Selecting Manuscripts for a High-Impact Journal Through Peer Review: A Citation Analysis of Communications That Were Accepted by *Angewandte Chemie International Edition*, or Rejected but Published Elsewhere

Lutz Bornmann
ETH Zurich, Professorship for Social Psychology and Research on Higher Education, Zähringerstr 24, CH-8092 Zurich. E-mail: bornmann@gess.ethz.ch

Hans-Dieter Daniel
ETH Zurich, Professorship for Social Psychology and Research on Higher Education, Zähringerstr 24, CH-8092 Zurich, and University of Zurich, Evaluation Office, Mühlegasse 21, CH-8001 Zurich. E-mail: daniel@evaluation.uzh.ch

All journals that use peer review have to deal with the following question: Does the peer review system fulfill its declared objective to select the "best" scientific work? We investigated the journal peer-review process at *Angewandte Chemie International Edition* (AC-IE), one of the prime chemistry journals worldwide, and conducted a citation analysis for Communications that were accepted by the journal ($n = 878$) or rejected but published elsewhere ($n = 959$). The results of negative binomial-regression models show that holding all other model variables constant, being accepted by AC-IE increases the expected number of citations by up to 50%. A comparison of average citation counts (with 95% confidence intervals) of accepted and rejected (but published elsewhere) Communications with international scientific reference standards was undertaken. As reference standards, (a) mean citation counts for the journal set provided by Thomson Reuters corresponding to the field "chemistry" and (b) specific reference standards that refer to the subject areas of *Chemical Abstracts* were used. When compared to reference standards, the *mean* impact on chemical research is for the most part far above average not only for accepted Communications but also for rejected (but published elsewhere) Communications. However, average and below-average scientific impact is to be expected significantly less frequently for accepted Communications than for rejected Communications. All in all, the results of this study confirm that peer review at AC-IE is able to select the "best" scientific work with the highest impact on chemical research.

Reputable scientific journals only publish manuscripts that have been subjected to peer review—that is, critical scrutiny by scientific experts. When a manuscript is submitted, referees who research and publish in the same field (peers) are asked to evaluate the content of the manuscript (as to significance and originality of the research findings, for example) and recommend to the editor that the manuscript be published, revised and then published, or rejected (Sense about Science, 2004). The goal of this process is to ensure that the valid manuscript is accepted, the messy manuscript improved, and the invalid manuscript rejected. Peer review of contributions to the primary research literature is the principal social mechanism for quality control in academic science (Braun, 2004). Peer review is the "best" available mechanism for quality control (Kostoff, 1997), but it is not perfect (Marsh, Jayasinghe, & Bond, 2008; Weller, 2002). Peers are not seers, after all, but ordinary human beings with their own opinions, strengths, and weaknesses (Ehses, 2004). All journals that use peer review have to deal with the following question: Does the peer-review system fulfill its declared objective to select the best scientific work? The goal of our study is to investigate whether the peer-review process of the journal *Angewandte Chemie International Edition* (AC-IE) is in fact capable of validly selecting the manuscripts most worthy of publication.

AC-IE is one of the prime chemistry journals in the world, with a higher annual Journal Impact Factor (JIF, provided by Thomson Reuters) than the JIFs of comparable journals (10.232 in the 2006 *Journal Citation Reports: Science Edition*; Institute for Scientific Information, 2006). AC-IE is a journal of the German Chemical Society (Gesellschaft Deutscher Chemiker (GDCh), Frankfurt am Main, Germany) and is published by Wiley-VCH (Weinheim, Germany). It introduced peer review in 1982, primarily in conjunction with one of the document types published in the journal,

Received February 25, 2008; revised May 2, 2008; accepted May 2, 2008

 • Published online 9 July 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.20901

"Communications," which are short reports on work in progress or recently concluded experimental or theoretical investigations. What the editors of AC-IE look for most of all is excellence in chemical research. Communications that referees deem to be of high quality are selected for publication. Communications that do not meet the high standards are rejected. As there is broad support for citation counts of scientific publications as a measure of the impact on scientific research (Cole, 2000; Daniel, 2005; van Raan, 2004), our assumption is that rejected Communications earn lower citation counts than accepted Communications. As AC-IE also archives manuscripts that have been rejected for publication, we explored this hypothesis via investigation of the citation rates of the accepted manuscripts and also of manuscripts that were rejected by AC-IE but published elsewhere.

Citation rates have been a controversial measure of both quality and scientific progress (Bornmann & Daniel, 2008). Nevertheless, Lokker, McKibbon, McKinlay, Wilczynski, and Haynes (2008) succeeded in demonstrating for clinical articles that publications regarded shortly after their appearance as important by experts in the appropriate research field were cited much more frequently in subsequent years than publications that were less highly regarded. The Chemistry Division of the National Science Foundation (Arlington, VA) carried out a citation analysis with the goal "to explore the use of this relatively new tool for what it might tell about the discipline and its practitioners." The results of the study "generally support the idea that citations are meaningful" (Dewitt, Nicholson, & Wilson, 1980, p. 265). Furthermore, the results of a comprehensive citation-content analysis conducted by Bornmann and Daniel (2007b) show that "an article with high citation counts had greater relevance for the citing author than an article with low citation counts" (p. 149).

Method

Manuscript Review at AC-IE

A Communication submitted to AC-IE is usually subject to internal and external review. First, editors at the journal evaluate whether the Communication contributes to the development of an important area of research (internal review). If the editors find that this is so, the submitted Communication is sent to several independent referees (external review), who review it using an evaluation form and a comment sheet. The journal editors then make the decision to accept or reject a Communication for publication on the basis of these reviews and on their own evaluations.

Database for the Present Study

For the investigation of manuscript review at AC-IE, we used information on all 1,899 Communications that were reviewed in the year 2000. The information was taken from documents archived by the publisher, Wiley-VCH. Of the 1,899 Communications, 46% ($n = 878$) were accepted for publication in AC-IE, and 54% ($n = 1,021$) were rejected. A search in the literature databases *Science Citation Index Expanded* (SCI, Thomson Reuters) and *Chemical Abstracts* (CA, Chemical Abstracts Services) revealed that of the 1,021 rejected manuscripts, 959 (94%) were then published in 136 other (different) journals: 723 in 21 journals (more than 9 Communications published in each) and 236 in 115 journals (fewer than ten published in each). Fifty or more rejected Communications were published later in each of the following journals: *Chemical Communications* ($n = 119$), *Organic Letters* ($n = 91$), *Journal of the American Chemical Society* ($n = 70$), *Tetrahedron Letters* ($n = 60$) and *Organometallics* ($n = 50$). Our results in Bornmann and Daniel (in press, Supporting Information) show that the authors of about 75% of the rejected Communications did not change or only marginally changed the content of the manuscript for publication elsewhere. About one fourth of the rejected Communications were either changed to a substantial extent, or the content of the rejected Communication was published together with other research results (see Cronin & McKenzie, 1992). The assessment of the extent of changes in the rejected Communications for publication elsewhere was carried out by a scientist with a doctoral degree in chemistry (in collaboration with other members of our research team).

