Journal Evaluation: Technical and Practical Issues

RONALD ROUSSEAU

ABSTRACT

THIS ESSAY PROVIDES AN OVERVIEW of journal evaluation indicators. It highlights the strengths and weaknesses of different indicators, together with their range of applicability. The definition of a "quality journal," different notions of impact factors, the meaning of ranking journals, and possible biases in citation databases are also discussed. Attention is given to using the journal impact in evaluation studies.

The quality of a journal is a multifaceted notion. Journals can be evaluated for different purposes, and hence the results of such evaluation exercises can be quite different depending on the indicator(s) used. The impact factor, in one of its versions, is probably the most used indicator when it comes to gauging the visibility of a journal on the research front. Generalized impact factors, over periods longer than the traditional two years, are better indicators for the long-term value of a journal. As with all evaluation studies, care must be exercised when considering journal impact factors as a quality indicator. It seems best to use a whole battery of indicators (including several impact factors) and to change this group of indicators depending on the purpose of the evaluation study. Nowadays it goes without saying that special attention is paid to e-journals and specific indicators for this type of journal.

INTRODUCTION

Few model-based approaches to journal evaluation can be found in the literature. A descriptive, but not explanatory model is the one used by the Leiden-based Centre for Science and Technology Studies (Tijssen & van Raan, 1990). Perhaps this overview will inspire fellow scientists to construct

an overall model explaining observed journal citation scores, and hence lead to a better understanding of their role in institutional and national evaluations. Theoretical issues dealt with in this article are restricted to giving precise formulations of indicators, in particular of the journal impact factor. No input-output model or explanation of dependent variables, such as journal citation counts, as a function of one or more independent variables (e.g., number of journals in the field or number of active scientists) is provided.

The study of the use and relative impact of scientific journals is an important application of citation analysis. Yet citations are only one aspect of a journal evaluation exercise. Indeed, journal evaluation can be performed with many purposes in mind. Impact factors measure only the (international) use of journals on the research front. Hence, they are of little direct use to a (special) librarian, because, as Line (1977) notes: Users of journals read, but many actually publish little or nothing at all. In this context, it is important to investigate the relation between in-house use and citation use. This has been done, for example, by Ming-yueh Tsay (1998, 1999) in a medical library. Numerous studies have shown that older volumes of scientific journals are less frequently used (read as well as cited) than more recent volumes. This phenomenon is generally described by the term "obsolescence" (Brookes, 1970; Line, 1993). A mathematical model describing the relation between the growth of the literature and obsolescence can be found in Egghe & Rousseau (2000).

It should also be pointed out that scientists read not only as a step in their scientific investigations, but also to keep informed of the latest findings in their field, or simply out of general interest. Further, the importance of scientific journals is not restricted to use (local or international). Geographic penetration in the sense of geographical distribution patterns of subscribers, authors, and citers, as well as the correlations between them, is still another indicator. Irene Wormell (1998) performed such an investigation of geographical distributions for the following journals: *College & Research Libraries, Computer Journal, Information Processing & Management, Journal of Documentation, Journal of the American Society for Information Science, Libri*, and *Scientometrics*. Studies like this one tell us whether international journals are really international in scope and impact. Among the journals considered by Wormell, *Libri* turned out to be the most international one, while *College & Research Libraries* is a very nationally oriented (i.e., U.S.) journal.

Many people are interested in journal evaluations: Librarians, scientists, science evaluators, publishers, etc. Librarians are interested in journal evaluations and local circulation data for selection and deselection purposes, and in the relation between impact and price (Van Hooydonk et al., 1994; Van Hooydonk, 1995; Abbott, 1999). Scientists want to find the most appropriate journal in which to publish their results. Funding agencies and governments want their grantees to publish in the most

prestigious journals (Pao & Goffman, 1990; Lewison & Dawson, 1997). Editors and publishers may relate high citation scores to a successful editorial practice and policy. Commercial publishers are interested in subscription data and sales. Information brokers are interested in finding those sources that have the most potential of satisfying their clients' needs. University research councils use journal impact and prestige scores as elements in local research evaluation studies in view of enlarging the visibility of the university's research.

Because economic indicators such as subscription data are essential for commercial publishers, an investigation, such as Peritz's (1995), of the relation between these and citation data is of great value. Let us just mention that, in most instances, Peritz found correlations between 0.25 and 0.5.

Besides serving as an archive for research findings, scholarly printed journals also provide professional, institutional, and disciplinary visibility, as well as recognition and prestige, to scientific authors. This, in turn, provides prestige to the journals themselves. Complex systems of "pecking orders" are based on the ranking of journals and a journal's position in them. The quality of the editorial board counts for much, of course, but the typography, quality of the paper used, quality of the illustrations, etc. all play their role. A truly excellent journal regularly garners papers from well-established authors and secures a larger number of institutional and individual subscriptions, thus making for a solid financial (economic) base.

The next sections cover the following topics: The definition of a quality journal, different definitions of impact factors, a general model for the citation distribution, electronic journals, the meaning of ranking journals, possible biases in citation databases, and how to use the journal impact in evaluation studies.

QUALITY JOURNALS

How has a quality journal been defined, what are the elements in such a definition, and how have they been used in practice? As early as 1970, Zwemer published the following list of characteristics of a "good journal":

1. High standards for acceptance of manuscripts (results must be based on new scientific information, reliable methods, adequate controls, and statistical treatment of data);
2. Having a broadly representative editorial board with appropriate representation of subdisciplines;
3. The editor uses a critical refereeing system;
4. Promptness of publication;
5. Being covered by major abstracting and indexing services;
6. Scientists using the articles published in the journal have a high confidence level in its contents;
7. Having a high frequency of citation by other journals.

These seven criteria are also among those used by the Philadelphia-based Institute of Scientific Information (ISI) to determine inclusion (or exclusion) of journals in their database (Garfield, 1990; Testa, 1998). The ISI management further mentions the following requirements:

8. Including abstracts or summaries in English;
9. Including authors' addresses;
10. Providing complete bibliographic information.

For new journals the reputation of the publisher and of the main editor is a good indicator of the possible importance or quality of the journal. If, for example, Elsevier, the American Chemical Society, or the IEEE launches a new journal, this will probably be a more important one than the newly established "Research Reviews of the Department of . . . of the . . . University."

Panels of (subject) experts have acted as judges to determine the value of journals and to draw formal ranked lists (Van Fleet, McWilliams, & Siegel, 2000). This approach is especially useful in the social sciences and humanities where the Science Citation Index (SCI) and Journal Citation Reports (JCR) cannot be used, and where local journals are often important. This is due to the local character of the investigations, as is the case in (national) law, or the literature or linguistics of small languages (Luwel et al., 1999).

