Managing Science.

Social Sciences in Operation.

Par Bert Klandermans

Excerpts from Bert Klandermans's valedictory lecture on the occasion of his retirement as Dean of the Faculty of Social Sciences and Professor of Applied Social Psychology at Vu University, Amsterdam.



Quality Assessment.

Rector manificus, ladies and gentlemen, the answer that is given increasingly within the science system reads, "Let us count." Let us count how many Euros have been acquired, how many publications are realized, and how many citations are generated. The higher the score, the better the researcher. However, it is not that simple. I showed how different the opportunities are for the three science domains to acquire research funds. A report of the Rathenau Institute about the working of various types of grants of Magw suggests that it is not necessarily the best researchers who are granted the Euros (Besselaar and Leydesdof, 2007).

The authors compare successful and unsuccessful grant applications of the years 2003, 2004, and 2005. The question they wanted to answer was: does the money go to the best researchers? In order to answer that question they employed the Social Science Citation Index to compare successful and unsuccessful applicants in terms of the number of publications and citations in the three years before the year when the application was submitted. Although the employment of publication and citation counts from citation indices is questionable—more about that in a while—, it remains of interest to know whether successful applicants have published more and are cited more frequently than unsuccessful applicants, if only because such figures are ever more frequently used as the equivalent of quality. The analyses produce unexpected results.

The 911 applicants who were unsuccessful did indeed publish less and were cited less than the 275 who were successful. So far, so good—apparently quality pays. Upon further inspection the figures are less comforting, though. The problem is that the data are rather skewed. In other words, among the unsuccessful applications are quite a few poor applications. If one takes the poor applications out a surprising picture emerges. On average the best 277 applications of those that were *not* granted are from better researchers than the 275 that were granted. In other words, once the wheat is separated from the chaff, it is no longer the quality of the researcher that controls where the Euros go; on the contrary, and that is not comforting.

One could argue that possibly the granted applications are of better quality than those that were not granted, but this is not true either. External reviewers evaluated all applications. As it turns out, on average the quality of the 275 successful proposals hardly differs from the best 277 unsuccessful proposals. The authors conclude that after elimination of the poor applications one could just as well draw lots among the remaining applications. The average quality of the awarded applications would not be affected, while the average quality of the awarded applicants would even improve. In sum, success and failure in the second stream of funding is not fitted very well as a criterion of quality.

Equally problematic are two other frequently employed criteria of quality—the number of publications and citations. To be sure, we all agree that good researchers publish and are cited, but how do we assess that someone publishes and is cited? Obviously, I can ask somebody's cv and count her publications and assess how frequently she is cited. Yet, this is easier said than done. In the first place, in order to count citations I need a database, and if I want to know whether someone is publishing in comparison a lot or a little I need databases as well. And that's where the misery begins, as the most frequently employed database—Web of Science—is demonstrably biased against the social sciences.

In order to underscore this argument I must take you to the world of the bookkeepers of the science system—the "scientometrist." As not everybody knows what that world looks like, let me take you on a short excursion. The Web of Science is a database where publications in so-called Isi journals are registered. In addition to the title and the text the references are entered. On the basis of these references citation indices can be calculated. These citation indices are an important source for scientometric analyses. Who cites who? How often is an article or an author cited? How often is a journal cited? Those are the questions that are dealt with. On the basis of such data the impact scores of journals can be calculated. The higher the impact factor of a journal, the more desirable a publication in that journal is. In order to be included in the Isi, journals must go through a procedure that takes six years. As a consequence, young, innovative journals are not taken into account initially (Adler and Harzing, 2009). The Isi primarily contains Anglo-Saxon journals and no books and other outlets.

The problem with the Web of Science is that only a limited proportion of the output of the social sciences is taken into account. Scientometrists at the Centre for Science and Technology (Cwts) of the Univeristy of Leiden have calculated that not even one third of the worldwide output of behavioral and social sciences between 1994 and 2003 was registered by the Web of Science (Leeuwen, 2006). In that same period between 80 and 95% of the science and medical science domains were covered. Since then not much has changed. For example, for the year 2006 the Cwts reports for the Uk a coverage of 24% for political science and public administration, 43% for economics, and 35% for the other social

sciences (Cwts, 2007). With a coverage of 75%, psychology seems to have figures comparable to those of science and medical science, but still a quarter of its output is not covered.

