DE GRUYTER
SAUR

Stefanie Haustein

MULTIDIMENSIONAL JOURNAL EVALUATION

ANALYZING SCIENTIFIC PERIODICALS BEYOND THE IMPACT FACTOR

KNOWLEDGE & INFORMATION

DE
G

Stefanie Haustein

Multidimensional Journal Evaluation

Analyzing Scientific Periodicals beyond the Impact Factor

De Gruyter Saur

D 61

ISBN 978-3-11-025494-5
e-ISBN 978-3-11-025555-3
ISSN 1868-842X

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

Printing: Hubert & Co. GmbH & Co. KG, Göttingen
∞ Printed on acid-free paper

Printed in Germany

www.degruyter.com

Knowledge & Information

Knowledge & Information

Studies in Information Science

Knowledge and Information (K&I) is a peer-reviewed information science book series. The scope of information science comprehends representing, providing, searching and finding of relevant knowledge including all activities of information professionals (e.g., indexing and abstracting) and users (e.g., their information behavior). An important research area is information retrieval, the science of search engines and their users. Topics of knowledge representation include metadata as well as methods and tools of knowledge organization systems (folksonomies, nomenclatures, classification systems, thesauri, and ontologies). Informetrics is empirical information science and consists, among others, of the domain-specific metrics (e.g., scientometrics, webometrics), user and usage research, and evaluation of information systems. The sharing and the distribution of internal and external information in organizations are research topics of knowledge management. The information market can be defined as the exchange of digital information on networks, especially the World Wide Web. Further important research areas of information science are information ethics, information law, and information sociology.

De Gruyter Saur

Contents

Chapter 1
Introduction

This book focuses on scholarly journals and deals with their role as the primary means of formal communication in the sciences. Their standing in the scientific community is evaluated from the perspective of quantitative informetric indicators. Two main research questions are investigated:

- Are there different dimensions of journal evaluation? If so, what indicators can be applied to represent each dimension of journal evaluation? How do metrics differ and what data can be used to adequately indicate journal impact?
- Various stakeholders of journal evaluation, i.e. scholarly authors, readers, librarians, publishers and editors, have different interests in using methods of journal evaluation. What indicators of what dimensions will each of the stakeholders find most valuable and apply in practice to journal selection?

In the form of a case study, a set of 45 physics journals will be used to demonstrate the various journal metrics and informetric methods in terms of practicability, validity and informative value.

Formal communication in the sciences depends primarily on the publication of research results and comprehensive reviews in the form of journal articles. Since the first scholarly journal was founded in the 17th century, the scientific periodical has become the most important medium in science, technology and medicine (STM). Through its fast and wide dissemination the peer-reviewed journal is the preferred formal publication outlet of researchers in these fields. In the social sciences and humanities its influence is growing. With constant growth in the number of scientists, research institutions and publications, the number of scholarly journals and the number of articles in them increases, too. The academic world has to deal with the problem of massive information overload. The current number of scientific peer-reviewed periodicals is estimated to be approximately 24,000 titles (Tenopir & King, 2009). Researchers are confronted with the impossible task of keeping up with the amount of information available (Moens, 2000; Keller, 2005).

> [I]nformation overload occurs when information received becomes a hindrance rather than a help when the information is potentially useful. (Bawden, Holtham, & Courtney, 1999, p. 249)

Bawden and Robinson (2009) quote a report based on experience from the area of biomedical research in 1986:

> Many medical practitioners have abandoned 'keeping up with the literature'. They are paralysed by the sheer enormity of the task: more than 20,000 journals in biomedicine are published each year and a consultant in a single medical subspeciality may need to read fifteen to twenty publications a month to keep up to date. (Bawden & Robinson, 2009, p. 183)

Price (1963) claims that "scientists have always felt themselves to be awash in a sea of scientific literature" (Price, 1963, p. 15). He was also the first to provide support for the subjective perception of the information flood in science by objective statistics on the number of scholarly journals. With the emergence of electronic publishing the number of documents available has further increased. With an ever growing output the challenge is to manage this information overload and select the most suitable sources. It is not only the reader who has to choose. Since they are evaluated on the basis of their publications, researchers need to develop a publication strategy and select the best periodicals in their role as authors as well. Scientometric indicators can help with these selection processes.

