtlety, should embrace so wholeheartedly such a simplistic technique as counting publications and grants for evaluation of its own performance. What is even more ironic is that within our universities we have experts on evaluation whose advice is hungrily sought out by governments and business. Relative even to where the field of evaluation was several decades ago, the typical evaluation of faculty or program quality in most universities could only be described as employing a worst-practice methodology.²

Publication counts are sometimes treated as indicators of quality, sometimes as indicators of productivity. The logical undergirding for either appellation is tenuous at best, and the fact that the same indicator is used as both a measure of productivity and of quality, which are two very different things, should give one pause. It is hard to think of another realm of endeavor where a single *quantitative* measure is used as an indicator of *quality*. Few people would judge the quality of a painter by the number of canvases he turns out, or the quality of a chef by the number of meals she prepares. Is it any more sensible to judge the quality of a professor by the number of papers he produces?

On the surface, it would seem that a simple quantitative index would have more potential to say something about productivity than about quality. In fact, in the literature on faculty work, research productivity is normally defined in terms of the number of publications (see e.g., reviews of literature on faculty research productivity like those of Blackburn & Lawrence, 1995, pp. 30-32; Ontario Council on University of Affairs, 1994, pp. 22-27). This way of defining research productivity, however, ignores the fact that normally the term productivity refers to a relationship between outputs and inputs. Rather, this definition equates productivity with product.

There has been substantial criticism from within academe regarding the emphasis that is placed upon publications counts as a performance indicator. At least until recently, most of this criticism has pertained to the alleged steering effect of the emphasis on number of publications; that is, that it leads faculty to put too much effort into research and publication instead of other activities, especially undergraduate teaching, with what W.S. Massy and R. Zemsky call a “loosening” of their institutional ties and responsibilities (Massy & Zemsky, 1994, p. 2; see also

Counting Publications

Barzun, 1991; Johnson, 1991; Smith, 1991; Sykes, 1988). Rather than repeating or adding to this established body of criticism (a recent survey of which is found in Lucas, 1996, pp. 189-199), my intention in this article is to explore the meaning of publication counts as an indicator of faculty research performance.

A useful perspective from which to begin this exploration is a different, and more recent, line of criticism of publication counts than the one which focuses on the implications for neglect of teaching. This line of criticism is concerned with the equity implications of assessing performance on the basis of number of publications. It alleges that there is gender bias in the practice. Shelley Park (1996) argues that work in the university is organized and allocated by gender and a higher valuation is placed upon work which is performed disproportionately more by men, that is, research. A key observation of Park's is that:

Men, as a group, devote a higher portion of their time to research activities, whereas women, as a group, devote a much higher percentage of their time to teaching and service activities than do men. *The result [my italics] is that men publish more extensively than do women.* (pp. 54-55)

Park alleges that it is mainly gender differences in values and power that account for the difference between males and females in the proportion of time spent in research. She cites literature which asserts that women place greater value than men on nurturing activity (teaching and student advising) and less value on competitiveness (grantsmanship and publication). Also, because women on average have less seniority and lower rank, they may be assigned heavier teaching loads. Park goes on to say that, since research is the most highly rewarded activity in the university and, since "most evidence suggests that research is assessed merely according to quantity, rather than quality of publication" (p. 63), men are more highly rewarded than women.

Robert T. Blackburn and Janet H. Lawrence dispute the assertion that male academics publish more than female academics. They report that while more than 50 studies show that women publish less than men, more recent studies report "near or exact equivalence between the sexes" (1995, p. 49). And in multiple regression analysis on their own data collected in a survey by the National Center for Research to Improve Postsecondary Teaching and Learning (NCRIPTAL), gender failed to predict publication rate, except slightly in one of nine institutional categories. Gender also failed to predict effort given to teaching; however, it is difficult to reconcile Blackburn and Lawrence's claim about what the literature says with Park's literature review. When Blackburn and Lawrence say that "more recent" studies show little or no difference in publication rates between the sexes, they refer to four U.S. studies between 1985 and 1990. Park buttresses her statement that there *are* gender differences in publication rates with seven citations between 1989 and 1994. None of the five sources from 1990 or later which Park gives are in Blackburn and Lawrence's bibliography. Thus, while Blackburn and Lawrence have called into question the view that men publish more than women, they have not overturned it.

