Fatal Attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods

To be published in Scientometrics Vol. 62, No. 1, 2005, p. 133-143

Anthony F.J. van Raan
Centre for Science and Technology Studies
Leiden University
Wassenaarseweg 52, P.O.Box 9555
2300 RB Leiden, The Netherlands

Abstract

Ranking of research institutions by bibliometric methods is an improper tool for research performance evaluation, even at the level of large institutions. The problem, however, is not the ranking as such. The indicators used for ranking are often not advanced enough, and this situation is part of the broader problem of the application of insufficiently developed bibliometric indicators used by persons who do not have clear competence and experience in the field of quantitative studies of science. After a brief overview of the basic elements of bibliometric analysis, I discuss the major technical and methodological problems in the application of publication and citation data in the context of evaluation. Then I contend that the core of the problem lies not necessarily at the side of the data producer. Quite often persons responsible for research performance evaluation, for instance scientists themselves in their role as head of institutions and departments, science administrators at the government level and other policy makers show an attitude that encourages 'quick and dirty' bibliometric analyses whereas better quality is available. Finally, the necessary conditions for a successful application of advanced bibliometric indicators as support tool for peer review are discussed.

1. Introduction: crucial problems in bibliometric analysis

Recently, ranking of universities has widely attracted the attention of policy makers, the scientific world and the public media. Particularly the ranking published by the *Jiao Tong University* in Shanghai (SJTU 2003) stirred the fire. Here, as in other cases, 'bibliometric' elements such numbers of publications and citations to these publications, play a crucial role. These rankings suggest a similar simplicity for the evaluation of scientific performance as in the case of a football league. The immediate observation that the well-known US top-universities take the lead, reinforces these suggestions. But things are not so simple.

Weingart (2003, 2004) pointed at the very influential role of the monopolist citation data producer ISI (Institute for Scientific Information, now Thomson Scientific) as its commercialization of these data (Adam 2002) rapidly increased the non-expert use of bibliometric indicators such as rankings. I think, however, that the 'danger' is not the data producer, but organizations greedy to buy ready-to-go indicators without any competence to understand what is measured. It is my experience that authoritative persons in

organizations still cultivate the unreasonable 'please press the button and I have the numbers that I want to have' mentality.

Before discussing the severe problems related to the use of rankings in the evaluation of 'scientific quality', I first present a brief overview of the basic elements of bibliometric analysis. Bibliometric assessment of research performance is based on one central assumption: scientists who have to say something important, do publish their findings vigorously in the open, international journal ('serial') literature. This assumption introduces unavoidably a 'bibliometrically limited view of a complex reality'. For instance, journal articles are not in all fields the main carrier of scientific knowledge; they are not 'equivalent' elements in the scientific process, they differ widely in importance; and they are challenged as the 'gold standard' by new types of publication behaviour, particularly electronic publishing. However, the daily practice of scientific research shows that inspired scientists in most cases, and particularly in the natural sciences and medical research fields, 'go' for publication in the better and -if possible- the best journals. A similar situation is developing in the social and behavioural sciences (Glänzel 1996; Hicks 1999), engineering and, to a lesser extent, in the humanities.

A first and good indication whether bibliometric analysis is applicable to a specific field is provided by the publication characteristics of the field, in particular the role of *international*, refereed journals. If international journals are the dominating or at least a major means of communication in a field, then in most cases bibliometric analysis is applicable. Therefore it is important to study first the 'publication practices' of a research group, department, or institute, in order to establish whether bibliometric analysis can be applied. A practical measure here is the share of CI-covered[1] publications in the total research output. For 'not-CI covered publications' a restricted type of analysis is possible, in as far as these papers are cited by CI-covered publications.

Work of at least some importance provokes reactions of colleague-scientists. Often, these colleague-scientists play their role as a member of the invisible college by referring in their own work to earlier work of other scientists. Thus, citation analysis is based on reference practices of scientists. The motives for giving (or not giving) a reference to a particular article may vary considerably (Brooks 1986; MacRoberts and MacRoberts 1988; Vinkler 1998). So undoubtedly the process of citation is a complex one, and it certainly not provides an 'ideal' monitor on scientific performance (MacRoberts and MacRoberts 1996). This is particularly the case at a statistically low aggregation level, e.g., the individual researcher. There is, however, sufficient evidence that these reference motives are not so different or 'randomly given' to such an extent that the phenomenon of citation would lose its role as a reliable measure of impact (van Raan 1998). Therefore, application of citation analysis to the entire work, the 'oeuvre' of *a group of researchers*