Informetrics and the sub-areas of scientometrics and bibliometrics represent an approach forcoping with the growing flood of scholarly literature by means of objective statistical methods (Tague-Sutcliffe, 1992; Stock & Weber, 2006; Bar-Ilan, 2008). Bibliometric methods quantitatively evaluate the structure of and processes involved in formal scholarly communication, in which the scientific journal plays a paramount role. Scientometric analysis of periodicals arose out of the need to identify significant sources of information. Initially, journal evaluation was applied with a local focus by librarians for collection management purposes. With the development of the Science Citation Index (SCI) and the advent of systematical journal-based research evaluation, journal rankings have become important and are applied by authors, readers, editors, publishers and research managers alike.

> The evaluation of scholarly journals is important for selection and cancellation decisions by librarians, the evaluation of faculty and librarians for promotion and tenure as well as annual performance reviews, manuscript submission decision by authors, monitoring of their journals by editors and publishers, and familiarizing new doctoral students or outsiders (such as members of a university-wide promotion and tenure committee evaluating faculty from other departments) with a field's journals. (Nisonger & Davis, 2005, p. 341)

> Journal citation measures are designed to assess significance and performance of individual journals, their role and position in the international formal communication network, their quality or prestige as perceived by scholars. Scientific journals may differ with respect to their importance of their position in the journal communication system, their status or prestige. (Glänzel & Moed, 2002, p. 171)

The most widely used indicator for selecting the most influential journals in a scientific field is the impact factor listed in the Journal Citation Reports (Stock, 2001). Developed by Eugene Garfield (1955) as a criterion to select a discipline's core journals for his citation indices, the impact factor has become a popular measure to indicate a journal's standing in the scientific community (Garfield, 1972). Librarians use the impact factor to compile literature within their budget, authors make use of it to choose the best suited journal for publication, readers use it to identify central sources of information, and editors and publishers apply it to analyze the market and observe competitors. The impact factor has become the most popular bibliometric indicator used inside and especially outside the scientometric community (Glänzel & Moed, 2002). As an average citation rate it is, however, not able to reflect all the aspects that contribute to a scholarly journal, not to mention its methodological shortcomings. As the impact factor became

popular, so, too, did its misuse. It is often applied directly to the articles a journal contains or even to the contributing authors. Publications in high-impact journals are rewarded financially, and with the help of cumulative impact factors, research grants are provided or careers decided upon (Jennings, 1998; Al-Awqati, 2007).

While alternatives exist and new and improved methods are being created by the bibliometric community, in the evaluation of journals by its users the influence, quality and prestige of a serial are mainly based on the impact factor. The journal's value is limited to a single quantitative indicator which divides the number of received citations by the number of published articles. Further methods and possibilities of journal evaluation are more or less disregarded by users outside the bibliometric community. Journal evaluation is unidimensional and therefore not able to reflect the actual impact of a scholarly periodical (Rousseau, 2002; Coleman, 2007).

This book argues that only a method that incorporates all facets of journal evaluation can adequately represent the journal's standing in the scientific community. Five conceptual dimensions are defined which contribute to the value of scholarly journals and are therefore to be included in a comprehensive and adequate approach to journal evaluation. The goal of this study is to provide the users with a variety of indicators to choose the most suitable method for their individual purpose regarding journal selection. Since the purposes of journal evaluation are individual (Björk & Holmström, 2006), it is not the aim to provide a cure-all or all-in-one device but to systematically review various metrics of journal evaluation including a transparent description of possibilities and shortcomings. The work focuses on the users of these methods, i.e. researchers in their role as authors and readers, librarians, publishers and editors, rather than the experienced bibliometrician and tries to guide them through the "maze of indicators" (Wouters, 1999, p. 128). This book tries to solve the lack of transparency of bibliometric methods and journal metrics, which causes users to restrict themselves to the well-known and readily available but limited and flawed impact factor. Whether or not alternative or complementary indicators are used depends above all on these indicators being known and understood by their potential users.