The idea that evaluating faculty performance on the basis of numbers of publications is gender biased provides both a compelling argument against the practice, and possibly a constituency to try to get this practice changed. As a step toward this change, I have heard that in some universities, in order to avoid gender bias in hiring, the information that now goes to hiring committees consists of a sample of applicants' actual publications rather than summary publication counts or full Curriculum Vitae which would show the number of publications. This allows hiring committees to concentrate on the quality of applicants' work rather than the quantity. Such an approach is consistent with similar changes being advocated in law firms to ameliorate the gender bias that is thought to be a property of strictly quantitative performance measures in the legal profession. It has been suggested that performance in legal work should be assessed on the basis of the quality of the work rather than by the number of hours billed, which is the analogue in the legal profession of the number of publications in the academic one.

Park's argument implies that measures of the difference between male and female faculty's numbers of publications are not measures of differential productivity but of differences in power and values. If this is true of gender differences, might something similar be true of differences in numbers of publications among faculty generally? It is known, for example, that there are well established differences in publication activity among disciplines. For example, as an indicator, number of publications is biased in favor of the natural sciences relative to the humanities owing to such factors as differences in the amount of support for publication, in the average amount of time needed to produce a journal article, and in professional culture (Skolnik, 1989). Summaries of the literature on factors associated with individual variation in numbers of publications by faculty have been produced by Mary Fox (1985) and by Blackburn and Lawrence (1995). Fox reports that conventional measures of faculty research productivity show no correlation with creativity or with measured ability (e.g., I.Q.), but do show a correlation with stamina. Blackburn and Lawrence report a study which showed Type A Behavior directly related to quantity of publication.

The emphasis on number of publications as an indicator of faculty performance came in for strong criticism several years ago in an inquiry into the tragedy at Concordia University in Montreal in 1992 in which four professors in the Faculty of Engineering and Computer Science were shot to death by another professor who had been thwarted in his many attempts to obtain tenure (Arthurs, Blais, & Thomson, 1993; Monahan, 1995). One of the claims made by the assassin was that more senior faculty had taken credit for papers that he had written in order to pad their Curriculum Vitae. The number of publications on some of the vitae in question were so staggering as to strain credulity. The inquiry committee reported that observers both inside and outside the Faculty of Engineering and Computer Science were skeptical that anyone "could contribute substantively to so many publications" (Arthurs et al., p. 66). The implication was that in this case publication

Counting Publications

numbers reflected not research productivity—and certainly not quality—but the power of someone in a position of authority to get his name listed as senior author when he has not contributed to the research. The inquiry committee expressed its disapproval of the “production-driven research culture” which undergirds the practice of evaluating faculty by their publication counts:

“Production,” as the past experience of the automotive industry demonstrates, can come to be measured primarily in terms of the quantity of units of output, rather than their quality, and to be maximized for its own sake, without regard to the externalities—the social, economic, cultural, and environmental consequences—which it generates. The analogy may be strained, but the implication is apt: too often university honours, research grants, and industrial contracts are awarded on the basis of numbers of publications, rather than on their quality or significance. . . there are strong pressures to be as prolific as possible, and [that] some of these strategies promote undesirable behavior. (Arthurs et al., p. 4)

The Relationship between Quantity and Quality

The report of the inquiry committee at Concordia University created a brief and modest stir in the Canadian university community, but there is no indication that it did the slightest damage to the production-driven research culture which it criticized. One of the reasons for this is that the habit of equating a scholar’s worth with the number of his or her publications is so ingrained. For example, when we listen to a speaker being introduced at an academic conference, we are more likely to be told that she is the author of over 100 journal articles or that he has written 28 books than about the quality of any of those articles or books. Numbers like these are meant to impress an audience, and they generally do.