Depending on the purpose and the type of journal, different journal indicators may be determined. Popular science journals, such as *Scientific American, Dr. Dobb's Journal,* and the *New Scientist,* are only marginally interested in impact factors. Besides practicing good (science) journalism, the number of subscriptions and corresponding revenues is what really counts for such journals.

The number of interlibrary lending (ILL) requests is still another local "use" indicator. Indeed, if a library does not subscribe to a journal, the librarian cannot directly determine its local use. In that case the number of local ILL requests for that journal can act as an indicator of its importance for the community served by the library.

Finally, a quality journal is indexed by many databases. Hence, the number of databases indexing this journal can be used as an indicator of its importance. However, as sheer numbers are not very important here, it is probably more relevant to investigate whether a scientific journal is covered by the most important database(s) in the field.

CITATION IMPACT

Investigations related to journal citations and impact received a considerable impetus since the annual publication (since 1976) of Journal Citation Reports (JCR) by the Institute of Scientific Information (then under the direction of Eugene Garfield). Generally speaking, the JCR is a statistical data set providing information on how often journals are cited,

how many items have been published, and how often, on the average, each item is cited. It also reports those source journals responsible for the references of each journal, the number of references each journal has published, and the distribution of those references in time (Egghe & Rousseau, 1990).

As early as 1960, Raisig suggested the use of a journal impact factor. He called it the "index of research potential realized" (p. 1418). Nowadays different "impact factors" are used. Defining exactly what is meant by the notion of an impact factor is not easy. Indeed, different impact factors exist, and a precise notation and some mathematical terminology is necessary in order to show their differences. First, it is stressed that citations, and hence impact, is always calculated with respect to a certain pool of journals. In practice these are usually all journals covered by ISI. For the moment, it is assumed that the journal of which the impact is calculated belongs to that pool. Impact factors are always quotients of the form: Number of citations received, divided by number of items published. They differ by the periods considered.

How to Calculate Impact Factors

The standard ISI (or Garfield) impact factor (Garfield & Sher, 1963) of a journal J in the year 2002 is obtained as follows:

- Collect the number of citations received in the year 2002 by journal J. Not all citations are used, however; only those related to articles published in the two previous years: 2001 and 2000. These numbers are denoted as $CIT_J(2002, 2001)$ and $CIT_J(2002, 2000)$.
- Find the number of articles published in journal J in the years 2001 and 2000. These numbers are denoted as $PUB_J(2001)$ and $PUB_J(2000)$.
- Form the quotient of the sum of $CIT_J(2002, 2001)$ and $CIT_J(2002, 2000)$, by the sum of $PUB_J(2001)$ and $PUB_J(2000)$. This is the ISI or Garfield impact factor of the journal J for the year 2002.

Written as a mathematical formula this is:

$$1. \ \frac{CIT(2002, 2001) + CIT(2002, 2000)}{PUB(2001) + PUB(2000)}$$

If now the symbol $CIT_J(Y, X)$ denotes the number of citations received (by a fixed journal J, from all members of the pool) in the year Y, by articles published in the year X, and the symbol $PUB_J(Z)$ stands for the number of articles published by this same journal in the year Z, then one can similarly define a Garfield impact factor for any year (not just the year 2002). The algorithm described above needs only little modifications. It becomes:

- Collect the number of citations received in the year Y by journal J. Use only citations pertaining to articles published in the two previous years: $Y - 2$ and $Y - 1$. These numbers are denoted as $CIT_J(Y, Y - 1)$ and $CIT_J(Y, Y - 2)$.

- Find the number of articles published in journal J in the years $Y - 1$ and $Y - 2$. These numbers are denoted as $PUB_J(Y - 1)$ and $PUB_J(Y - 2)$.
- Form the quotient of the sum of $CIT_J(Y, Y - 1)$ and $CIT_J(Y, Y - 2)$, by the sum of $PUB_J(Y - 1)$ and $PUB_J(Y - 2)$. This is the ISI or Garfield impact factor of the journal J for the year Y.

As a mathematical formula this is:

$$2. \ \frac{CIT(Y, Y-1) + CIT(Y, Y-2)}{PUB(Y-1) + PUB(Y-2)}$$

ISI defines the so-called immediacy index in the year Y as the number of citations obtained during the year of publication, divided by the number of publications. This is:

$$3. \ \frac{CIT(Y)}{PUB(Y)}$$

Actually, formulae 2 and 3 are biased in favor of "immediate" (i.e., short-term) citations. It is clear that 2 can easily be generalized to include more than two years. This leads to a generalized (n year) synchronous impact factor, denoted as IF(Y, n) (Rousseau, 1988), where now citations and publications over n years are taken into account (the exact formula, equation 4, is presented in the Appendix).

If it is clear from the context which year is meant, or if the exact year does not matter, one simply writes IF(n). Hence, ISI's or Garfield's impact factor is IF(2). ISI's five-year impact factors are denoted as IF(5). All synchronous impact factors, however, suffer from the same problem: They mix different publication years. This practice, however, should not be followed in research evaluation studies. Indeed, the more aspects (in this case the publication year) are kept constant the better. Consequently, a diachronous impact factor, denoted as IMP, keeping the year of publication fixed (see Appendix for a precise formulation) is the preferred index for evaluation studies by the Centre for Science and Technology studies (Moed, Frankfort, & van Raan, 1985; de Bruin et al., 1993; van Raan, 2000). In my LUC evaluation studies (Rousseau, 1995, 1998a, 1998b), I used IMP with a four-year citation window. For a description of the difference between synchronous and diachronous impact factors and their use in research evaluation, the reader is referred to Ingwersen et al. (2001).

Obviously, for a librarian, the long-term impact (perhaps ten years) is of considerable more importance than the short-term (two-year) impact of a journal. Using different generalized impact factors, or different windows, allows one to compare the long-term versus the short-term journal impact. Garfield (1998) performed such an investigation. He found that some journals, such as *Cell, The New England Journal of Medicine, Proceedings of the National Academy of Sciences, Nature,* and *Science,* always had a high impact,

whatever the period (two, seven, or fifteen years). Other journals moved up or down significantly. Letters journals in particular suffered considerable downward changes in ranking.

Until now the journal for which the impact was calculated has been assumed to be a member of the pool. This leads one to question how to measure the impact of a journal that is not in the pool (e.g., a non-ISI journal). This will be explained for the ISI impact factor, equation 2; then, comments will be given on the diachronous impact factor (equation 5, see Appendix).