[1] The Science Citation Index, the Social Science Citation Index, the Arts & Humanities Citation Index and the 'specialty' citation indexes (CompuMath, Biochemistry & Biophysics, Biotechnology, Chemistry, Material Science, Neurosciences) are produced and published by the Institute for Scientific Information (ISI/Thomson Scientific) in Philadelphia. In this paper we use the term 'CI' (Citation Index) for the above set of indexes.

as a whole over a longer period of time, does yield in many situations a strong indicator of scientific performance.

With citation analysis as core element, 'bibliometric' indicators measure the impact of published research in an international perspective. Why should we apply bibliometric analysis of research performance? Peer review undoubtedly is and has to remain the principal procedure of quality judgment. But peer review and related expert-based judgments may have serious shortcomings and disadvantages (Moxham and Anderson 1992; Horrobin 1990). Subjectivity, i.e., dependence of the outcomes on the choice of individual committee members, is one of the major problems. This dependence may result in conflicts of interests, unawareness of quality, or a negative bias against younger people or newcomers to the field. Basically, the methodological problem of determining the quality of a subject is still far from solved, as illustrated by the results of re-review of previously granted research proposals, see, for instance, Nederhof (1988) and the studies mentioned by Weingart (2004).

I do not plead for a replacement of peer review by bibliometric analysis. Subjective aspects are not merely negative. In any judgment there must be room for the intuitive insights of experts. I claim, however, that for a substantial improvement of decision-making on matters of scientific activities, advanced bibliometric indicators have to be used in parallel to a peer-based evaluation procedure. Properly designed and constructed (Moed *et al* 1995; Van Raan 1996; 2000), they can be applied as a powerful support tool to peer review. Also for interdisciplinary research fields this is certainly possible (Van Raan and Van Leeuwen 2002). Bibliometric indicators have great potential, more than many people think, as these indicators provide necessary and even unexpected insight into scientific developments. This is, however, only the case under the following two fundamental conditions: (1) the *technical system* and (2) the *methodology* on which these indicators are based, must be sufficiently advanced and sophisticated. This means that the very many pitfalls and sources of error in citation analysis have to be known in detail and that all relevant corrections have to be made.

Inappropriate design, improper calculation and sloppy application of bibliometric indicators negatively influence the appreciation of bibliometric methods in general by the scientific community, and thus severely hamper the application of carefully constructed indicators. The question of the appropriateness of the methodology is certainly in both a scientific and in a practical sense most crucial. But in contrast to what is thought generally, still the quality of the technical system is the first source of problems to be tackled. Methodological problems are certainly also present, but it is not very useful to discuss them if basic technical problems -in order to arrive at a reliable data system on which, as a next step, a bibliometric indicator methodology can be based- are not solved.

2. Technical problems

The most central technical process on which citation analysis is based entirely, is the *matching* of *citing* publications with *cited* publications. With this process the following is meant. When in a publication (which is the 'citing publication') a reference is given to

another publication (which is the 'cited publication'), the reference has to be identified as -an earlier- 'source publication' in the citation indexes. This is in fact the most important *technical* procedure that has to be executed by the producer of the citation indexes. In most cases this 'identification-by-matching' procedure is done well, but we remind that the original purpose (Wouters 1999) of the citation indexes is its role as an *information retrieval* database, and not as a database suited for *evaluation*.

The reason for this statement is that still a considerable amount of errors occurs in the citing-cited matching process leading to a 'loss' of citations to a specific publication. In average, the number of non-matching references -although they are citation index covered source papers- is about 7% of the citations matched (Moed 2002). Frequently occurring non-matching problems relate to publications written by 'consortia' (large groups of authors), to variations and errors in author names particularly -but certainly not only- authors from non-English speaking countries, errors in journal volume numbers, errors in initial page numbers, discrepancies due to journals with dual volume-numbering systems or combined volumes, or to journals applying different article numbering systems. Thus, these non-matching citations are highly unevenly distributed in specific situations, which may cause an increase of the percentage of lost citations up to 30% (Moed 2002). So if the citation indexes are used for *evaluation* purposes, all these possible errors have to be corrected as much as possible.