There may be at least a few members of an audience, however, who are unimpressed, having a good appreciation of how much time and effort are required to produce a really first rate article or book. Still, it is one thing to feel this and another thing to voice it, because the quantitative norm of “more is better” is so strong in academe. As Christopher Lucas notes:

One who suggests that publications counts are equivalent of the emperor’s new clothes—“widely admired though few can see anything there at all”—not only is not apt to be taken seriously, but risks being branded “irresponsible, threatened, or incompetent.” (Lucas, 1996, p. 196)

The persistence and pervasiveness of this norm is especially remarkable when one considers the way that academe has increasingly been fractured by ideological differences among faculty. As a consequence of these differences, some of the previously long-standing norms of academe, such as universalism, have fallen by the wayside. But the idea that academic quality can be measured by number of publications is honored by all, be they conservative, liberal, Marxist, traditionalist, postmodernist, deconstructionist, or whatever persuasion.

Well, almost all. There has been, as I have noted, some questioning of this norm by some feminist scholars. Park (1996), I think, goes right to the heart of the issue when she suggests that one might "hypothesize an inverse relationship between the quantity and quality of research" (p. 63). This is an intriguing and important hypothesis about which so far as I can tell there has been no research.³ There is some *prima facie* basis for expecting that quantity and quality of publication might be inversely related. In the world of fiction writing, it would seem that there are assembly line or formula writers who can churn out book after book in no time, and whose books read like it; and others who take years, finely crafting each sentence. One need only compare the ceaseless output of say, a John O'Hara or Danielle Steele with that of the far less frequent works of a William Styron or Thornton Wilder.

In academic work, the processes of planning and executing research, analyzing and reflecting, writing, revising, and editing are extremely time-consuming. If an individual spent the same amount of time preparing one article as producing ten, it is reasonable to expect that the one would be of higher quality than the ten. Whether the individual will, or should, spend that time on one article or on ten depends upon both his or her own intrinsic valuation of quality versus quantity, and the way that these properties are valued in the reward structure in which he or she operates. The implication of the observations of Park, and of the Concordia Inquiry Committee, cited above, is that the academic reward structure is heavily weighted in the direction of quantity relative to quality. Blackburn and Lawrence (1995) justify the weight given to quantity on the basis of "the importance of research and our need for new knowledge" (p. 115); however, one also hears complaints about the mountain of trivial articles that no one reads coming off academic presses. In any event, deciding how to allocate effort between handling fewer cases better and more cases less well is one of the most fundamental and important choices to be made in any profession. Yet it is an aspect of academic life that rarely, if ever, is the subject of explicit discussion.

Time: The Most Precious Input

How an individual values quantity relative to quality of publications is just one of many factors which determines the number of publications that he or she produces. Earlier I noted Park's emphasis on differences in the proportion of their time that professors devote to research as an important determinant of differences in numbers of publications. Two individuals, however, could devote the same percentage of their time to research, and yet one could have substantially more time for research if that one spent more time in total.

Although studies of faculty work frequently indicate quite a range of hours worked per week, there do not seem to have been many studies which show the relationship between time spent on academic activity and number of publications produced. This is an important gap in the data, because productivity in other sectors

Counting Publications

is defined in terms of output per hour of work. Without data on hours worked to match the data on publications, we cannot calculate true indices of faculty research productivity, and accordingly we can use this term only in a quite misleading way.

One source of such data that I have been able to find is the study of academic work by the Ontario Council on University Affairs referred to earlier. Using data from the Academic Profession in Canada Survey (Lennards, 1987), the study compares the 25 percent of Canadian university faculty who produced the greatest number of refereed journal articles in the preceding three years with the 25 percent who produced the fewest (Ontario Council on University Affairs, p. 26). Faculty in the highest publications output group reported working an average of 51 hours per week which included 24 hours on research and 20 hours on teaching. Faculty in the lowest publications output group reported an average of 43 total hours per week which included 12 hours on research and 24 hours on teaching. Thus, the high producers of refereed articles put in 12 more hours per week on research than the low producers of articles. The high producers found two-thirds of this additional time for research (8 hours) by working a longer week, and one-third (4 hours) by spending less time on teaching.

The Ontario study did not break down the data by gender, or consider possible reasons for variation in time spent on academic activities, such as differences in time demands for child care and other household responsibilities. Other factors which might contribute to differences in time spent on academic work are health, energy level, support systems, and interests. So far as I am aware, in the literature on faculty and faculty performance, there has not yet been any empirical study of how life circumstances influence research output. A genre of work where this relationship has been discussed though, is that of biographies of scholars and scientists. Probably the case that has received the most attention is that of Charles Darwin.