Conducting the Citation Analysis

For accepted and rejected (but published elsewhere) Communications, we determined the total number of citations up to the end of 2006 and the number of citations for a fixed time window of three years after the publication year. "Fixed citation windows are a standard method in bibliometric analysis, in order to give equal time spans for citation to articles published in different years, or at different times in the same year" (Craig, Plume, McVeigh, Pringle, & Amin, 2007, p. 243). According to Harnad (2007), the reliability of the results of citation analysis can be tested by using two different citation windows (here, we used a fixed three-year window, and the time window from publication up to the end of 2006).

In the citation search we included self-citations, because (a) it is not expected that the number of self-citations varies systematically for the accepted and rejected (but published elsewhere) manuscripts, and (b) the number of self-citations of a publication can be modeled in the multiple-regression analysis (the results of which are reported below) using the number of authors of a manuscript. As Herbertz (1995) shows, a greater number of authors is associated with a greater number of self-citations of a publication (see also Leimu & Koricheva, 2005).

The citation analyses for the present study were conducted based on both SCI and CA. The SCI includes multidisciplinary data from journals in the sciences (see http://isiwebofknowledge.com/). CA is a comprehensive database of publicly disclosed research in chemistry and related sciences, including the world's largest collection of substance information (see http://www.cas.org/). A study by Whitley (2002) comparing citation searching in SCI and CA showed that the two indices lead to different results. According to Whitley, CA misses an estimated 17% of the citations found

TABLE 1. Mean number of citations of Communications accepted by AC-IE, and of Communications rejected by AC-IE but published elsewhere, searched in *Science Citation Index* and in *Chemical Abstracts* for two different citation windows.

			Citations					
Publication year	Editorial decision	Number of communications	Minimum	Maximum	Arithmetic average	Standard deviation	Median
Science Citation Index (citation window: from year of publication to the end of 2006)[a]							
2000	Accepted	528	0	194	31.14	28.66	23.00
	Rejected	254	0	160	21.79	21.73	15.00
	Difference				**9.35**		**8.00**[b]
2001	Accepted	350	1	263	29.66	28.79	22.00
	Rejected	538	0	165	19.05	19.46	13.00
	Difference				**10.61**		**9.00**[c]
Chemical Abstracts (citation window: from year of publication to the end of 2006)[a]							
2000	Accepted	528	0	218	33.86	29.97	25.00
	Rejected	254	0	173	23.39	22.73	16.00
	Difference				**10.47**		**9.00**[d]
2001	Accepted	350	1	282	31.95	29.16	24.00
	Rejected	538	0	182	20.75	20.42	15.00
	Difference				**11.20**		**9.00**[e]
Science Citation Index (citation window: the first three years after the publication year)[f]							
2000–2003	Accepted	878	0	126	15.99	14.12	12
	Rejected	939	0	93	10.64	10.67	8
	Difference				**5.35**		**4.00**[g]
Chemical Abstracts (citation window: the first three years after the publication year)[f]							
2000–2003	Accepted	878	0	136	17.62	14.90	14
	Rejected	939	0	107	11.72	11.36	9
	Difference				**5.90**		**5.00**[h]

Note. [a]167 rejected Communications that were published elsewhere between 2002 and 2006 are not shown, because for those years there are no accepted Communications available to serve as a comparison group.

[b]Mann-Whitney U-Test: $Z(n = 782) = -5.72$, $p < 0.05$

[c]Mann-Whitney U-Test: $Z(n = 888) = -7.59$, $p < 0.05$

[d]Mann-Whitney U-Test: $Z(n = 782) = -6.31$, $p < 0.05$

[e]Mann-Whitney U-Test: $Z(n = 888) = -7.99$, $p < 0.05$

[f]All of the Communications that were published between 2000 and 2003 by AC-IE or elsewhere could be included in the analysis ($n = 1{,}817$). For 20 rejected Communications that were published elsewhere between 2004 and 2006, a citation window of three years (one year after publication up to the end of 2006) was not available.

[g]Mann-Whitney U-Test: $Z(n = 1{,}817) = -10.64$, $p < 0.05$

[h]Mann-Whitney U-Test: $Z(n = 1{,}817) = -11.36$, $p < 0.05$

in SCI. Conversely, researchers using only SCI to search citations miss an estimated 23% that are covered only in CA.

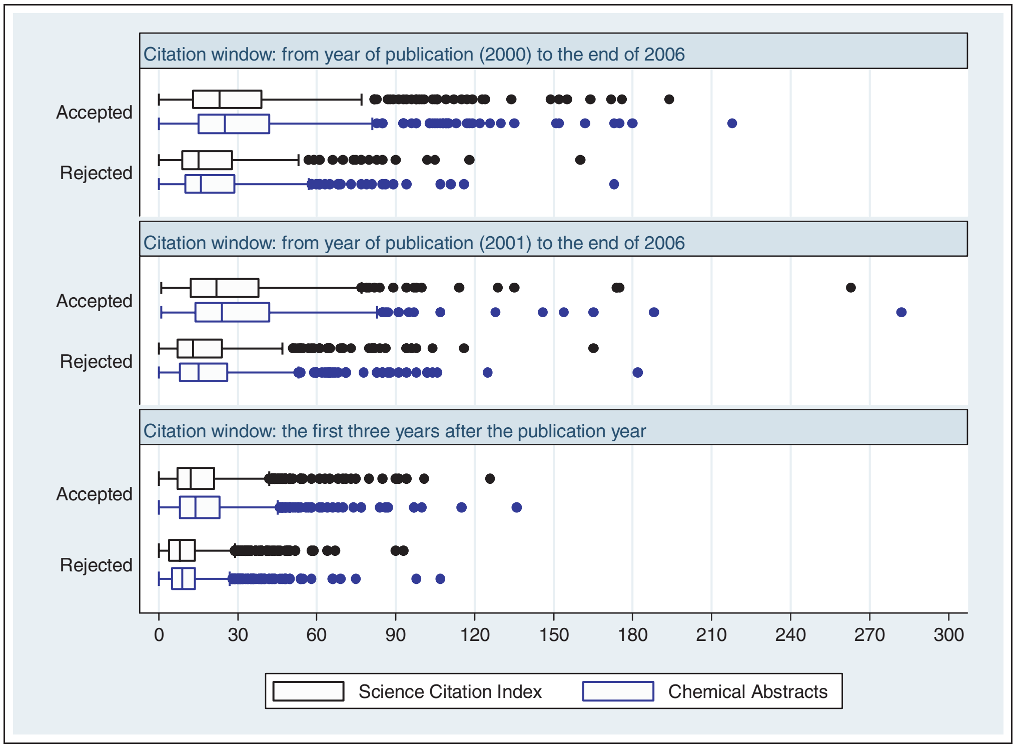
Results

Did manuscript review at AC-IE actually achieve its goal of selecting the best Communications, showing the highest impact on chemical research? As shown in Table 1, the findings provide evidence that it did.