In order to calculate an analogue of the ISI impact factor for a non-ISI journal, one simply adds this journal to the pool of ISI source journals. One determines how often this particular journal is cited by ISI journals (during the period under investigation) and adds the number of times the journal cites itself. Then one simply divides by the number of articles published by the non-ISI journal (Spaventi et al., 1979; Sen, Karanjai, & Munshi, 1989; Stegmann, 1997, 1999). Although this is a simple procedure, there are two caveats. First, ISI always includes journal self-citations, but for these "constructed impact factors" this is not done. For journal evaluation purposes, it may indeed be more appropriate to remove journal self-citations for ISI-covered journals as well (Stegmann, 1997). Second, if this new impact factor is used to compare the non-ISI journal with ISI-journals, the ISI-journals' impact factor must also be recomputed, because the pool of journals has changed.

In the case of the diachronous impact factor, the method (and the caveats) are the same. There is, however, one important benefit here. It becomes possible now to calculate the (diachronous) impact of a book containing conference proceedings or contributions written by different authors. This has been done for *Informetrics 87/88* (Rousseau, 1997a). Besides the obvious benefits for research evaluation, this fact is also interesting from a theoretical point of view. Indeed, one can even determine a volume, issue, or section diachronous impact factor, leading to a possibly finer grained statistical study of the visibility and impact of a journal.

Although the impact factor is a size-independent measure (or at least a size-limited one), since it is defined as a ratio, with the number of publications in the denominator, it suffers from other limitations. According to Pinski & Narin (1976), the most important drawback of the "traditional" impact factor(s) is the fact that citations are not weighted. All citations are counted as equally important, regardless of the citing journal. To remedy this limitation (and related ones) Pinski & Narin (1976) proposed a new weighted measure for journals. Unfortunately, this measure is seldom used for journal evaluations. Most evaluators stick to some form of the traditional impact factor. Yet, the Pinski-Narin measure inspired the makers of the Internet search engine Google to take the strength of hyperlinks into account for their search output-ranking algorithm (Brin & Page, 1998; Kleinberg, 1999).

Meaning of Self-cited and Self-citing Rates

The self-citing rate of a journal relates a journal's self-citations to the total number of references it gives. The self-cited rate relates a journal's self-citations to the number of times it is cited by all journals in the database. A high self-cited rate is an indicator of a journal's low visibility. A high self-citing rate is an indicator of the isolation of the field covered by the journal (Egghe & Rousseau, 1990). The self-cited (SCD) and self-citing (SCG) rates of a journal over a fixed period are calculated as follows:

If A denotes the number of references in journal J to journal J;

B denotes the total number of citations received by journal J; and

C denotes the total number of references in journal J, then

$$4.\ SCD = \frac{A}{B} \qquad SCG = \frac{A}{C}$$

An interesting (and little known) indicator is the so-called popularity factor of journal J (Yanovsky, 1981): This is the ratio of the number of journals citing (in a particular period) journal J, over the number of journals cited by J. It tells us something about whether the journal exports knowledge (ratio larger than one) or rather imports knowledge (ratio smaller than one). For those willing to evaluate journals by a whole battery of indicators, this is certainly one that deserves inclusion.

THE BASIC CITATION MODEL AND ITS CONSEQUENCES

Recall that a citation curve is a curve showing the number of citations received by a source (usually a journal, but it can also be an author, institute, or country) over a certain period. It is generally agreed that citation curves can be modeled as unimodal graphs, having a mode at the year two (i.e., two years later than the publication of the journal) or later. This is in accordance with Price's theory on the immediacy effect (Price, 1970): The number of references to literature of a specific age rises until the cited literature is two or three years older than the citing literature, and then falls off gradually. At the mode the curve levels off, so that the number of citations obtained three years after the publication of the article—CIT(Y, Y − 3)—is larger than the average of the number of citations received one and two years after the publication of the article. Wouters (1999, p. 176) offers a nice real-world example of this phenomenon. Of course, it is well known that there are exceptions to this model. This often happens in very dynamic fields, such as biomedicine. Another well-known exception is the self-citation curve of a journal (Rousseau, 1999).

For this basic model it is further assumed that the number of publications does not decrease in time. This means that PUB(Y − 3) = PUB(Y − 2) = PUB(Y − 1), because, for example, PUB(Y − 2) denotes the number of items published two years before year Y, while PUB(Y − 3) denotes the number of articles published three years before year Y. The assumption that sources

(e.g., journals) do not decrease their production over time is a very natural one. Indeed journals, and certainly successful ones, generally increase the number of articles they publish (Rousseau & Van Hooydonk, 1996).

Rousseau et al. (2001) shows that IF(3), the synchronous impact factor calculated over a three-year period, is, in the basic model, always larger than IF(2), the "standard" impact factor.

From the literature, it is known that the basic model can be described by certain statistical distributions, such as the lognormal or the Weibull distribution. Using realistic parameters for these distributions, one can show that it follows from the shape of these curves that the three-year synchronous impact factor is always larger than the two-year one—IF(3) > IF(2). This has been done in Rousseau (1993). The basic model, and, in particular, its consequences concerning the synchronous impact factor, were confirmed by Rousseau (1988) for mathematics journals, and for a random sample of journals in ISI's database by Dierick & Rousseau (1988). Other studies related to the basic model were published by Rao (1973) and Nakamoto (1988). A recent investigation by Rousseau et al. (2001) using the Chinese Science Citation Database did not confirm the basic model.

ELECTRONIC JOURNALS

The calculation of impact factors for printed journals or for online journals (i.e., e-journals) is exactly the same. Of course, besides impact, both kinds of journals have specific indicators. Subscription data are not meaningful for free e-journals, while counting links from Web sites or other e-journals to a particular e-journal is a typical aspect of e-journal evaluation.

One of the many criticisms of citation counts as an indicator for use (or visibility) is the fact that they only measure a special kind of use. They offer no information on reading, browsing, or other forms of use. For e-journals, though, it is possible to collect use data on a finer scale. One can not only count how many persons visit a journal's site, but one can collect viewers' data per article. This corresponds roughly to measuring the number of times a printed article is examined in a library (maybe several times by the same person). If this electronic article does not only exist in HTML format, but also in a complete downloadable PDF or PostScript format (as is often the case), then one can also count the number of download operations. This distinguishes "browsers" or occasional visitors from persons who are genuinely interested in the article. Finally, one can count the number of links made to this article. This corresponds to an electronic citation (sometimes called, with a pun, a sitation (Rousseau, 1997b)). Note that some e-journals, such as *Conservation Ecology* (Holling, 1999), already collect some of these data. Hence, this yields three visibility indicators for articles in e-journals: The number of visits to the article's page, the number of downloads, and the number of links (sitations). This leads to an appreciable increase of usage information with respect to citation counts that,

however, would continue to play their role as another kind of visibility or use indicator. Admittedly, there are, at the moment, some problems with this approach. Some people download (via the "save as" option in popular browsers) or directly print the HTML version. However, downloading a complete article in this way requires that one saves different objects (text, graphs, pictures) separately, which is not handy. Further, printing the HTML file usually leads to a poorer quality copy than that obtained by printing the PDF or PS version. Hence, for these reasons, download counts would miss only a small percentage of all interested scientists.