Nearer home the two Amsterdam universities reveal a similar picture (Visser, Raan, and Nederhof, 2009). Although the percentages for the social sciences are more favorable than those mentioned above, the Web of Science misses a large part of the output of the social sciences of both universities. While science and medical sciences have a coverage higher than 85%, psychology reaches 75% and 71 %, the social sciences, 47 %, and the humanities, 27 and 22%.

Such differences do not result from differences in quality, but from differences in publication and citation culture. Diana Hicks, for example (2008; see also 1999 and 2004), an international expert in scientometrics, shows that journals where social scientists choose to publish, more often than those in science and medical sciences, are *not* included in citation indices, and, while the latter hardly publish books and national journals, social scientists frequently choose such outlets.

Yearly, Dutch universities must register the scientific output of their personnel. The registration distinguishes between scientific publications and professional publications, books, etc. Scientific publications might be included in the citation indices; professional publications certainly are not. Only 5% of the publications in science and medical sciences are so called professional publications, against 28% in the social sciences and humanities. In quality assessments that employ Web of Science, such output is not taken into account. The effect of this for the social sciences and humanities is many times larger than that for the science and medical sciences (see Butler and Visser, 2006).

In sum, Web of Science fails as an indicator of the quality of research and researchers in the social sciences. In comparisons between disciplines this could lead to the erroneous conclusion that researchers from the one discipline are doing better or worse than those from another. This is already true for comparisons within one domain, let alone for comparisons between domains. To be true, scientometrists warn against such comparisons (see Leeuwen, 2006, p. 138), but such warnings are not heard by university administrators, policy makers, selection committees, and constructors of rankings. This lead Adler and

Harzing (2009), two other scientometrists, to call for a “moratorium on rankings.”

However, for some time, Google Scholar has been offering an alternative to Web of Science. Scientometrists are still a bit hesitant—partly because Google is not very forthcoming with information about its procedures, possibly because of matters of copyrights—, but there seems to be a compromise in the making. In any event, the impression of scientometrists is that Google Scholar is rapidly improving and becoming more reliable. Moreover, software is available now (Harzing, 2009) that enables the same analyses as Web of Science offers. This is good news for the social sciences as Google Scholar covers a much wider range of publications than Web of Science. Let me illustrate that with the example of two scholars. For the sake of anonymity, I name them Beta and Gamma. Beta is a scholar from a domain Web of Science covers for around 80%; Gamma is from a domain covered by Web of Science for around one third of the publications.

The differences are spectacular. For scholar Beta, Web of Science finds 225 publications, 3092 citations, and an h-index of 28. Google Scholar finds 447 publications, 4084 citations and an h-index of 31. For scholar Gamma, Web of Science finds 33 publications, 588 citations, and an h-index of 11; Google Scholar, on the other hand, finds 320 publications, 4294 citations, and an h-index of 28. Better than these figures I cannot illustrate what goes wrong. Viewing Google Scholar the two researchers are more or less the same: to be true, Beta has more publications but citations and h-index are approximately the same. Web of Science, however, suggests an enormous difference between the two; and indeed, in line with the reported figures on coverage of Web of Science the difference between the two databases is much larger for scholar Gamma than for scholar Beta. Erroneously, one would conclude from Web of Science that Scholar Beta is many times better than scholar Gamma.

Databases such as Web of Science or Google Scholar remain problematic in many ways. Hicks (2008) and Leeuwen (2006) recommend falling back on original output sources when possible. Examples of such sources are the annual reports of Dutch universities. The report from Vu University (2008) distinguishes between three types of publications: PhD dissertations, scientific publications,

and professional publications. For the sake of comparison I calculated the number of publications per research fte.

Output at VU University in 2008, standardized per research fte.

	Diss./fte	Sci publ./fte	Sci + prof. Publ./fte	% research fte
Science/Med. Sciences	.17	3.5	3.7	73%
Social Sciences	.22	5.0	7.0	21%
Arts/Humanities	.22	7.5	10.3	6%

The figures are clear—with apologies to my colleagues from science and medical sciences, who perhaps were not aware of the matter. Whatever set of publication we take into consideration the output per research fte in the science and medical sciences domains lags substantially behind that of the social sciences and humanities. The last column reminds us that the distribution of research ftes between the three domains is precisely the reverse. There appears to be a strong, negative correlation between the number of ftes invested in a domain and the output per fte.

Why is There No Protest?