Then there is a *second* extremely important technical problem. This problem relates to the *attribution* of publications -and with that, of the citations to these publications- to specific organizations such as institutes, university departments, and even on a high aggregation level to the main organization, for instance universities. Very often it is thought the citation indexes can simply be scanned in order to find 'all' publications of University X. This assumption is based on the argument that all these publications mention somewhere in the address data of the publication clearly 'University X' as the main affiliation of the authors. But this assumption is naïve and totally wrong.

Next to variations in the name of the same university, we must also realize that departments and institutes (in many variations) are to a non-negligible extent mentioned without proper indication of the university. Furthermore, groups or institutes of a national research organization (such as the French CNRS) are quite often mentioned instead of the university where the research actually takes place. The same problems occur for 'graduate schools'. Even the name of a city as such will not always be sufficient, as parts of universities may be located in suburbs.

Another difficulty with large research organizations is to distinguish the many institutes within the main organization. Very problematic in this respect is capturing the *medical research* of universities, as often only the medical school, and/or the name of the hospital without mentioning the university is indicated. One explicitly needs the names of all those hospitals in a specific city –and also in the suburbs of a city- that are in fact university hospitals. So again problems related to names of cities and their suburbs play a crucial role. Again, large efforts in cleaning and re-building the original citation indexes are necessary to solve this problem.

For instance, in the commercial database on highly cited scientists we find five variants for Leiden University: Leiden University, Universiteit Leiden, Leiden Observatory, Leiden University Medical Center, Leids Universitair Medisch Centrum. In the Shanghai study the number of highly cited scientists determines 20% of the ranking position of a university. If a university is not properly 'unified' (e.g., all variants in the above example has to be re-defined as 'Leiden University' only), it may loose a considerable part of the score (e.g., highly cited scientists in the observatory or in the university hospital, typically disciplines in a university with a high chance for highly cited scientists!). Furthermore, these highly cited scientists are mainly determined by 'life time citation counts', which enhances the 'old boys' effect and does not specifically focus on the impact at the research 'front' of today by younger researchers.

On top of that, further problems arise when two or more universities are within one city. In some cases these problems are so large (e.g., Vrije Universiteit Brussel and the Université Libre de Bruxelles, both are indexed as 'Free University (of) Brussels') that it is virtually impossible to distinguish both universities on citation-index based address data only. Similar problems occur for Leuven and Louvain-la-Neuve. Next, there are major differences in research systems between countries affecting the 'definition' of a university. For instance, the University of London is not a university anymore in the usual sense. It is an 'umbrella organization' covering several different virtually autonomous universities. But in Paris and other French cities not such an umbrella structure exists, there we deal with completely autonomous universities that were part of originally one 'mother-university'. As a consequence, it is very cumbersome to distinguish between departments, medical schools and hospitals of these different universities within a city.

All these affiliation-related problems require very careful and mostly cumbersome cleaning and, in addition and most importantly as discussed above, *unification* of addresses and attribution of publications to the right main organization, including 'unification of cities' with their suburbs. As in the case of the earlier mentioned technical problems related to discrepancies in cited references, large and continuous efforts are necessary to 'rebuild' the citation indexes into a data system that is accurately enough to use it for the calculation of bibliometric indicators and to apply it for evaluation purposes.

All these discrepancies not only affect the 'measurement' of highly cited scientists as one of the determining factors for the position of a university on the ranking list, but all three citation-relation indicators in the Shanghai study, summing up to 60% of the score used for the ranking. In a first assessment, I estimate that this may introduce an uncertainty in ranking of possibly 5 to 10 positions in the European list, and about 25 to 50 positions in the world list.

3. Methodological problems

After having done all the corrections, cleaning and unifications related to two main technical problems as mentioned above, the methodological problems concerning the

design and the calculation of these indicators, are still pertinent and has to be solved too. It is the next job you have to do. Methodology directly relates to the aims of a study. The objective of the Shanghai study is to obtain a worldwide ranking of universities in terms of their 'scientific strength'. The crucial point is then, with which indicator, or with which combination of a set of indicators (and with what weights per individual indicator) the ranking has to be constructed.