The announcement of the publication of a paper on the Web, by a newsgroup or another alerting service, may lead to an enormous increase in the hit rate for this paper. This effect has been termed the Slashdot effect (Adler, 1999). Similarly, a catchy phrase in the title of a Web-based article or a site is probably even more effective in generating traffic to that paper or site. Hence, a "catchy phrase effect" is predicted for Web sites and articles.

Yet, notwithstanding a "slashdot or catchy phrase effect" for separate articles or sites, e-journals themselves have, until now, not been able to generate high impact factors (Harter, 1998).

RANKING JOURNALS: THE MEANING OF A RANK

Impact factors, such as those published in the JCR, lead to a global ranking of journals. It is, however, clear at a glance that the top of this general list is dominated by certain types and fields: Multidisciplinary and review journals and journals in biomedicine are obviously at an advantage with respect to journals in engineering or the library and information sciences. Indeed, such general rankings exhibit an inherent bias against journals from small fields. Even within fields, rankings are often heavily influenced by the uneven impact of subfields on the broader field.

Consequently, ISI has devised a field classification scheme and journal rankings can also be viewed per subfield (subject category listings). The idea to devise a "disciplinary impact factor" dates already from 1978 (Hirst) and is regularly taken up again. Sometimes field rankings use the whole database as citation pool, sometimes only journals in the field are considered to be sources of citations. Both approaches have positive and negative aspects. In the second case, there is a clear discrimination of journals that try to act as a bridge between several subdomains, or between the applied and the basic side of a discipline. In the first case, it is possible that a journal receives more citations from outside the field than from inside, and perhaps that too is not always desirable. Again, trying to use both approaches (if possible) is the appropriate way to proceed.

As mentioned before, there are significant differences in the citation potentials of different scientific fields, that is, in the maximum number of times any given article—and, hence, also any journal—will be cited in its lifetime. It is clear that the number of research workers in the field is an impor-

tant factor here. Yet, Garfield (1979) claims that the major determinant of these citation potentials is the average number of references per article.

What Is the Meaning of a Rank?

Lists of ranked journals (ranked according to, for example, impact factor) are said to help users to identify sources with significant contributions (Todorov & Glänzel, 1988). Yet rankings of journals according to the number of citations received or the impact factor are only meaningful as long as fluctuations reflect a real rise or drop in the importance or influence of the journal, and is not only the result of noise or of a purely random process. In order to account for the random effect on citation measures, Schubert & Glänzel (1983) devised a method for estimating the standard error of mean citation rates per publication and applied this method to find confidence intervals for the impact factor. Nieuwenhuysen & Rousseau (1988) devised a "quick and easy" method to find a lower bound on the size of fluctuations of the impact factor. As there are many more journals with a low impact factor than journals with a high one, rankings for the low impact ones are less stable than for the high impact ones. Table 1 (a hypothetical example) illustrates the influence of fluctuations on a journal's impact ranking.

It suggests that, for high impact journals, noise and fluctuations have only a small influence on the impact, and do not lead to any change in ranking. For low impact journals, on the other hand, noise and random effects may lead to a considerable change in ranking (i.e., it is possible that journal E actually ranks third and not fifth). This example agrees with McGrath's observation (1993) that rankings of anything are often unreliable, particularly if those ranks are based on data with large variability. Consequently, adjacent values of data, when ranked, are often not significantly different.

Different types of articles lead to different citation potentials. This ef-

Table 1. Influence of Fluctuations on a Journal's Impact Ranking.

Journals	# Citations	# Publications	Rank	Error on Citation Counts	Highest Impact	Lowest Impact
A	100	20	1	±4	104/20 = 5.20	96/20 = 4.80
B	50	20	2	±4	54/20 = 2.70	46/20 = 2.30
C	22	20	3	±3	25/20 = 1.25	19/20 = 0.95
D	20	20	4	±3	23/20 = 1.15	17/20 = 0.85
E	18	20	5	±3	21/20 = 1.05	15/20 = 0.75

fect leaks down to the journal level if journals "specialize" in certain types of articles. Besides the possible effect of letters to the editor, due to a wrong methodology (Moed & van Leeuwen, 1996), Peritz (1983) showed that, at least in sociology, methodological papers are more cited than theoretical or empirical ones. Rousseau & Van Hooydonk (1996) clearly showed that the impact factors of review journals are much higher than those of "normal" journals, while the impact factors of translations are much lower. In general, they found that the more articles a (normal) journal publishes the higher its impact factor.

BIAS?

ISI's database and hence all measures derived from it are often accused of being biased. They are said to be biased in favor of American journals, in favor of English language publications, or in favor of certain fields (mainly basic science), etc. This is probably true to some extent, but until a scientifically valid definition of bias is given (Garfield, 1997), it is impossible to say to what extent this bias is inherent in the scientific community as a whole, or in the way American scientists (the largest community) behave, or is due to commercial decisions of ISI. It is true, though, as stated by Spinak (1995, p. 353), that research processes are not "objective and neutral" but are part of a social milieu, and, as a result, can vary from one society to another. Using ISI's products as the only standard would reduce evaluation studies to the North American standard, which is not necessarily that of other communities (again this problem is more severe in the social sciences and the humanities than in the sciences). Local citation indices, such as the Chinese Science Citation Database (Jin & Wang, 1999) and the Chinese Scientific and Technical Papers and Citations (CSTPC) database, may provide a solution to this problem.

As stated above, an impact factor is always calculated with respect to a pool of journals. So it is a legitimate question to ask what would happen if ISI covered more or other journals? What if ISI or another organization had started with an initial set of French, Chinese, or Spanish language journals? Would this have led to a different pool of international journals (Rousseau & Spinak, 1996)? Nothing can be stated with certainty, of course, but the question is worth investigating. To some extent, this challenge has been taken up by Leo Egghe, who, in two articles, (Egghe, 1998, 1999) studied limiting properties of a stochastic process describing the evolution of core collections, including the quality of the original set of source journals (Egghe, 1999).