Differences in Mean Citation Counts

Table 1 shows the mean number of citations found in SCI and CA for two different citation windows for Communications. For example, Communications that were accepted by AC-IE and published in the year 2000 were cited, according to the SCI, on average 31.14 times (arithmetic mean) up to the end of 2006; Communications that were rejected by AC-IE and were published elsewhere in 2000 were cited on average 21.79 times up to the end of 2006. The difference between the mean citation counts of accepted und rejected Communications is 9.35. As the results in Table 1 show, independent of publication year, literature database in which the citation search was conducted, and citation time window, accepted Communications are cited more frequently on average than rejected (but published elsewhere) Communications. Similar results were obtained by using Scopus (Elsevier, Amsterdam, The Netherlands) as a data source for citation counts (Bornmann & Daniel, in press).

As arithmetic-mean citation counts can be affected by outliers (since high citation counts skew the mean high), we prepared three graphs (see Figure 1) with the distributions of citation counts for the accepted and rejected Communications.



Note. See Table 1 for further information, for example for the numbers of accepted and rejected Communications and for the statistically significant differences between the individual median values.

FIG. 1. Box plots for citation counts, found in *Science Citation Index* and in *Chemical Abstracts* by citation searching for two different citation windows, of Communications accepted by AC-IE and of Communications rejected by AC-IE but published elsewhere.

The box plots in Figure 1 show the interquartile ranges (length of the boxes) and medians (vertical lines through the boxes) of citations found in SCI and in CA for both citation time windows (see Kohler & Kreuter, 2005). As the graphs show, the distributions of the citation counts for accepted and rejected Communications are characterized by a multitude of outliers (this finding accords with the high standard deviations in Table 1; see also Bornmann & Daniel, 2007a). Unlike the arithmetic mean, the median is not affected by outliers. Figure 1 and Table 1 show that independent of publication year, literature database in which the citation search was conducted, and citation window, the median citation counts for accepted Communications are consistently statistically significantly higher than the median citation counts for rejected (but published elsewhere) Communications (see the results of the statistical tests in the note to Table 1).

Results of Negative Binomial-Regression Models

While the arithmetic means and medians in Table 1 and Figure 1 suggest that the AC-IE review process indeed selects the "better" Communications among the submissions for publication, factors other than their scientific contribution to the development of an important area of research could in principle have been responsible for the higher citation counts. Bibliometric studies have demonstrated that several factors have a *general* influence on citation counts. By considering these factors in the statistical analysis, it becomes possible to establish the adjusted covariation between editorial decisions and citation counts of accepted and rejected (but published elsewhere) Communications.

The probability of citation may be influenced by the number of authors (Glänzel, Debackere, Thijs, & Schubert, 2006; Leimu & Koricheva, 2005; see above), the number of pages in a publication (Bornmann & Daniel, 2007c), the language of the journal in which a publication appears, and the field or discipline to which the journal can be assigned (Bornmann & Daniel, 2008). In addition, consistent with Robert K. Merton's interpretation of the Matthew effect in science (Merton, 1968) and Cozzens' (1985) "success-breeds-success" phenomenon, publications by authors whose works have been very frequently cited in the past can be expected to be cited more often than publications by authors who have not published highly cited works in the past. This means that publications of the same intrinsic worth will be cited differently depending on the status of the author (see also Garfield, 2002). The JIF (Callaham, Wears, & Weber, 2002) was not considered as a predictive factor in the statistical analysis because "article citation rates determine the journal impact factor, not vice versa" (Seglen, 1997). For Leimu and Koricheva (2005) there is a widespread belief "that publication in a high-impact journal might by itself enhance the citation rate of an article by increasing its visibility or persuasiveness of the arguments presented" (p. 29). Their study results

TABLE 2. Description of the independent variables.

Independent variable	Values	Mean value
Models A and B: *Science Citation Index* and *Chemical Abstracts* (citation window: from year of publication to the end of 2006) (n = 1,829)[a]		
Editorial decision	rejected (0) → accepted (1)	0.48
Number of authors of the publication	1 → 16	4.26
Length of the publication (in number of pages)	2 → 19	4.08
Language of the journal in which the publication appeared	English (0) → multiple or non-English language (1)	0.07
Organic chemistry[b]	Other area (0) → organic chemistry (1)	0.51
Physical, inorganic, and analytical chemistry[b]	Other area (0) → physical, inorganic, and analytical chemistry (1)	0.30
Macromolecular chemistry[b]	Other area (0) → macromolecular chemistry (1)	0.06
Applied chemistry[b]	Other area (0) → applied chemistry (1)	0.03
Number of authors listed in ISIHighlyCited.com	0 → 0.5	0.03
Models C and D: *Science Citation Index* and *Chemical Abstracts* (citation window: the first three years after the publication year) (n = 1,810)[a]		
Editorial decision	rejected (0) → accepted (1)	0.48
Number of authors of the publication	1 → 16	4.25
Length of the publication (in number of pages)	2 → 19	4.06
Language of the journal in which the publication appeared	English (0) → multiple or non-English language (1)	0.07
Organic chemistry[b]	Other area (0) → organic chemistry (1)	0.51
Physical, inorganic, and analytical chemistry[b]	Other area (0) → physical, inorganic, and analytical chemistry (1)	0.30
Macromolecular chemistry[b]	Other area (0) → macromolecular chemistry (1)	0.06
Applied chemistry[b]	Other area (0) → applied chemistry (1)	0.03
Number of authors listed in ISIHighlyCited.com	0 → 0.5	0.03

Note. [a]Of the total of 1,837 (models A and B) or 1,817 (models C and D) publications, 8, or 7, could not be included in the regression analyses due to missing information on either the language of the journal in which the publication appeared or the section assignment (only those cases can be included in the statistical analyses that have no missing values for the variables entered into the model).

[b]In the regression analysis, the biochemistry section forms the reference category.

"do not support this 'journal effect' hypothesis, because there was considerable variation in citation rates, especially for papers published in high-impact journals" (p. 29).

We performed a multiple-regression analysis, which reveals the factors that exert a primary influence on a certain outcome. The coefficients in the regression model, called "partial" regression coefficients (Rabe-Hesketh & Everitt, 2004), represent the effects of each factor, controlling for all other factors in the model. Since the skewness of citation counts (see Table 1 and Figure 1) suggests the use of a negative binomial specification (Glänzel & Schubert, 1993), we calculated a negative binomial-regression model (NBRM; Long & Freese, 2006, section 8.3; see also Bornmann & Daniel, 2006). The citation counts for the accepted and rejected (but published elsewhere) Communications enter into the NBRM as a dependent variable. As we performed the citation search for the Communications in CA and SCI for two different citation windows, we calculated a total of four different models (models A–D).