In a review of the most recent biography of Darwin, Stephen Jay Gould (1996) offers a general discussion of the important role that a "fortuitous combination of external circumstances (personal, familial, societal, and historical)" play in enabling scholarly accomplishment (p. 12). Gould notes that it is "a staple of recent Darwin biographies" to observe that Darwin's success was abetted by being a member of an "ever so-blessed group-upper-class white males of substantial wealth and great opportunity," and goes on to say that Darwin:

...used his wealth, his illnesses, his country residence, his protective wife for one overarching purpose: to shield himself from ordinary responsibility and to acquire *precious time for intellectual work* [my italics]. (p. 14)

While one can overestimate the importance of circumstance relative to ability in explaining scholarly output, it is interesting to hear Darwin's own self assessment in his autobiography: "With such moderate abilities as I possess, it is truly surprising that I should have influenced to a considerable extent the belief of scientific men on some important points" (quoted in Gould, p. 12). Darwin's example raises interest-

ing questions about the relationships among ability, circumstance, and intellectual output. (Creationists, no doubt, wish that he had been less blessed with wealth, opportunity, and time). Without intending to suggest that we could all accomplish what Darwin did, if given sufficient time and money, the example shows how favorable circumstances can facilitate scholarly output simply by providing the abundant time necessary for intellectual work.

The Meaning of Publications Numbers

In making decisions on the basis of any numerical indicator, it is important to understand the dynamics of how such indicators are generated. Some of the uses of the indicator, number of faculty publications, which is the object of interest here, are determining tenure, promotion, salary, qualifications for teaching, and overall professional status. On logical grounds and some of the evidence available, questions arise as to whether too much weight has been given to this indicator. For example, there is reason to suspect that prominent among the underlying factors that explain differences in numbers of publications are differences in privilege, power, and preference. Insofar as this is true, there is room for legitimate debate as to how much universities should wish to reward such factors as privilege, power, and preferences.⁴

In view of these suspicions, or at least our lack of understanding about what number of publications produced signifies, it is both perplexing and dismaying that these figures hold such sway in academe, and that their criticism has thus far attracted so little credibility or support. I do not count here the ritualistic acknowledgment one hears in personnel committees that quantity is not a surrogate for quality, before they blithely go on to behave as if it were!

To raise such questions is to begin to deconstruct this performance indicator. Some of the data needed for this deconstruction are not available. What we do know suggests that perhaps as much as measuring ability, brilliance, or creativity, number of publications is an indicator of: personal preferences; personal traits such as ambition, competitiveness, and stamina; cleverness in getting the most publications from a given corpus of research, and in avoiding teaching and administrative work; and power to command the resources and time required for production, and, in some cases, as the Concordia University example indicated, to get one's name on work done by others. It indicates also how well one's personal life—by fortune and design—provides the time, support, and congenial atmosphere for intellectual work. In addition, there is a nagging concern about the likelihood, other things equal, of an inverse relationship between quantity and quality, and about the downsides of a "production-driven research culture." All things considered, there is really little justification for the importance which we attach to sheer counting of publications, and there are alternatives to reliance on it, one of which I suggested earlier—if we are serious about wanting to emphasize quality rather than quantity.