It is clear that the fact of whether a journal is included in the ISI database or not may have a profound impact on its visibility, and hence on its standard impact factor. The inclusion of journal self-citations plays an important role here (Gómez et al., 1997), as some journals derive a large part of their impact factor from self-citations.

Although complaints about bias in citation-based measures continue to be heard, using prestige rankings by peers does not offer a solution, as these are certainly biased. Christenson & Sigelman (1985) found that scholarly journals in sociology and political sciences tend to establish reputations that endure in spite of what they merit. Once a journal has been placed on a discipline's prestige ladder, it tends to retain its place because its reputation is accepted at face value. Such journals are not re-evaluated in the light of changing circumstances. Comparing prestige scores with impact scores showed that good and bad reputations tend to be exaggerations of what impact scores suggest are merited. This clearly is a form of the Matthew effect (Merton, 1968): Already famous persons (or journals) receive more credit than they actually deserve, while recognition of less prestigious scientists (or journals) is withheld. The Matthew effect derives its name from the following quote from the Gospel according to St. Matthew:

> For unto everyone that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath. (25:29)

Bonitz, Bruckner, & Scharnhorst (1997, 1999) studied the Matthew effect for countries. They found that:

> Few countries with high expectations [i.e. expected number of citations, based on journal impact factors] receive more citations than expected while many countries with low expectations receive fewer citations than expected. (1999, p. 362)

This redistribution effect originates in a relatively small number of journals, headed by *Nature, Physical Review B, Science,* and *Physical Review Letters.* Countries such as China, the former Soviet Union, and Nigeria are among the greatest losers.

USE OF THE JOURNAL IMPACT IN EVALUATION STUDIES

Quality journals in science generally contain coherent sets of articles, both in contents and in professional standards. This coherence stems from the fact that most journals are nowadays specialized in relatively narrow subdisciplines and their gatekeepers, that is, editors and referees, share views on questions like relevance, validity, and quality with the invisible college to which they belong (Schubert & Braun, 1993). This is the main reason why journals can play a legitimate role in evaluation studies (de Bruin et al., 1993; Spruyt, de Bruin, & Moed, 1996).

When gauging the impact of research groups, comparisons are made with their peers. The two most interesting indicators are the ratio of the average of the group's citations (per article) with the average of the journals in which they have published, and the ratio of the average of the group's citations with the average of the field (or fields) in which they are active (de Bruin et al., 1993; Rousseau, 1998a, 1998b). When calculating

the impact of a field, two approaches are possible: Either one just takes the average of the impact factors of all journals in the field (this is called the average impact of this set of journals), or one calculates a global average (Egghe & Rousseau, 1996). The latter is the better approach. The difference between these two approaches is shown—mathematically—as follows. If C_j denotes the number of citations (over a certain period) of journal j, and if P_j denotes the number of publications in journal j, then I_j denotes the impact of journal j (citations per publication). The average impact factor is then defined as:

$$5.\ AIF = \frac{1}{n}\sum_{j-1}^{n}\frac{C_j}{P_j} = \frac{1}{n}\sum_{j-1}^{n}I$$

The global impact, on the other hand, is calculated as:

$$6.\ GIF = \frac{\sum_{j-1}^{n}C_j}{\sum_{j-1}^{n}P_j} = \frac{\mu_C}{\mu_P}$$

where μ_C and μ_P denote the mean number of citations and the mean number of publications. Hence, the first one is an average of quotients, while the second one is a quotient of averages. An example (Table 2), will illustrate the numerical difference between these two approaches.

The global impact of the meta-journal consisting of the four journals A, B, C, and D is 1.96, while the average of these journals' impact is only 1.35. This difference is due to the fact that (here) the journals with the lowest impact publish the lower number of articles.

Problems with Using Impact as a Quality Measure

It is clear that there are problems with using impact as a quality measure: These two notions clearly cannot be substituted for each other. Some of these problems were discussed in the previous sections. They are briefly recalled here and some other ones are highlighted.

Some fields are very useful for science as a whole, but by their particu-

Table 2. Artificial Meta-journal and the Calculation of the Average Impact.

Journal	# of Articles	# of Citations	Impact
A	20	8	0.40
B	20	10	0.50
C	100	250	2.50
D	200	400	2.00
			1.35 Average Impact
Meta-journal	340	668	1.96 Global Impact

lar nature, cannot be cited much. If the impact factor (or similar measures) would become the main determinant to judge journal quality this could eliminate whole subfields, and undermine the health of many others. A case in point is basic taxonomy (Valdecasas, Castroviejo, & Marcus, 2000). Doing high-quality work in taxonomy is expensive and time consuming. Good taxonomy articles have continued to be cited for more than a century after their publication. Moreover, taxonomy lies at the basis of all biodiversity studies. Yet, during the short period used to calculate impact factors they will attract few or no citations. This, however, tells us nothing about the quality of taxonomy journals. Similar cases can be made for other fields of science: An enormous gap lies between popular research areas (with many thousands of authors, papers, and citations) and less popular ones (Schoonbaert & Roelants, 1998). Neglecting these less popular fields because of citation counts would lead to an impoverishment of science.

There is also the following technical problem: How should multidisciplinary journals be evaluated? Specifically, how are articles published in these journals treated? It would be best if individual articles were assigned to the proper category and its citations compared with the citation results of that category. This means that one needs a (preferably automatic) method to assign articles to categories, and to delineate categories. This assignment problem of individual articles has been studied, for example, in de Bruin & Moed (1993) and Glänzel, Schubert, & Czerwon (1999).

Finally, with an eye to future developments, I would like to make the following remark concerning the future of journal impact factors. Although journals will always consist of articles, and journal impact will always be a kind of "average" measure of its articles' impact, it is clear that for electronic journals the emphasis will be much more on the individual article, and less on the journal. This trend will probably erode the value given to journal impact factors.

A review of the use of bibliometric techniques for research and institutional evaluation can be found in Russell & Rousseau (in print).

CONCLUSION

The quality of a journal is a multifaceted notion. Journals may be evaluated for different purposes, and the results of such evaluation exercises can be quite different, depending on the used indicator(s). The impact factor, in one of its versions, is probably the most used indicator when it comes to gauging the visibility of a journal on the research front. Generalized impact factors, over longer periods than the traditional two-year period, are better indicators for the long-term value of a journal. The diachronous approach is strongly favored.

As with all evaluation studies, care must be exercised when considering journal impact factors as an indicator of quality. It seems best to use a

whole battery of indicators (including several impact factors) and to change this group of indicators depending on the purpose of the evaluation study. Moreover, in the case of journal evaluation, it should be pointed out that calculating impact factors for one particular year is not very instructive. Trend analyses of impact factors over several years have much more value for the evaluation of journals (in the same field, of the same type!).