Table 2 shows a description of the independent variables that were included in the NBRMs. In addition to the editorial decision (approval or rejection), the models take the number of authors, number of pages of the publications, the language (English, or non-English or multiple) of the journal in which the publication appeared, and the field or discipline into account. CAS categorizes publications in 80 different sections, each section covering only one broad area of scientific inquiry. Each abstract in CA appears in only one CA section, according to the most important aspect of the publication. The 80 individual sections are listed according to five broad headings, for the five main areas of chemical research: (a) organic, (b) physical, inorganic, and analytical, (c) macromolecular, (d) applied, and (e) biochemistry. The accepted and rejected manuscripts' assignments to the five main areas were entered into the regression analysis (see Table 2). The status of the authors of a publication as most highly influential scientists in an area of research was entered into the model as the number of authors that are listed in ISIHighlyCited.com, which is a part of the *ISI Web of Knowledge* platform provided by Thomson Reuters. ISIHighlyCited.com provides "a tool to identify individuals . . . that have made fundamental contributions to the advancement of science and technology in recent decades . . . These individuals are the most highly cited within each category for the period 1981 to 1999" (ISIHighlyCited.com, 2008). There are 21 subject categories in life sciences, medicine, physical sciences, engineering, and social sciences.

In models A and B the publication year of each accepted and rejected (but published elsewhere) Communication was included in the models as exposure time (Long & Freese, 2006, pp. 370–372). By using the "exposure" option provided in the statistical package Stata (StataCorp, 2007), the amount of time that an article is "at risk" of being cited is considered. In models C and D this option is not needed, as the individual citation counts refer to the same citation window of three years. The violation of the assumption of independent observations caused by including citation counts of more than one publication per journal is considered in the models by using the "cluster" option in Stata (StataCorp). This option specifies that the citation counts are independent across papers

TABLE 3. Negative binomial-regression models for variables predicting citations found through citation searching in *Science Citation Index* and in *Chemical Abstracts* with two different citation windows.

	Citation window: From year of publication to the end of 2006		Citation window: The first three years after the publication year	
	Model A: *Science Citation Index*	Model B: *Chemical Abstracts*	Model C: *Science Citation Index*	Model D: *Chemical Abstracts*
Publication year (from 2000 to 2006)	(exposure)	(exposure)		
Editorial decision	0.411* (4.81)	0.419* (4.90)	0.341* (4.41)	0.345* (4.45)
Number of authors	0.0249 (1.57)	0.0236 (1.57)	0.0338* (2.44)	0.0318* (2.45)
Length of publication (number of pages)	−0.00650 (−0.58)	−0.00683 (−0.58)	0.0138 (1.13)	0.0147 (1.13)
Language of the journal	−0.613* (−3.67)	−0.570* (−3.48)	−0.553* (−3.57)	−0.507* (−3.31)
Organic chemistry	0.126* (2.20)	0.146* (2.67)	0.126* (2.36)	0.141* (2.72)
Physical, inorganic, and analytical chemistry	0.295* (4.81)	0.279* (4.44)	0.240* (4.56)	0.246* (4.38)
Macromolecular chemistry	0.275* (3.60)	0.283* (3.52)	0.243* (3.49)	0.261* (3.28)
Applied chemistry	0.0721 (0.52)	0.123 (0.87)	0.0197 (0.14)	0.105 (0.77)
Number of authors listed in ISIHighlyCited.com	0.609* (4.18)	0.623* (4.67)	0.579* (2.99)	0.619* (3.25)
n	1,829	1,829	1,810	1,810

Note. Standard errors are adjusted for the dependency in the dataset caused by the publication of several accepted and/ or rejected Communications in one journal.

t statistics are shown in parentheses.

* $p < 0.05$.

published in different journals, but are not necessarily independent within papers of the same journal (see Hosmer & Lemeshow, 2000, section 8.3).

In Table 3 the results of models A–D for predicting citation counts for accepted and rejected (but published elsewhere) Communications show similar results. As to the variable *editorial decision* the models yield the following results: a statistically significant greater number of citations are expected for accepted Communications than for rejected Communications. The calculation of the percent-change coefficients for the editorial decisions following the NBRM (see Long & Freese, 2006) shows that being accepted by AC-IE increases the expected number of citations by about 50% for models A and B, and about 40% for models C and D—holding all other variables constant. As assessed by SCI and CA citations for both citation windows, AC-IE editors were therefore able to accomplish the difficult task of assessing the scientific merit of the Communications absolutely accurately, and of selecting the "best" manuscripts among submissions.

In all of the models, statistically significant effects could be found in the expected directions for three factors that in bibliometric studies have been demonstrated to have a general influence on citation counts: (a) More citations are expected for a publication in an English-language journal than for a publication in a multiple-language or non-English language journal. (b) The results also show influence of the research area: Publications in organic; physical, inorganic, and analytical; and macromolecular chemistry were more frequently cited than publications in biochemistry. (c) A higher number of authors listed in ISIHighlyCited.com is associated with a greater number of citations of a publication. However, contrary to expectations, all of the models show no statistically significant effect for number of pages of a publication. This finding may be due to the fact that all of the publications that were entered into the model calculations were originally submitted to AC-IE as Communications and were thus similar in length. Even if a number of the rejected manuscripts were published elsewhere not as Communications but as full papers, it appears that their information content—despite any variation in number of pages—still usually accorded with a Communication (on this, see Bornmann & Daniel, 2007c).

As to number of authors, the results of the NBRMs vary greatly depending on the citation window: Whereas *number of authors* shows a statistically significant effect for the fixed citation window of three years, no statistically significant