In this paper I will not discuss the methodology of the Shanghai group in detail. I already stressed, that *first of all* the technical problems have to be solved before discussion of the applied methodology makes any sense. Nevertheless, I here give a number of general remarks that are at the basis of the methodology. It is clear that the Shanghai study heavily depends on the ISI citation index data. First, the citation index *coverage* is problematic particularly for engineering, social and behavioral sciences, and certainly for the humanities. This problem is directly related to the well-known fact that application of citation index data depends on the *role of journal articles* in the different fields. So the strength of a university in engineering, in the social and the behavioral sciences or in the humanities may contribute little -or even hardly- to the position of that university in the ranking.

Second, there is the problem of a *US bias* in citation data. As always in rankings based entirely on citation-index based bibliometric indicators, US universities and research institutions dominate by far the top-positions of the ranking. No doubt that the US top-universities are institutions of undisputed world-class. But there is also the effect -to a certain extent- of US 'dominance' in the overall 'publication and -particularly- citation traffic'. It is not easy to find out to what extent, and to correct accurately for it. More methodological work is necessary to come to grips with this problem, and before this problem is solved, it is not useful and even 'dangerous' to make worldwide rankings as they suggest that, for instance, Europe is scientifically far beyond the US. This may be true in a number of situations, but certainly not to the extent as is suggested purely on the basis of bibliometric rankings.

A further important point is related to the type of article: 'normal articles', 'letters', 'notes' and 'reviews'. There are large differences between, for instance, *reviews* and *normal articles* both on the 'citing side' (the number of references, for reviews this number is usually large to very large) and on the 'cited side' (the number of citations received, for reviews in many cases again large). Therefore, the average impact (mean number of received citations) of research groups (as parts of universities) may be dominated substantially by reviews. However, reviews are in most cases not original scientific work, the authors present state-of-the-art overviews of developments in their field. So it is absolutely necessary to take article type into account in all normalization procedures (i.e., comparison of reviews with reviews, etc.) in the calculation of impact-indicators used for the ranking procedure. As far as I can see, the Shanghai group does not take the different article types into account. In fact, they indicate that in the calculation of their fourth indicator, "Articles in Science Citation Index-expanded and Social Science Citation Index", only publications 'of article type' are taken into account.

I assume that they here refer to the 'normal article type', but in that case the important article type 'letters' does not play any role in this ranking indicator.

A next point concerns language. Recent work (Grupp *et al* 2001; van Leeuwen *et al* 2001) shows that the utmost care must be taken in interpreting bibliometric data in a comparative evaluation of national research systems (May 1997). The measured value of impact indicators of research activities at the level of an institution and even of a country strongly depends upon whether one includes or excludes publications in CI-covered journals written in *languages other than English*. This is due to the simple fact that the CI covers non-English language (particularly French and German-language) journals of which the papers have a considerably lower impact than those in the English-language journals. So if researchers use these non-English language journals for publications, these publications are indeed formally entered into the citation indexes. But generally the impact of publications of these French and German-language journals is (very) low. So in the calculation of impact indicator these publications count on the output side, but they contribute very little, if any at all, on the impact side.

Therefore, such non-English language publications considerably 'dilute' the measured impact of a university or a department. We have clear empirical evidences (van Leeuwen *et al* 2001) that the use of German-language journals covered by the citation indexes may lead to about 25% lower measured impact. Simply by removing the publications in these German-language journals and only using the English-language journals (which is fair in an international comparison and certainly in a comparison with, for instance, the United States and the UK), the measured impact will 'improve' with this 25% for a whole medical faculty of a university! No doubt that there will be the same effect for French-language journals. These findings clearly illustrate again that indicators need to be interpreted against the background of their inherent limitations, such as, in this case, effects of publication language, even at the 'macro-level' of entire countries, but certainly at the level of institutions.

The most serious problem of these rankings is that they are considered as 'quasi-evaluations' of the universities considered. This is absolutely unacceptable. Next to all the technical and methodological problems discussed above, I must also stress that if bibliometric analysis is used for evaluation of universities or university departments, it is a *condition sine qua non* that publication lists are verified by the institutions under evaluation. But even on the basis of *exactly the same* data and *exactly the same* technical and methodological starting points, *different types* of impact-indicators can be constructed, for instance one focusing entirely on normalized impact, and another in which also scale (size of the institution) is taken in to account. Rankings based on these different indicators are not the same, as proven very clearly by a recent CWTS study for the European Commission (Van Raan and Van Leeuwen 2001).