Journal impact and scores of research groups with respect to the impact of the journals used as publication outlets are just two elements in evaluation studies. Ranking projects, institutes, or research groups on the basis of impact factors only makes sense for scientists working in the same field. Indeed, evaluation, whether of journals, scientists, or institutes, is only a means to an end, not a goal in itself.

We hope that more people with a library and information sciences degree will be involved in journal evaluation studies, not only with the aim of finding an optimal set of journals for local use, but also when it comes to institutional evaluation exercises. Consequences are too heavy to leave the job to computer scientists or alumni of a management school. A librarian's daily task involves handling, buying, canceling, copying, binding, and discussing journals. They have the expertise to be part of an evaluation team, at least when it comes to having a well-founded opinion on the quality of journals. The author hopes this article helps them in better understanding the mathematical technicalities.

Finally, the subject of journal evaluation and the use of journals in research evaluation exercises have attracted scores of empirical articles. Yet, relatively few model-based approaches can be found in the literature. Perhaps the time is ripe to make a "grant model" that can be used to explain observed journal citation scores, and hence their role in institutional evaluations.

ACKNOWLEDGMENTS

APPENDIX: MATHEMATICAL FORMULATIONS OF THE SYNCHRONOUS AND DIACHRONOUS IMPACT FACTORS

The n-year synchronous impact factor is defined as:

$$7. \; IF(Y, n) = \frac{\sum_{i=1}^{n} CIT(Y, Y-1)}{\sum_{j=1}^{n} PUB(Y-j)}$$

Taking n = 2 yields the standard, or Garfield, impact factor. The n-year diachronous impact factor for the year Y is defined as:

$$8. \ IMP(n) = \frac{\sum_{i=k}^{n} CIT(Y + i, \ Y)}{PUB(Y)}$$

with k = 0 or 1. Sometimes one includes the publication year (k = 0), sometimes one does not (k = 1).

GLOSSARY OF TERMS

average impact factor
The average impact factor of a group of journals (or meta-journal), as opposed to the global impact factor. See text for a mathematical formulation.

basic citation model
A citation model. The number of citations to a fixed journal issue is assumed to reach a top quickly (after two or three years) and then start a slow decline. During the first period, when the journal becomes "better" with time, the Burgundy effect (getting better with age) prevails. The basic citation model also assumes that the number of publications in a journal does not decline over time.

catchy phrase effect
Term to denote that articles with a special or trendy phrase in the title attract more attention than other ones, especially on the Internet.

Chinese Science Citation Database (CSCD)
Database compiled by the Documentation and Information Center of the Chinese Academy of Sciences (DICCAS). It has a similar purpose as the Science Citation Index, but uses only Chinese sources. Source of the *Chinese Scientometric Indicators*.

Chinese Scientific and Technical Papers and Citations (CSTPC)
Database compiled by the Institute of Scientific and Technical Information of China (ISTIC). It has a similar purpose to the Science Citation Index, but uses only Chinese sources. It is the source of the *Chinese S&T Journal Citation Reports*.

citation pool
The set of documents whose references are used in counting citations.

diachronous impact factors (IMP)
A group of impact factors using citations received in different years, but referring to one specific publication year. See Appendix for an exact mathematical formulation.

Garfield impact factor
Popular name for the synchronous impact factor referring to a two-year citation window.

global impact factor
The impact factor calculated for a group of journals considered as one whole (meta-journal). See text for a mathematical formulation.

HTML format
HyperText Markup Language (HTML) is a high-level programming language used to write hypertext documents with corresponding text and hyperlinks. It allows nonprogrammers to design Web pages by specifying their structure and content, but leaves the detailed presentation and extraction of information to the client's Web browser.

immediacy index
An indicator used by ISI to determine the impact of a journal's publications during the year of publication.

indicator
Statistic used to determine the state of an activity. This is usually an economic activity, but the term is used in bibliometric studies to study science or information-related entities, such as journals, research output of institutes, Web-activity, and so on.

Institute of Scientific Information (ISI)
ISI, the company founded by Eugene Garfield, is now a Thomson Scientific Company, and part of The Thomson Corporation. The company, through its Science Citation Index, the Web of Science, and related products, indexes the most influential scientific and technical journals from 1945 onwards. ISI captures all bibliographic information including the citations or references that are part of a peer-reviewed article or item. ISI's databases may be used for information retrieval and for science evaluation purposes.

ISI impact factor
See Garfield impact factor.

Journal Citation Reports (JCR)
The Journal Citation Reports, a product of ISI, provides quantitative measures for ranking, evaluating, categorizing, and comparing journals. The impact factor is one of these.

journal impact factor
This is a measure giving the relative number of citations received by a journal. There exist several different versions (see synchronous and diachronous impact factor) which are all useful in clarifying the significance of absolute citation frequencies.

journal self-cited rate (SCD-rate)
The self-cited rate relates a journal's self-citations to the number of times it is cited by all journals in the citation pool. See text for a mathematical description of the SCD rate.

journal self-citing rate (SCG-rate)
The self-citing rate of a journal relates a journal's self-citations to the total number of references it gives. See text for a mathematical description of the SCG rate.

Matthew effect
The term refers to the observation that already famous people (or journals) receive more credit than they actually deserve, while recognition of less prestigious scientists (or journals) is withheld. The term derives its name from the Gospel according to St. Matthew.

meta-journal
A group of journals considered for evaluation (or other) purposes as one large journal.

PDF format
Adobe® Portable Document Format (PDF) is a universal file format that preserves all of the fonts, formatting, colors, and graphics of any source document, regardless of the application and platform used to create it. PDF files are compact and can be shared, viewed, navigated, and easily printed.

popularity factor
The ratio of the number of journals citing a journal (during a particular period) over the number of journals cited by this journal.

PostScript (PS) format
PostScript is a device-independent high-level programming language for describing the appearance of text and graphics on a printed page.

Science Citation index (SCI)
The ISI Science Citation Index provides access to current and retrospective bibliographic information, author abstracts, and cited references found in 3,500 leading scientific and technical journals covering more than 150 disciplines. The Science Citation Index Expanded format available through the Web of Science and the online version, SciSearch, covers more than 5,700 journals.

standard impact factor
See Garfield impact factor.

synchronous impact factors (IF)
A group of impact factors using citations received in the same year, but referring to different publication years. See Appendix for an exact mathematical formulation.

REFERENCES
Abbott, A. (1999). University libraries put pen to paper in journal pricing protest. *Nature, 398,* 740.