DOI: 10.1002/asi

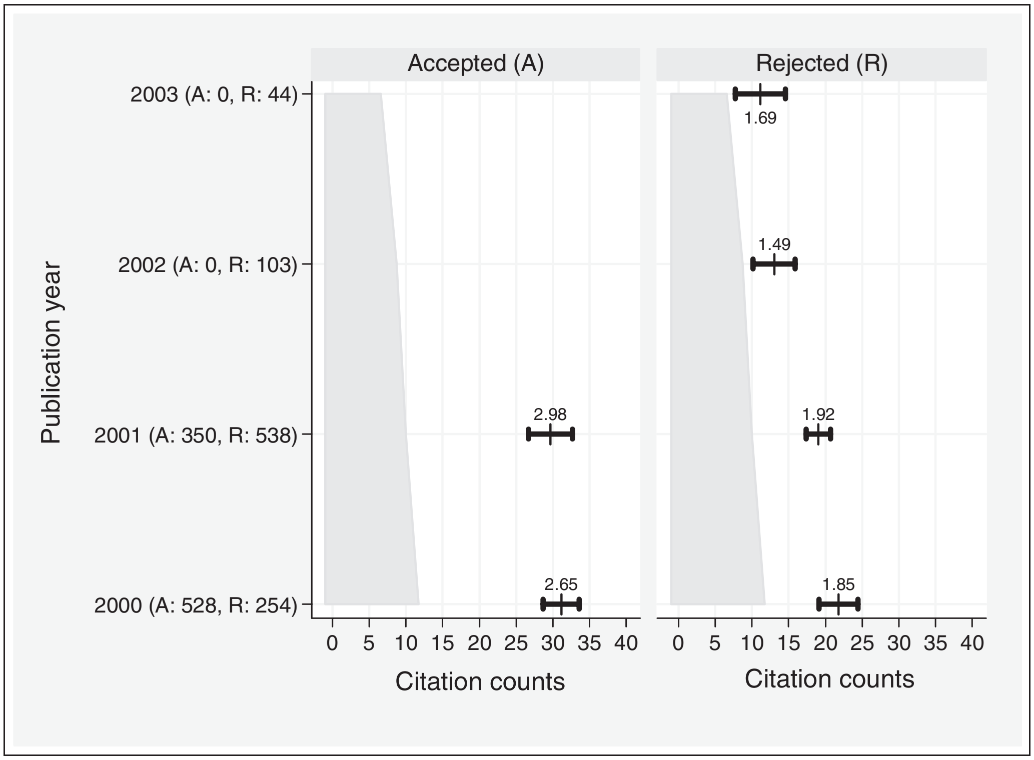


FIG. 2. Mean citation counts (*Science Citation Index*, citation window: from year of publication to the end of 2006), with 95% confidence intervals, of Communications accepted by AC-IE and of Communications rejected by AC-IE but published elsewhere ($n = 1{,}817$). The number of accepted (A) and rejected (R) Communications is shown in parentheses following the publication year. The gray-shaded area marks for each publication year citation counts below the citation-impact baseline for "chemistry." The value that is shown for the confidence intervals above or below the mean citation count is R_w (Relative Subfield Citedness).

effect was found for the citation window from year of publication to the end of 2006. The number of citations of a publication in the initial years after publication thus seems to be decisively dependent on the authors' self-citations (see above). In agreement with this assumption is the finding by Rousseau (1999) that citations of a publication by the authors themselves (self-citations) reach a clearly earlier peak than citations by colleagues (external citations).

Comparison with Reference Standards

Even if the findings in Table 1, Figure 1, and Table 3 show that the editors select Communications that have a higher impact than rejected (but published elsewhere) Communications, we still do not know whether the journal publishes "scientific excellence." Our intention when conducting bibliometric analyses is not only to find out whether AC-IE peer review is able to select the "better" chemical research papers but also to be able to identify "high-impact" submissions (see Research Evaluation and Policy Project, 2005). The latter question can be answered only by comparing the performance of approved and rejected Communications with international scientific reference standards. For this, Vinkler (1997) recommends a worldwide reference standard (see also van Raan, 1999): "Relative Subfield Citedness (R_w, where W refers to "world") relates the number of citations obtained by the set of papers evaluated to the number of citations received by the same number of papers published in journals dedicated to the respective discipline, field, or subfield" (p. 164, see also Vinkler, 1986).

To calculate R_w for the Communications in our study, as a first step we used mean citation counts for the journal set provided by Thomson Reuters (see Essential Science Indicators, ESI) corresponding to the field "chemistry" as a reference standard (Thomson Reuters, 2008). To determine R_w, we divided the (arithmetic) mean number of citations for accepted or rejected (but published elsewhere) Communications by the (arithmetic) mean number of citations of all publications in this journal set. In a recently published article in *Science*, Wuchty, Jones, and Uzzi (2007) define highly cited work "as receiving more than the mean number of citations for a given field" (p. 1037), that is, with $R_w > 1$. According to van Raan (2004) the R_w quotient allows more specific determination of whether the citation impact of the accepted and rejected (but published elsewhere) Communications is far below ($R_w < 0.5$), below ($R_w = 0.5$–0.8), approximately the same as ($R_w = 0.8$–1.2), above ($R_w = 1.2$–1.5), or far above ($R_w > 1.5$) the international (primarily the Western world) citation-impact baseline for the chemistry field. With R_w values above 1.5, the probability of identifying excellent contributions is very high (van Raan, 2004).

Figure 2 shows mean citation counts (SCI, citation window: from year of publication to the end of 2006), with 95% confidence intervals, of accepted Communications by AC-IE

DOI: 10.1002/asi

and of Communications that were rejected by AC-IE but published elsewhere, classified according to year of publication, 2000–2003. The mean citation counts and confidence intervals presented are restricted to these four publication years (the years 2004–2006 are not shown), because only for each of these years are 10 or more Communications available for a reliable estimation of the statistical values. The 95% confidence intervals in the figure show the range within which, with high probability, the "true" mean value of the citations for accepted or rejected Communications published in a given year lies (see Greenwood, 2007; Kline, 1998). Since citation counts are not normally distributed but follow a negative binomial distribution (see above), the confidence intervals were estimated using the bootstrap resampling method (see Stine, 1990). Figure 2 shows the confidence intervals; the value that is shown above or below the mean citation count is R_w. The (light) gray-shaded area in the figure depicts for each publication year citation counts that are below the citation-impact baseline for the chemistry field.

As the R_w values in Figure 2 for accepted or rejected (but published elsewhere) Communications published in a certain year show, five of the six values are above 1.5 (between 1.69 and 2.98), and one is between 1.2 and 1.5 (1.49). This ratio value just below 1.5 is for rejected Communications published in the year 2002. Depending on the number of accepted or rejected Communications in a publication year, the size of the confidence intervals in Figure 2 and thus also the exactness of the estimate of the "true" mean value of the citations for the Communications vary greatly. The differences between the lower and upper ends of the confidence intervals vary between 3.33 citations (rejected Communications published in 2001) and 6.85 citations (rejected Communications published in 2003). Despite this uncertainty in the estimates—as the relative positions of the 95% confidence intervals to the gray-shaded areas in the figure show, all lower ends of the confidence intervals are above the citation-impact baseline for "chemistry." According to the definition by Wuchty et al. (2007) therefore, with high certainty the research is (on average) highly cited in all of the publication years.

As AC-IE "publishes articles from the full spectrum of chemistry" according to the journal's chief editor Peter Gölitz (2005, p. 5539), the Thomson Reuters journal set "chemistry" provides only a very rough reference standard for assessing the scientific impact of accepted and rejected (but published elsewhere) Communications (see also the findings in Bornmann, Leydesdorff, & Marx, 2007). Citation practices vary even in different areas (or clusters) within a single subfield (Klamer & van Dalen, 2002; Lewison & Dawson, 1998). For this reason, Neuhaus and Daniel (in press) propose the use of specific reference standards that refer to the subject areas of CA (see also van Leeuwen, 2007). CAS categorizes chemical publications into 80 different subject areas (called sections; see above). Every publication becomes associated with a single principal entry, which makes clearly apparent the most important aspect of the work (Daniel, 1993/2004). In contrast to the journal sets provided by Thomson Reuters, CA sections are assigned on a paper-by-paper basis (Bornmann, Mutz, Neuhaus, & Daniel, 2008).