> What will ultimately determine which of this new battery of measurements succeed and which fail, either individually or as composite measures, is likely to be how strongly they resonate with the communities they serve. The best ideas do not always make the best products, but instead simplicity and transparency can be the difference between success and obscurity. (Craig & Ferguson, 2009, p. 188 f.)

This book tries to provide transparency and lets the user choose from a toolbox of indicators and combine the most suitable for his particular needs, i.e. an author to select the most appropriate publication venue, the reader to select journals to fulfill his information needs, a librarian to compile a collection, and an editor and/or publisher to monitor the performance of his own journals and that of competitors.

In the following, the scholarly journal is defined and its emergence as the most important form of formal scholarly communication briefly described. Afterwards section 1.2 provides an overview of the usage and users of evaluation methods. Section 1.3 introduces the five dimensions of multidimensional journal evaluation, which are discussed and analyzed in detail in the respective chapters. Section 1.4 describes the 45 physics journals which throughout the study function as a test set to apply and compare various indicators.

1.1 The Scholarly Journal

Since the founding of general science journals over 300 years ago, the publishing landscape has grown, specialized and become more and more heterogeneous and unmanagable for a researcher to cover. After an attempt to define what constitutes a scholarly journal (section 1.1.1), a brief historical overview of the emergence and development (section 1.1.2) is given.

1.1.1 Definitions

As this work is based on the evaluation of scholarly journals, a definition of it is due. Although a vast amount of literature exists, an exact, formal definition of the scholarly journal is, however, lacking. Most works assume that it is understood. In addition, the terms 'journal', 'periodical' and 'serial' are used synonymously (Buchet, 2004; Keller, 2005; Mierzejewska, 2008). Page, Campbell, and Meadows (1997) define 'serial' and 'periodical' as follows:

> **Serial**: a publication issued in successive parts, bearing numerical or chronological designations and intended to be continued indefinitely.
>
> **Periodical**: a publication appearing at stated intervals, each number of which contains a variety of original articles by different authors. (Page et al., 1997, p. 1)

In this study, the terms 'journal', 'periodical' and 'serial' are used synonymously and all refer to scholarly or learned journals which, on a regular basis and to a great extent, publish original research articles and/or reviews and apply some form of peer review. Academic journals can be clearly distinguished from popular magazines and newspapers and so-called trade journals which, on the one hand, publish current, non-technical and unscholarly content often written by in-house writers and commissioned journalists and, on the other hand, inform members of one particular industry sector or branch of trade.[1] As electronic journals or e-journals are merely a new form of appearance of the same medium established 300 years ago, no distinction is made between print and electronic journals. Even though future visions of the electronic journal predicted significant changes of scholarly communication tapping the full potential of the digital age, the concept of the academic journal has so far hardly been altered by its electronic form (Keller, 2005). Although the electronic, i.e. PDF, format improves access, searchability and intra- and inter-document navigation, the scholarly article as such, and hence the academic journal, has not changed.

In the area of STM, academic journals play the major role in formal scholarly communication. The extent to which scholarly output is published in journals, varies between disciplines. As table 1.1 shows, over 80% of all scientific output in the biological, chemical, physical and medical research fields is published in journals, while in the social sciences and arts and humanities, where book publications play a more significant role, it is less than half as much. As bibliographical databases such as the Web of Science (WoS) and Scopus are limited to periodical literature and not able to cover the entire scholarly output, the databases' ability to reflect formal scholarly communication depends on the discipline-specific publication behavior, on the one hand, and database coverage, on the other. Due to these field differences, this study

1 http://www.unity.edu/library/scholarly_journal.htm

focuses on physics journals (see section 1.4), but methods and indicators can be equally applied to periodicals in any other STM research field. When evaluating social sciences and arts and humanities journals, one should be aware of capturing only a fraction of the scholarly output. In order to adequately reflect formal scholarly communication in non-STM research, books have to be considered as well.