Adler, S. (1999). The slashdot effect. Retrieved 15 February 2002 from http://ssadler.phy.bnl.gov/adler/SDE/SlashDotEffect.html

Bonitz, M.; Bruckner, E.; & Scharnhorst, A. (1997). Characteristics and impact of the Matthew effect for countries. *Scientometrics, 40*(3), 407–422.

Bonitz, M.; Bruckner, E.; & Scharnhorst, A. (1999). The Matthew index—Concentration patterns and Matthew core journals. *Scientometrics, 44*(3), 361–378.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. In *Proceedings of World-Wide Web '98* (WWW7). Retrieved 15 February 2002 from www7.scu.edu.au/programme/fullpapers/1921/com1921.htm

Brookes, B. C. (1970). Obsolescence of special library periodicals: Sampling errors and utility contours. *Journal of the American Society for Information Science, 21*(5), 320–329.

Christenson, J. A., & Sigelman, L. (1985). Accrediting knowledge: Journal stature and citation impact in social science. *Social Science Quarterly, 66*(4), 964–975.

de Bruin, R.; Kint, A.; Luwel, M.; & Moed, H. F. (1993). A study of research evaluation and planning: The University of Ghent. *Research Evaluation, 3,* 25–41.

de Bruin, R., & Moed, H. F. (1993). Delimitation of scientific subfields using cognitive words from corporate addresses in scientific publications. *Scientometrics, 26*(1), 65–80.

Dierick, J., & Rousseau, R. (1988). De impactfactor voor tijdschriften: een parameter bij het bepalen van een—al dan niet defensief—collectiebeleid? In J. Van Borm & L. Simons (Eds.), *Het oude en het nieuwe boek. De oude en de nieuwe bibliotheek* (pp. 593–601). Kapellen: DNB/Pelckmans.

Egghe, L. (1998). The evolution of core collections can be described via Banach space valued stochastic processes. *Mathematical and Computer Modelling, 28*(9), 11–17.

Egghe, L. (1999). An application of martingales in the limit to a problem in information science. *Mathematical and Computer Modelling, 29*(5), 13–18.

Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics: Quantitative methods in library, documentation and information science.* Amsterdam: Elsevier.

Egghe, L., & Rousseau, R. (1996). Average and global impact of a set of journals. *Scientometrics, 36*(1), 97–107.

Egghe, L., & Rousseau, R. (2000). Aging, obsolescence, impact, growth, and utilization: Definitions and relations. *Journal of the American Society for Information Science, 51*(11), 1004–1017.

Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics, 1*(4), 359–375.

Garfield, E. (1990). How ISI selects journals for coverage: Quantitative and qualitative considerations. *Current Contents, 22,* 5–13.

Garfield, E. (1997). A statistically valid definition of bias is needed to determine whether the Science Citation Index discriminates against third world journals. *Current Science, 73*(8), 639–641.

Garfield, E. (1998). Long-term vs. short-term journal impact: Does it matter? *The Scientist, 12*(3), 10–12.

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation, 14*(3), 195–201.

Glänzel, W.; Schubert, A.; & Czerwon, H.-J. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics, 44*(3), 427–439.

Gómez, I.; Coma, L.; Morillo, F.; & Cami, J. (1997). Medicina Clinica (1992–1993) vista a través del Science Citation Index. *Medicina Clínica, 109*(13), 497–505.

Harter, S. P. (1998). Scholarly communication and electronic journals: An impact study. *Journal of the American Society for Information Science, 49*(6), 507–516.

Hirst, G. (1978). Described impact factors: A method for determining core journal listings. *Journal of the American Society for Information Science, 29*(4), 171–172.

Holling, C. S. (1999). Lessons for sustaining ecological science and policy through the Internet. *Conservation Ecology, 3*(2), 16. Also available at http://www.consecol.org/vol3/iss2/art16.

Ingwersen, P.; Larsen, B.; Rousseau, R.; & Russell, J. (2001). The publication-citation matrix and its derived quantities. *Chinese Science Bulletin, 46*(6), 524–528.

Jin, B., & Wang, B. (1999). Chinese Science Citation Database: Its construction and application. *Scientometrics, 45*(2), 325–332.

Kleinberg, J. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM, 46*(5), 604–632.

Lewison, G., & Dawson, G. (1997). The effect of funding on the outputs of biomedical research. In B. Peritz & L. Egghe (Eds.), *Proceedings of the 6th Conference of the International Society for Scientometrics and Informetrics* (pp. 229–238). Jerusalem: Hebrew University of Jerusalem.

Line, M. (1977). On the irrelevance of citation analyses to practical librarianship. In W. E. Batten (Ed.), *EURIM II: A European conference on the application of research in information services in libraries* (pp. 51–53). London: Aslib.

Line, M. (1993). Changes in the use of literature with time—Obsolescence revisited. *Library Trends, 41*(4), 665–683.

Luwel, M.; Moed, H.; Nederhof, A.; De Samblanx, V.; Verbrugghen, K.; & van der Wurff, L. J. (1999). *Towards indicators of research performance in the social sciences and humanities.* VLIR report (d/1999/2939/9).

McGrath, W. E. (1993). The reappearance of rankings: Reliability, validity, explanation, quality, and the mission of library and information science. *Library Quarterly, 63*(2), 192–198.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56–63.

Moed, H. F.; Burger, W. J. M.; Frankfort, J. G.; & van Raan, A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy, 14,* 131–149.

Moed, H. F., & van Leeuwen, T. N. (1996). Impact factors can mislead. *Nature, 381,* 186.

Nakamoto, H. (1988). Synchronous and diachronous citation distributions. In L. Egghe & R. Rousseau (Eds.), *Informetrics 87/88* (pp. 157–163). Amsterdam: Elsevier.

Nieuwenhuysen, P., & Rousseau, R. (1988). A quick and easy method to estimate the random effect on citation measures. *Scientometrics, 13*(1–2), 45–52.

Pao, M. L., & Goffman, W. (1990). Quality assessment of schistosomiasis literature. In L. Egghe & R. Rousseau (Eds.), *Informetrics 89/90* (pp. 229–242). Amsterdam: Elsevier.

Peritz, B. C. (1983). Are methodological papers more cited than theoretical or empirical ones? The case of sociology. *Scientometrics, 5*(4), 211–218.

Peritz, B. C. (1995). On the association between journal circulation and impact factor. *Journal of Information Science, 21*(1), 63–67.

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with applications to the literature of physics. *Information Processing and Management, 12*(5), 297–312.