For the present study, we asked the Central Information Service for the institutes of the Chemical Physical Technical (CPT) Section of the Max Planck Society (located at the Max Planck Institute for Solid State Research in Stuttgart, Germany) to generate reference standards for 33 of the 80 CA sections. For each of these 33 sections, we have in the sample ten or more Communications for a reliable estimation of the statistics (mean and 95% confidence interval). The reference standard for a section is based on the publications of the year 2001 and the citations of these publications in the years 2002–2004 (fixed three-year citation window). The manuscripts accepted and rejected by AC-IE are mainly published as Communications and research articles. Because CAS "does not provide a distinct document type for research articles" (Neuhaus & Daniel, in press), the reference standards were generated by excluding publications with nonrelevant document types, such as conference proceedings and reviews. We compared the quotients of citations and publications (that is, the CA-section baselines, or area-specific reference standards) with the mean citation counts (CA, citation window: the first three years after the publication year) for accepted or rejected Communications that were published between 2000 and 2003.

Figure 3 shows (a) the mean citation counts for accepted or rejected (but published elsewhere) Communications with 95% confidence interval, shown according to the 33 CA sections, and (b) a gray-shaded area, which indicates citation counts below the citation-impact baseline for each section. The color of the confidence interval in Figure 3 indicates whether the "true" mean value for accepted or rejected Communications categorized in a section (with high certainty) is below or above the baseline. With green confidence intervals, the lower end of the citation-impact interval is above the baseline; with red confidence intervals it is below the baseline. For the confidence intervals, the section-specific R_w is shown above the mean citation count.

As the R_w values in Figure 3 show, a large part of the mean citation counts for the Communications is far above the baseline: of the total 66 mean citation counts, 59 are far above, 3 are above (rejected Communications categorized in "organometallic and organometalloidal compounds," $R_w = 1.44$; "enzymes," $R_w = 1.25$; and "electric phenomena," $R_w = 1.23$), 2 are the same as (rejected Communications categorized in "alkaloids," $R_w = 1.06$; and accepted Communications categorized in "general biochemistry," $R_w = 0.90$), and 2 are below (rejected Communications categorized in "biochemical genetics," $R_w = 0.79$; and "general biochemistry," $R_w = 0.64$) the baseline. Depending on the number of accepted or rejected (but published elsewhere) Communications categorized in a section, the exactness of the estimate of the "true" mean value of the citations for the Communications varies greatly (see the size of the 95% confidence intervals in the figure). Still, of the total 66 lower ends of the confidence intervals, 50 lie above (green confidence intervals) and only 16 below (red confidence intervals)

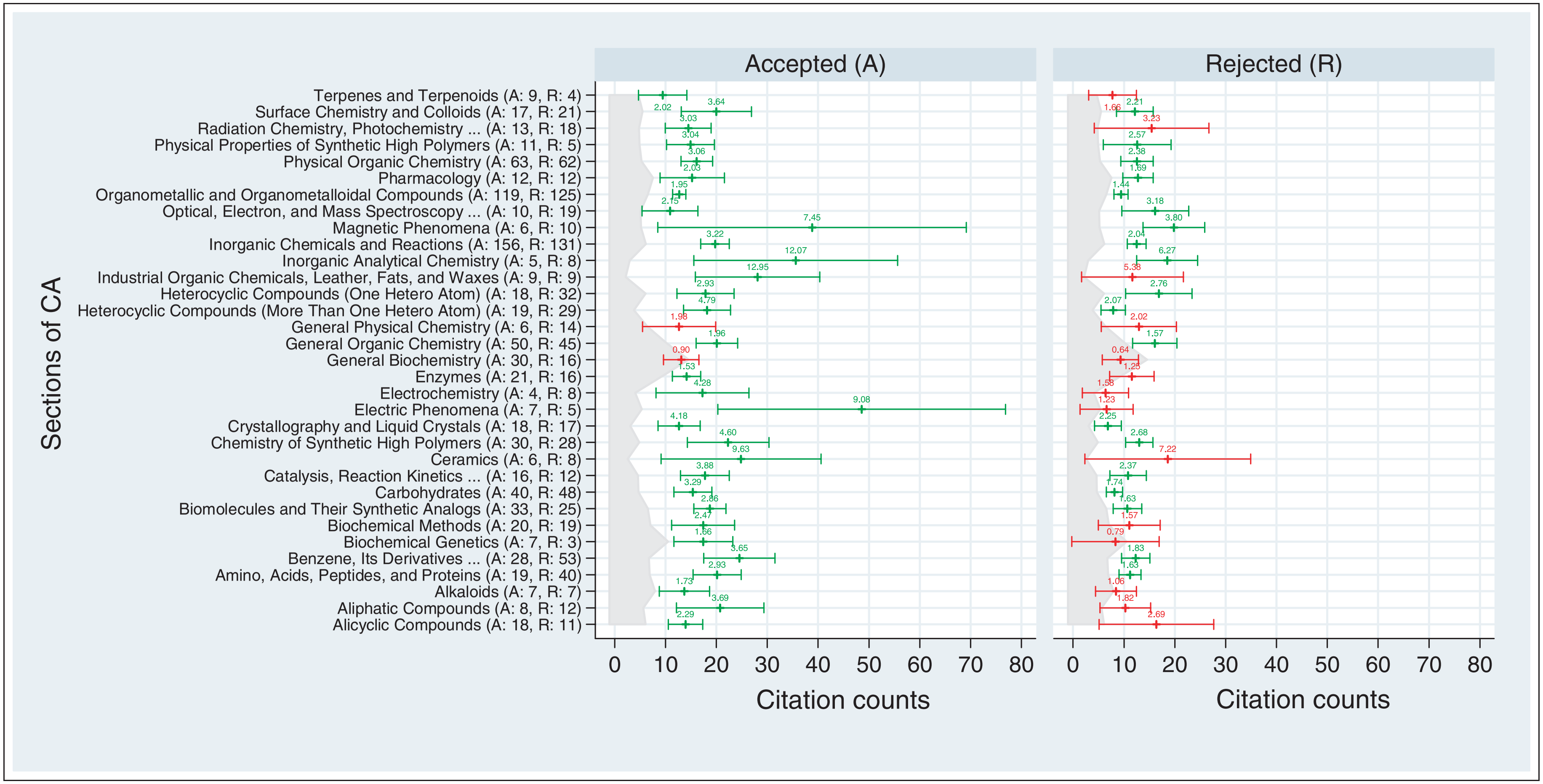


FIG. 3. Mean citation counts (*Chemical Abstracts*, CA, citation window: the first three years after the publication year) with 95% confidence intervals of Communications accepted by AC-IE and of Communications that were rejected by AC-IE but published elsewhere ($n = 1{,}707$). For each section the number of accepted (A) and rejected (R) Communications is shown in parentheses. The gray-shaded area indicates, for each section of *Chemical Abstracts*, citation counts below the citation-impact baseline. Looking at the confidence intervals, the value shown above or below the mean citation count is R_w (Relative Subfield Citedness). With the green confidence intervals, the lower end of the citation-impact interval is above the baseline; with red confidence intervals the lower end is below the baseline.

the citation-impact baseline for the corresponding section. Of these 16 values, 2 refer to the group of accepted Communications ("general biochemistry" and "general physical chemistry") and 14 to the group of rejected (but published elsewhere) Communications. With regard to the distribution of these values, the difference between accepted and rejected Communications is statistically significant, χ^2 (1, $n = 66) = 11.88$, $p < 0.05$. The effect size of this result is medium (Cramer's V = −0.42).