Notes

1. An early version of the ideas underlying this paper was presented at the 26th Anniversary Meeting of the Canadian Society for the Study of Higher Education held at Brock University, St. Catharines, Ontario, May 24-26, 1996. The fact that it has taken so long from first expressing these ideas to getting an exposition of them into print indicates that I have had firsthand experience of the factors which I suggest limits a professor's publication counts—but I hope that readers will find some compensatory evidence of quality in the article! I am indebted to Jamie-Lynn Magnusson of the Ontario Institute for Studies in Education of the University of Toronto for encouragement and for helpful suggestions on the manuscript.
2. Magnusson has made a related point, that assessment practices in higher education adhere to traditional paradigms of social science inquiry at a time when there is no longer widespread adherence to those paradigms in the social sciences (Magnusson, 1996, p. 3).
3. It is common for studies of research performance to note that publication counts do not take account of quality—and then to go on with the analysis of numbers. Typical of such statements is a footnote in the study by the Ontario Council on University Affairs which observes that "... simple numerical counts do not take into consideration the qualitative aspects of the output such as the effort required to produce it, its originality, the contribution it makes to the advancement of knowledge in its field, or its potential social and economic value. A more thorough analysis of research productivity would have to consider the relation between quantity and quality of research outputs." (Ontario Council on University Affairs, 1994). Blackburn and Lawrence (1995) are more sanguine about the prospects of such analysis, noting that "Quality is a social construct, and no doubt will always have a debatable element" (p. 119).
4. It is important to emphasize that what is being criticized here is the practice of assessing faculty research performance solely on the basis of publication counts, not individual faculty members who are forced to comply with this convention. According to Lucas, "the sadness of it all, critics claim, is that so many academics have allowed themselves to become part of a system that forces them to write when, as is painfully obvious, they have nothing of any great importance to say" (1996, p. 84). It is difficult, however, to see how the individual academic could do otherwise. Moreover, clearly not all publication is the result of "scholarship at gunpoint," to use Barzun's term (elaborated in Lucas, pp. 189-199). Many academics have a genuine passion for writing and would try continue to find a way to keep writing regardless of changes in the academic reward structure.

References

- Arthurs, H.W. (Chair), Blais, R.A., & Thompson, J. (1994, April). *Integrity in scholarship: A report to Concordia University*. Report of the Independent Committee of Inquiry into Academic and Scientific Integrity. Montreal, Quebec, Canada: Concordia University.
- Barzun, J. (1991). *Begin here: The forgotten conditions of teaching and learning*. Chicago, IL: University of Chicago Press.
- Blackburn, R.T. & Lawrence, J.H. (1995). *Faculty at work: Motivation, expectation and satisfaction*. Baltimore, MD: The Johns Hopkins University Press.

- Fox, M.F. (1985). Publication, performance, and reward in science and scholarship. In J.C. Smart (Ed.), *Higher education: Handbook of theory and research*. Vol. 1, (pp. 255-282). New York: Agathon.
- Gould, S.J. (1996). Why Darwin? *The New York Review*, 43(6), April 4, 10-14.
- Johnson, I.C. (1991). Myth conceptions of academic work. *The Canadian Journal of Higher Education*, 21(2), 108-116.
- Judge, H. (1982). *American graduate schools of education*. New York: Ford Foundation.
- Knowles, J.G., & Cole, A.L. (1998). Setting and defining the context of reform, In A.L. Cole, R. Elijah, & J.G. Knowles (Eds.), *The heart of the matter: Teacher educators and teacher education reform* (pp. 15-36). San Francisco, CA: Caddo Gap Press.
- Lennards, J. (1987). The academic profession in Canada: Summary report. Unpublished report, Department of Sociology, Glendon College, York University, North York, Ontario, Canada.
- Lucas, C.J. (1996). *Crisis in the academy: Rethinking higher education in America*. New York: St. Martin's Press.
- Massy, W.F., & Zemsky, R. (1994). Faculty discretionary time. *The Journal of Higher Education*, 65(1), 1-22.
- Monahan, E.J. (1995). The Fabrikant Case at Concordia University: Some lessons for the better management of universities and improved academic ethics. *Minerva*, 33, 129-148.
- Magnusson, J. (1996, May). Higher education assessment: Social science in a political arena. A paper presented at the 26th Anniversary Meeting of the Canadian Society for the Study of Higher Education, Brock University, St. Catharines, Ontario, Canada.
- Ontario Council on University Affairs. (1994, August). *The structure of academic work: Background paper*. Toronto, Ontario, Canada: Ontario Council on University Affairs, Task Force on Resource Allocation.
- Park, S.M. (1996). Research, teaching, and service: Why shouldn't women's work count? *The Journal of Higher Education*, 67(1), 46-84.
- Skolnik, M.L. (1989). How academic program review can foster intellectual conformity and stifle diversity of thought and method in the university. *The Journal of Higher Education*, 60(6), 619-643.
- Smith, P. (1991). *Killing the spirit: Higher education in America*. New York: Penguin.
- Sykes, C. (1988). *Profscam: Professors and the demise of higher education*. Washington, DC: Regnery Gateway.