Price, D. J. de Solla (1970). Citation measures of hard science, soft science, technology, and nonscience. In C. E. Nelson & D. K. Pollock (Eds.), *Communication among scientists and engineers* (pp. 3–22). Lexington, MA: Heath.

Raisig, L. M. (1960). Mathematical evaluation of the scientific serial. *Science,* 131(3411), 1417–1419.

Rao, I. K. R. (1973). Obsolescence and utility factors of periodical publications: A case study. *Library Science with a Slant to Documentation, 10*(3), 297–307.

Rousseau, R. (1988). Citation distribution of pure mathematics journals. In L. Egghe & R. Rousseau (Eds.), *Informetrics 87/88* (pp. 249–262). Amsterdam: Elsevier.

Rousseau, R. (1993). A note on maximum impact factors. In *Information as a global commodity—Communication, processing and use* (pp. 120–125). (21st Annual Conference of the Canadian Association for Information Science.)

Rousseau, R. (1995). A scientometric study of the scientific publications of LUC. Period 1981–1993. LUC report.

Rousseau, R. (1997a). The proceedings of the first and second international conferences on bibliometrics, scientometrics and informetrics: A data analysis. In B. Peritz and L. Egghe, *Proceedings of the 6th Conference of the International Society for Scientometrics and Informetrics,* Jerusalem (pp. 371–380). Jerusalem: Hebrew University of Jerusalem.

Rousseau, R. (1997b). Sitations: An exploratory study. *Cybermetrics, 1*(1). http://www.cindoc.csic.es/cybermetrics/articles/v1i1p1.html

Rousseau, R. (1998a). Indicadores bibliométricos e econométricos para a avaliaçao de instituiçoes científicas. *Ciência da Informaçao, 27*(2), 149–158.

Rousseau, R. (1998b). *A scientometric study of the scientific publications of LUC. Period 1984–1996.* LUC report.

Rousseau, R. (1999) Temporal differences in self-citation rates of scientific journals. *Scientometrics, 44*(3), 521–531.

Rousseau, R.; Jin, B.; Yang, N.; & Liu, X. (2001). Observations concerning the two- and three-year synchronous impact factor, based on the Chinese Science Citation Database. *Journal of Documentation, 57*(3), 349–357.

Rousseau, R., & Spinak, E. (1996). Do a field list of internationally visible journals and their journal impact factors depend on the initial set of journals? A research proposal. *Journal of Documentation, 52*(4), 449–456.

Rousseau, R., & Van Hooydonk, G. (1996). Journal production and journal impact factors. *Journal of the American Society for Information Science, 47*(10), 775–780.

Russell, J., & Rousseau, R. (in print). Bibliometrics and institutional evaluation. In *Encyclopedia of Life Support Systems (EOLSS)*, Part 1.30, Science and Technology Policy, Rigas Arvantis (Ed.). Paris: UNESCO.

Schoonbaert, D., & Roelants, G. (1998). Impact takes precedence over interest. *Nature, 391*, 222.

Schubert, A., & Braun, T. (1993). Reference standards for citation based assessments. *Scientometrics, 26*(1), 21–35.

Schubert, A., & Glänzel, W. (1983). Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics, 5*(1), 59–74.

Sen, B. K.; Karanjai, A.; & Munshi, U. M. (1989). A method for determining the impact factor of a non-SCI journal. *Journal of Documentation, 45*(2), 139–141.

Spaventi, J.; Tudor-Silovic, N.; Maricic, S.; & Labus, M. (1979). Bibliometrijska analiza znanstvenih casopisa iz Jugoslavije (Bibliometric analysis of scientific journals from Yugoslavia). *Informatologia Yugoslavica, 11*(3), 11–23.

Spinak, E. (1995). Quantitative analyses of scientific literature and their validity for judging Latin American production. *Bulletin of PAHO, 29*(14), 352–359.

Spruyt, E. H. J.; de Bruin, R. E.; & Moed, H. F. (1996). Are bibliometric indicators appropriate tools in a young university? *Higher Education Management, 8*(3), 141–154.

Stegmann, J. (1997). How to evaluate journal impact factors. *Nature, 390*, 550.

Stegmann, J. (1999). Building a list of journals with constructed impact factors. *Journal of Documentation, 55*(3), 310–324.

Testa, J. (1998). A base de dados ISI e seu processo de seleçao de revistas. *Ciência da Informaçao, 27*(2), 233–235.

Tijssen, R. J. W., & van Raan, A. F. J. (1990). Net citation balances: A measure of influence between scientific journals. *Journal of the American Society for Information Science, 41*(4), 296–304.

Todorov, R., & Glänzel, W. (1988). Journal citation measures: A concise review. *Journal of Information Science, 14*(1), 47–56.

Tsay, M.-Y. (1998). The relationship between journal use in a medical library and citation use. *Bulletin of the Medical Library Association, 86*(1), 31–39.

Tsay, M.-Y. (1999). Library journal use and citation age in medical science. *Journal of Documentation, 55*(5), 543–555.

Valdecasas, A. G.; Castroviejo, S.; & Marcus, L. F. (2000). Reliance on the citation index undermines the study of biodiversity. *Nature, 403*, 698.

Van Fleet, D. D.; McWilliams, A.; & Siegel, D. S. (2000). A theoretical and empirical analysis of journal rankings: The case of formal lists. *Journal of Management, 26*(5), 839–861.

Van Hooydonk, G. (1995). Cost and citation data for 5399 scientific journals in connection with journal price-setting, copyright laws and the use of bibliometric data for project review. *The Serials Librarian, 27*(1), 45–58.

Van Hooydonk, G.; Gevaert, R.; Milis-Proost, G.; Van de Sompel, H.; & Debackere, K. (1994). A bibliotheconomic analysis of the impact factors of scientific disciplines. *Scientometrics, 30*(1), 65–81.

van Raan, A. F. J. (2000). The Pandora's box of citation analysis: Measuring scientific excellence—The last evil? In B. Cronin & H. B. Atkins (Eds.), *The web of knowledge. A Festschrift in honor of Eugene Garfield* (pp. 301–319). Medford, NJ: Information Today.

Wormell, I. (1998). Informetric analysis of the international impact of scientific journals: How 'international' are the international journals? *Journal of Documentation, 54*(5), 584–605.

Wouters, P. (1999). *The citation culture*. Unpublished doctoral disseration, University of Amsterdam.

Yanovsky, V. I. (1981). Citation analysis significance of scientific journals. *Scientometrics, 3*(3), 223–233.

Zwemer, R. L. (1970). Identification of journal characteristics useful in improving input and output of a retrieval system. *Federation Proceedings, 29*, 1595–1604.