Rector magnificus, ladies and gentlemen, behold the state of the social sciences: financing that fails, quality assessment that fails, funding which is demonstrably lagging behind despite excellent achievements. In the field where I am at home best—the study of protest behavior—this would be characterized as illegitimate inequality—known as the engine of protest, but apparently not for social scientists in the Netherlands, as they do not seem to protest. This raises the question “Why not’” To be sure, there are protests, but staged—oddly enough in view of what I said before—by scholars from science. This by the way is in line

with the literature: in case of decline, protest is more often staged by those who are in comparison better off. But—back to the question of why social scientists do not protest. Why do they not gather at the doorsteps of the Ministry of Education, at the offices of Now, the Vsnu, or our University administrators? This relates to characteristics of the science system.

Scholarly work still is in many ways an individualistic endeavor; that holds for the social sciences and humanities even more than for science and medical sciences. In fact, social scientists are a collection of single-worker enterprises that compete with each other for scarce resources. In such settings, failing to acquire funds to do research is easily viewed as individual failure. Under such circumstances, structural underfunding, like in the case of the social sciences, manifests itself to the onlooker as an underachieving or at least unsuccessful discipline. Indeed, the science system prefers to see itself as a meritocracy that awards quality. In a system of which people pretend that it gives everybody equal chances, inequality if noted at all is seldom defined as illegitimate. Awareness of shared grievances—a necessary condition for protest to occur—is unlikely to develop in such situations.

For the same reason, another condition for protest to occur—the formation of collective identity—is unlikely to be met. To be sure, we are all social scientists, but when it comes to that point, we are, if not each other's competitor, at least coming from different disciplines, such as psychology, sociology, communication science, economics, and so on. This explains also why another condition of protest is so difficult to meet, namely the establishment of effective organizations that interpret grievances, represent interests, and organize protest if needed. The only institution in this country that brings representatives of all disciplines to one table is the so called Disciplinary Consultation of Social Sciences (Dsw), also called Deans' Consultation. This group of gentlemen and one lady of which I had the honor to be the chair during the last years is not the most suited to operate as a protest organization and to mobilize for collective action, although we did warn Nwo, the Academy of Sciences (Knaw) and the Vsnu repeatedly—without much effect, though.

Illustration : Willemvdk, « Flag of the Netherlands », 24.5.2010, Flickr, (Creative Commons license).

Résumé

Quality Assessment.Rector manificus, ladies and gentlemen, the answer that is given increasingly within the science system reads, “Let us count.” Let us count how many Euros have been acquired, how many publications are realized, and how many citations are generated. The higher the score, the better the researcher. However, it is not that simple. I ...

Bibliographie

Nancy J. Adler and Anne-Wil Harzing. “When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings” in *Academy of Management Learning & Education*, vol. 8, 2009, pp. 72–95.

Peter van den Besselaar and Loet Leydesdorff, *Past Performance as Predictors of Successful Grant Applications: A Case Study*, Den Haag, Rathenau Instituut, 2007.

Linda Butler and Martijn S. Visser, “Extending Citation Analysis to Non-Source Items” in *Scientometrics*, vol. 66, 2006, pp. 327–43.

Anne-Wil Harzing, *The Publish or Perish Book,* Melbourne, Tarma Software Research Pty Ltd, 2010.

Diana M. Hicks, “The Difficulty of Achieving Full Coverage of International Social Science Literature and the Bibliometric Consequences” in *Scientometrics*, vol. 44, 1999, pp. 193–215.

—, “The Four Literatures of Social Science” in Henk Moed (dir.), *Handbook of Quantitative Science and Technology Research,* Dordrecht, Kluwer Academic p, 2004, pp. 1–18.

—, “The Four Literatures of Social Sciences,” lecture, “Kennismakers: Dag van de Onderzoeker,” Vlaanderen, Fwo, 2008.

Thed van Leeuwen, “The Application of Bibliometric Analyses in the Evaluation of Social Science Research: Who Benefits from It, and Why It Is Still Feasible” in *Scientometrics*, vol. 66, 2006, pp. 133–54.

Faculty of Social Sciences, *Research Assessment in Social Sciences: Self-Evaluation Report, 2001–2006*, Amsterdam, Vu University, 2007.

Centre for Science and Technology Studies (Cwts), *Scoping Study on the Use of Bibliometric Analysis to Measure the Quality of Research in UK Higher Education Institutions. Report to Hefce*, Leiden, Leiden University, 2007.

M.S. Visser, A.F.J. van Raan, and A.J. Nederhof, *Bibliometric Benchmarking Analysis of the Amsterdam Universities, 2002–2007*. Leiden, Leiden University, 2009.