Thus, when compared to international scientific reference standards, the *mean* impact on chemical research is for the most part far above average not only for accepted Communications but also for rejected (but published elsewhere) Communications. However, if on the basis of the 95% confidence intervals we look at the areas in which, with high certainty, the "true" mean values of the citations for the Communications lie, and compare these areas with the reference standards, it becomes clear that average and below-average scientific impact is to be expected significantly less frequently for accepted Communications than for rejected Communications.

Discussion and Conclusions

In this study on manuscript review for the selection of Communications for publication in the journal AC-IE, we analyzed the process with regard to whether the journal is achieving its goal of selecting the best manuscripts. Assessing the quality of selection decisions requires a generally accepted criterion for the impact on scientific research. Citation counts are considered to be an indicator of research impact, since they measure the "*actual* influence on surrounding research activities at a given time" (Martin & Irvine, 1983, p. 70).

The results of the citation analyses in this study show that the review process at AC-IE indeed achieves the goal of selecting for publication the best Communications with the highest impact on chemical research. Independent of the publication year, the literature database in which the citation search was performed (SCI or CA), and the citation window, accepted Communications were on average (clearly) more frequently cited than rejected (but published elsewhere) Communications (the differences are between 5.35 and 11.2 citations). According to the results of our regression models with CA and SCI citations as dependent variables, for accepted Communications approximately 50% (models A and B) or 40% (models C and D) more citations are to be expected than for rejected (but published elsewhere) Communications, holding all other variables in the models constant. Moreover, a comparison of average citation counts of accepted and rejected Communications with international scientific reference standards reveals that for accepted Communications "true" mean citation counts below the baseline are significantly less frequently expected than for rejected Communications.

In a *Nature* "Web Focus" article on journal peer review, Jennings (2006) states, "the most important question is how accurately the peer-review system predicts the longer-term judgments of the scientific community. One way to address this would be through citation data." The results of this study confirm that the editorial decisions of AC-IE are able to, as Jennings puts it, predict the longer-term judgments of the chemical scientific community. Similar results regarding the (on average) accurate prediction of longer-term judgments by the journal peer-review system have been reported for selection decisions at four different journals. Based on mean citation rates for accepted manuscripts and rejected manuscripts that were nevertheless published elsewhere, the decisions made by the editors of the *Journal of Clinical Investigation* (Wilson, 1978), the *British Medical Journal* (Lock, 1985), *Angewandte Chemie* (Daniel, 1993/2004), and *Cardiovascular Research* (Opthof, Furstner, van Geer, & Coronel, 2000) reflect a high degree of predictive validity. Daniel (1993/2004), for example, showed for Communications submitted to *Angewandte Chemie* in the mid-1980s that accepted Communications were on average cited twice as frequently as manuscripts that had been rejected on the basis of reviewers' recommendations but were later published elsewhere.

All in all, an interesting methodology has been developed and successfully applied in our study to answer the question, does the journal peer-review system fulfill its declared objective to select the best scientific work? All journals that use peer review should find an answer to this question. In the case of AC-IE, the verification of the higher scientific impact of accepted manuscripts in contrast to rejected (and published elsewhere) ones provides evidence for the proper operation of the journal peer-review system (see Jennings, 2006).

Acknowledgments

The entire research study, which is also investigating quality assurance of open-access journals, is supported by a grant from the Max Planck Society. The authors would like to thank Dr. Christophe Weymuth (formerly at the Organic Chemistry Institute of the University of Zurich and now at BIOSYNTH AG, Switzerland) for investigation of the manuscripts rejected by *Angewandte Chemie International Edition* and published elsewhere. We also thank Dr. Werner Marx and Dr. Hermann Schier of the Central Information Service for the institutes of the Chemical Physical Technical (CPT) Section of the Max Planck Society (located at the Max Planck Institute for Solid State Research in Stuttgart, Germany) for conducting the citation search for citations of the accepted and rejected (but published elsewhere) Communications in the literature databases *Science Citation Index* and *Chemical Abstracts*, and for calculating the section-specific reference standards (*Chemical Abstracts*).

We thank Dr. Peter Gölitz, Editor-in-Chief of *Angewandte Chemie*, the Editorial Board of *Angewandte Chemie*, and the German Chemical Society (GDCh, Frankfurt am Main, Germany) for permission to conduct the evaluation of the peer-review process of the journal, and to the members of the editorial office for their generous support during the study. As the data that are analyzed in this study are strictly confidential, they are not publicly available for replication studies without permission of the journal.

The authors wish to express their gratitude to three anonymous referees for their helpful comments.

References

Bornmann, L., & Daniel, H.-D. (2006). Selecting scientific excellence through committee peer review: A citation analysis of publications previously published to approval or rejection of postdoctoral research fellowship applicants. Scientometrics, 68, 427–440.

Bornmann, L., & Daniel, H.-D. (2007a). Convergent validation of peer review decisions using the *h* index: Extent of and reasons for type I and type II errors. Journal of Informetrics, 1, 204–213.

Bornmann, L., & Daniel, H.-D. (2007b). Functional use of frequently and infrequently cited articles in citing publications. A content analysis of citations to articles with low and high citation counts. In D. Torres-Salinas & H.F. Moed (Eds.), Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics (Vol. 1, pp. 149–153). Madrid, Spain: Spanish Research Council (CSIC).

Bornmann, L., & Daniel, H.-D. (2007c). Multiple publication on a single research study: Does it pay? The influence of number of research articles on total citation counts in biomedicine. Journal of the American Society for Information Science and Technology, 58, 1100–1107.

Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. Journal of Documentation, 64, 45–80.

Bornmann, L., & Daniel, H.-D. (in press). The effectiveness of the peer review process: Interreferee agreement and predictive validity of manuscript refereeing at *Angewandte Chemie*. Angewandte Chemie International Edition.

Bornmann, L., Leydesdorff, L., & Marx, W. (2007). Citation environment of *Angewandte Chemie*. CHIMIA, 61, 104–109.

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Use of citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. Ethics in Science and Environmental Politics, 8, 7.

Braun, T. (2004). Keeping the gates of science journals: Gatekeeping indicators of national performance in the sciences. In H.F. Moed, W. Glänzel, & U. Schmoch (Eds.), Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems (pp. 95–114). Dordrecht, The Netherlands: Kluwer Academic.

Callaham, M., Wears, R.L., & Weber, E. (2002). Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. Journal of the American Medical Association, 287, 2847–2850.

Cole, J.R. (2000). A short history of the use of citations as a measure of the impact of scientific and scholarly work. In B. Cronin & H.B. Atkins (Eds.), The web of knowledge. A Festschrift in honor of Eugene Garfield (pp. 281–300). Medford, NJ: Information Today.