4. Further Observations and Conclusions

From the above considerations we conclude that the Shanghai ranking should not be used for evaluation purposes, even not for benchmarking. Several aspects of the Shanghai methodology have to be analyzed in more detail. For instance, the number of Nobel laureates as a partial indicator and factor in the ranking procedure. Also here 'affiliation' is a serious problem. A scientist may have an (emeritus) position at University A at the time of the award (which seems to be the criterion in the Shanghai study), but the prize-winning work was done at University B. The 1999 physics Nobel Laureate Veltman is a striking example (A= University of Michigan, Ann Arbor; B = University of Utrecht).

Rankings such as the Shanghai one are part of a larger problem in the science evaluation circus. Quite often I am confronted with the situation that responsible science administrators in national governments and in institutions request the application of bibliometric indicators that are not advanced enough. They are aware of this insufficient quality level, but they want to have it 'fast', in 'main lines', and not 'too expensive'. Many of the problems mentioned by Weingart (2004) such as the 'easy' use of journal impact factors for research performance assessments of individuals scientists or groups simply should not exist: there are already better indicators for quite a long time. But still heads of institutions demand their librarians to do 'quickies', i.e., rapid and, particularly, cheep 'evaluations' with help of standard journal impact factors. The fault of these leading scientists and administrators is asking too much and offering too little. The responsible persons do not want to pay a reasonable amount for a study of better quality. They do pay a considerable amount for the data, but want to have a competent evaluation study based on these data for a small fee: 'the data are already there, so please press the button'. Thus, it is not so much the commercialization of the monopolist data producer ISI that makes the problems. These heads of institutions, government administrators and policy makers are the first to blame that the intermediary research groups that hitherto cleaned the crude ISI data, prepared the data for the construction of reliable bibliometric indicators and developed the competence and skill to interpret the indicators (Weingart 2004) are being squeezed out of the market.

In the line of the above discussion, I think that the cases presented by Weingart (2004) concerning unintended steering effects of bibliometric measures are due to the simple fact that the applied bibliometric analyses are not sophisticated enough. Again: the real problem is not the use of bibliometric indicators as such, but the application of *less-developed* bibliometric measures. I already mentioned the simple use of the standard journal impact factors, although already for a long time bibliometric researchers warned against this use with clear arguments (Moed and Van Leeuwen 1995) -even in top-journals such as Nature (Moed and Van Leeuwen 1996).

My final remarks concerns the principles with which Weingart concludes his paper. These principles are completely in line with the experiences of our Leiden group and I fully support them. As discussed in detail in this paper, the 'raw' publication and citation have to be cleaned, corrected and enriched in many aspects. This is the first task of professional, experienced bibliometricians who publish their findings regularly in the open, academic literature. The second task of the professional bibliometric community is the construction of user-oriented, as advanced as possible indicators of research

performance on the basis of the corrected data and building on already developed methodology and best practice. Also here I have to point my finger in another direction than at ISI: within the bibliometric community there are too much efforts to 're-invent the wheel' without proper reference to, and building on earlier work. That was at least my impression of the recent Jülich conference (see ref. to Weingart 2003).

The third task follows immediately from the second: as peer review has to remain the primary process of research evaluation, bibliometric indicators must act as a support tool in order to make peer review more objective and transparent (Rinia *et al* 1998). Professional bibliometricians can act here as advisers, not as 'number crunchers', in order to add value to the peer review process and to avoid misleading use which cause damage to universities, institutes and individual scientists. A recent example of 'best practice' is the application of bibliometric indicators combined with peer review in the evaluation of all university research groups in chemistry and chemical engineering in the Netherlands (VSNU 2002).

This paper is written in addition to the author's contribution to the conference "Bilbiometric Analysis in Science and Research" led in Jülich (Germany), on 5-7 November, 2003.

References

Adam, D. (2002). The counting house. *Nature* 415, 726-729.

Brooks, T.A. (1986). Evidence of complex citer motivations. *Journal of the American Society for Information Science* 37, 34-36.

Glänzel, W. (1996). A bibliometric approach to social sciences, national research performances in 6 selected social science areas, 1990-1992. *Scientometrics* 35, 291-307.

Grupp, H., U. Schmoch, and S. Hinze (2001). International alignment and scientific regard as macro-indicators for international comparisons of publications. *Scientometrics* 51, 359-380

Hicks, D. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics* 44, 193-215.

Horrobin, D.F. (1990). The philosophical basis of peer review and the suppression of innovation. *Journal of the American Medical Association (JAMA)* 263, 1438-1441.

van Leeuwen, Th.N., H.F. Moed, R.J.W. Tijssen, M.S. Visser, and A.F.J. van Raan (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics* 51, 335-346.

MacRoberts, M.H. and B.R. MacRoberts (1988). Author motivation for not giving citing influences- a methodological note. *Journal of the American Society for Information Science* 39, 432-433.

MacRoberts, M.H. and B.R. MacRoberts (1996). Problems of citation analysis. *Scientometrics* 36, 435-444.

May, R.M. (1997). The scientific wealth of nations. *Science* 275, 793-796.

Moed, H.F. (2002). The impact factors debate: the ISI's uses and limits. *Nature* 415, 731-732.

Moed, H.F. and Th.N. van Leeuwen (1995). Improving the accuracy of the Institute for Scientific Information's Journal Impact Factors. *Journal of the American Society for Information Science* (JASIS) 46, 461-467.

Moed, H.F. and Th.N. van Leeuwen (1996). Impact Factors Can Mislead. *Nature* 381, 186.

Moed, H.F., De Bruin, R.E., and Th.N. Van Leeuwen (1995). New Bibliometric Tools for the Assessment of National Research Performance: Database Description, Overview of Indicators and First Applications. *Scientometrics* 33, 381-422.

Moxham, H. and J. Anderson (1992). Peer review. A view from the inside. *Science and Technology Policy* 7-15.

Nederhof, A.J. (1988). The validity and reliability of evaluation of scholarly performance. In: Van Raan, A.F.J. (ed). (1988). *Handbook of Quantitative Studies of Science and Technology*. Amsterdam: Elsevier/North-Holland, pp.193-228 (ISBN 0-444-70537-6).

van Raan, A.F.J. (1996). Advanced Bibliometric Methods as Quantitative Core of Peer Review Based Evaluation and Foresight Exercises. *Scientometrics* 36, 397-420.

van Raan, A.F.J. (1998). In matters of quantitative studies of science the fault of theorists is offering too little and asking too much. *Scientometrics* 43, 129-139.

van Raan, A.F.J. (2000). The Pandora's Box of Citation Analysis: Measuring Scientific Excellence, the Last Evil? In: B. Cronin and H. Barsky Atkins (eds.). *The Web of Knowledge. A Festschrift in Honor of Eugene Garfield*. Ch. 15, p. 301-319. Medford (New Jersey): ASIS Monograph Series, 2000. ISBN 1-57387-099-4

van Raan, A.F.J. and Th.N. van Leeuwen (2001). Identifying the fields for mapping RTD excellence in the life sciences -a first approach. Report to the European Commission, Brussels, under contract COPO-CT-2001-00001.

van Raan, A.F.J. and Th.N. van Leeuwen (2002). Assessment of the scientific basis of interdisciplinary, applied research. Application of bibliometric methods in nutrition and food research. *Research Policy* 31, 611-632

Rinia, E.J. Th.N. van Leeuwen, H.G. van Vuren, and A.F.J. van Raan (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria. Evaluation of condensed matter physics in the Netherlands. *Research Policy* 27, 95-107.

SJTU, Shanghai Jiao Tong University (2003). Ranking Methodology. See website http://ed.sjtu.edu.cn/rank/methodology.htm

Vinkler, P. (1998). Comparative investigation of frequency and strength of motives toward referencing, the reference threshold model- comments on theories of citation? *Scientometrics* 43, 107-127.

VSNU (2002). *Quality Assessment of Research. Chemistry and Chemical Engineering Research in the Netherlands*. Utrecht: VSNU. The bibliometric part of this report is available via the CWTS website: www.cwts.nl.

Weingart, P. (2003). Evaluation of research performance: the danger of numbers. In: *Bibliometric analysis in science and research. Applications, Benefits and Limitations*. Second Conference of the Central Library, Forschungszentrum Jülich, p. 7-19 (ISBN 3-89336-334-3).

Weingart, P. (2004). Impact of bibliometrics upon the science system: inadvertent consequences? In: H.F. Moed, W. Glänzel and U. Schmoch (eds). Handbook on Quantitative Science and Technology Research. Dordrecht (The Netherlands): Kluwer Academic Publishers, 2004.

Wouters, P. F. (1999), The Citation Culture, PhD thesis, University of Amsterdam.

****