Cozzens, S.E. (1985). Comparing the sciences: Citation context analysis of papers from neuropharmacology and the sociology of science. Social Studies of Science, 15, 127–153.

Craig, I.D., Plume, A.M., McVeigh, M.E., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact? A critical review of the literature. Journal of Informetrics, 1, 239–248.

Cronin, B., & McKenzie, G. (1992). Documentation note the trajectory of rejection. Journal of Documentation, 48, 310–317.

Daniel, H.-D. (1993/2004). Guardians of science. Fairness and reliability of peer review. Weinheim, Germany: Wiley-VCH.

Daniel, H.-D. (2005). Publications as a measure of scientific advancement and of scientists' productivity. Learned Publishing, 18, 143–148.

Dewitt, T.W., Nicholson, R.S., & Wilson, M.K. (1980). *Science Citation Index* and chemistry. Scientometrics, 2, 265–275.

Ehses, I. (2004). By scientists, for scientists: The Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—and how it functions. B.I.F. Futura, 19, 170–177.

Garfield, E. (2002). Highly cited authors. Scientist, 16(7), 10.

Glänzel, W., Debackere, K., Thijs, B., & Schubert, A. (2006). A concise review on the role of author self-citations in information science, bibliometrics and science policy. Scientometrics, 67, 263–277.

Glänzel, W., & Schubert, A. (1993). A characterization of scientometric distributions based on harmonic means. Scientometrics, 26, 81–96.

Gölitz, P. (2005). Who is going to read all this? Angewandte Chemie International Edition, 44, 5538–5541.

Greenwood, D.C. (2007). Reliability of journal impact factor rankings. BMC Medical Research Methodology, 7.

Harnad, S. (2007). Open access scientometrics and the UK Research Assessment Exercise. In D. Torres-Salinas & H.F. Moed (Eds.), Proceedings of the 11th Conference of the International Society for Scientometrics and Informetrics (Vol. 1, pp. 27–33). Madrid, Spain: Spanish Research Council (CSIC).

Herbertz, H. (1995). Does it pay to cooperate? A bibliometric case-study in molecular biology. Scientometrics, 33, 117–122.

Hosmer, D.W., & Lemeshow, S. (2000). Applied logistic regression (2nd ed.). Chichester, UK: Wiley.

Institute for Scientific Information (2006). Journal Citation Reports: Science Edition. Database of the Institute for Scientific Information. Philadelphia, PA: Institute for Scientific Information.

ISIHighlyCited.com (2008). About ISIHighlyCited.com (para. 1 and 2). www.ISIHighlyCited.com

Jennings, C.G. (2006). Quality and value: The true purpose of peer review. What you can't measure, you can't manage: The need for quantitative indicators in peer review. Retrieved July 6, 2006, from http://www.nature.com/nature/peerreview/debate/nature05032.html

Klamer, A., & van Dalen, H.P. (2002). Attention and the art of scientific publishing. Journal of Economic Methodology, 9(3), 289–315.

Kline, P. (1998). The new psychometrics: Science, psychology, and measurement. London, UK: Routledge.

Kohler, U., & Kreuter, F. (2005). Data analysis using Stata. College Station, TX: Stata.

Kostoff, R.N. (1997). The principles and practices of peer review. Science and Engineering Ethics, 3, 19–34.

Leimu, R., & Koricheva, J. (2005). What determines the citation frequency of ecological papers? Trends in Ecology & Evolution, 20, 28–32.

Lewison, G., & Dawson, G. (1998). The effect of funding on the outputs of biomedical research. Scientometrics, 41, 17–27.

Lock, S. (1985). A difficult balance: Editorial peer review in medicine. Philadelphia, PA: ISI Press.

Lokker, C., McKibbon, K.A., McKinlay, R.J., Wilczynski, N.L., & Haynes, R.B. (2008). Prediction of citation counts for clinical articles at two years using data available within three weeks of publication: Retrospective cohort study. British Medical Journal, 336(7645), 655–657.

Long, J.S., & Freese, J. (2006). Regression models for categorical dependent variables using Stata (2nd ed.). College Station, TX: Stata.

Marsh, H.W., Jayasinghe, U.W., & Bond, N.W. (2008). Improving the peer-review process for grant applications—reliability, validity, bias, and generalizability. American Psychologist, 63(3), 160–168.

Martin, B.R., & Irvine, J. (1983). Assessing basic research: Some partial indicators of scientific progress in radio astronomy. Research Policy, 12(2), 61–90.

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

Neuhaus, C., & Daniel, H.-D. (in press). A new reference standard for citation analysis in chemistry and related fields based on the sections of *Chemical Abstracts*. Scientometrics.

Opthof, T., Furstner, F., van Geer, M., & Coronel, R. (2000). Regrets or no regrets? No regrets! The fate of rejected manuscripts. Cardiovascular Research, 45(1), 255–258.

Rabe-Hesketh, S., & Everitt, B. (2004). A handbook of statistical analyses using Stata. Boca Raton, FL: Chapman & Hall/CRC.

DOI: 10.1002/asi

Research Evaluation and Policy Project. (2005). Quantitative indicators for research assessment: A literature review (REPP Discussion Paper 05/1). Canberra, Australia: The Australian National University, Research School of Social Sciences.

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

Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. British Medical Journal, 314(7079), 498–502.

Sense about Science. (2004). Peer review and the acceptance of new scientific ideas. London, UK: Sense about Science.

StataCorp. (2007). Stata statistical software: Release 10. College Station, TX: Stata.

Stine, R. (1990). An introduction to bootstrap methods: Examples and ideas. In J. Fox & J.S. Long (Eds.), Modern methods of data analysis (pp. 325–373). Newbury Park, CA: Sage.

Thomson Reuters. (2008). Journal Citation Reports. Retrieved June 17, 2008, from http://scientific.thomson.com/products/jcr/.

van Leeuwen, T.N. (2007). Modelling of bibliometric approaches and importance of output verification in research performance assessment. Research Evaluation, 16, 93–105.

van Raan, A.F.J. (1999). Advanced bibliometric methods for the evaluation of universities. Scientometrics, 45, 417–423.

van Raan, A.F.J. (2004). Measuring science. Capita selecta of current main issues. In H.F. Moed, W. Glänzel, & U. Schmoch (Eds.), Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems (pp. 19–50). Dordrecht, The Netherlands: Kluwer Academic.

Vinkler, P. (1986). Evaluation of some methods for the relative assessment of scientific publications. Scientometrics, 10, 157–177.

Vinkler, P. (1997). Relations of relative scientometric impact indicators. The relative publication strategy index. Scientometrics, 40, 163–169.

Weller, A.C. (2002). Editorial peer review: Its strengths and weaknesses. Medford, NJ: Information Today.

Whitley, K.M. (2002). Analysis of SciFinder scholar and Web of Science citation searches. Journal of the American Society for Information Science and Technology, 53, 1210–1215.

Wilson, J.D. (1978). Peer review and publication. Journal of Clinical Investigation, 61, 1697–1701.

Wuchty, S., Jones, B.F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036–1039.

DOI: 10.1002/asi