Table 1.1 Percentage ($P_{\%}$) of references (of articles and reviews published in 2002) per discipline published in journals and covered by WoS (WoS). Source: Craig and Ferguson (2009, p. 163).

Discipline	$P_{\%}$ in journals	$P_{\%}$ in WoS	$P_{\%}$ WoS coverage
Molecular biology and biochemistry	96	97	92
Biological sciences related to humans	95	95	90
Chemistry	90	93	84
Clinical medicine	93	90	84
Physics and astronomy	89	94	83
WoS average	*84*	*90*	*75*
Applied physics and chemistry	83	89	73
Biological sciences (animals and plants)	81	84	69
Psychology and psychiatry	75	88	66
Geosciences	77	81	62
Other social sciences (medicine and health)	75	80	60
Mathematics	71	74	53
Economics	59	80	47
Engineering	60	77	46
Social sciences	41	72	29
Humanities and arts	34	50	17

As major communication channels in STM, academic journals fulfill the four basic functions of registration, certification, dissemination and archiving (Mierzejewska, 2008). Through publication researchers are able able to publish their findings quickly and claim priority. Peer review and editorial processes guarantee quality control and validity of published contents. The journal provides a platform of wide dissemination in the scholarly community and at the same time makes findings permanent through archiving (Mabe & Amin, 2002; Meier, 2002). While registration, certification, awareness and archiving represent the primary functions of an academic periodical, secondary functions can be identified as well. They mainly refer to social aspects of scientific communities, i.e. forming and developing communities by providing a communication platform, identifying and empowering influential authors through editorial board membership and by providing the framework for scientific evaluations (Mierzejewska, 2008).

1.1.2 Emergence and Development

The first scholarly journals emerged in the 17th century in order to make science more efficient. Journal publishing was intended to serve the purpose of avoiding duplication

and advance scientific knowledge by building on results of colleagues (Mierzejewska, 2008). Published in France in January of 1665, Le Journal des Sçavans is said to be the first scholarly journal since it published articles and news items on many different topics for scholars. Only shortly after, the first issue of Philosophical Transactions of the Royal Society of London appeared, which had a greater resemblance to the modern scholarly journal since it consisted of 407 instead of 20 pages (Keller, 2005; Tenopir & King, 2009). Philosophical Transactions was launched by the Secretary of the Royal Society, Henry Oldenburg

> to inform the Fellows of the Society and other interested readers of the latest scientific discoveries. As such, Philosophical Transactions established the important principles of scientific priority and peer review, which have become the central foundations of scientific journals ever since.[2]

The emergence of scholarly journals in the 17th century fundamentally changed the entire process of scholarly communication, which up to that point had been carried out through society meetings, books and letters (Price, 1963; Zuckerman & Merton, 1971). Although the latter represented a timely method of publishing results and claim priority for discoveries, scientific letters were still a personal form of communication and thus limited in terms of distribution (Keller, 2005; Tenopir & King, 2009). Price (1963) emphasizes the initial purpose of the newly founded journals as

> monitoring and digesting of the learned publications and letters that now were too much for one man to cope with in his daily reading and correspondence. (Price, 1963, p. 15)

Scientific communication advanced from personal correspondence and became structured and distributed on a regular basis. This made it possible to keep a record of and archive knowledge systematically. In addition, publication in a journal allowed researchers to claim their discoveries (Keller, 2005).

Since the founding years of the scholarly journal, the number of titles has increased, although Price's famous estimations of exponential growth which predicted one million titles in 2000 did not occur (compare chapter 2). Annual growth rates increase gradually rather than exponentially (Mierzejewska, 2008). Mabe (2003) calculated an almost constant annual growth rate of 3.46% of active refereed scholarly journal titles from 1665 to 2001 listed in Ulrich's Periodicals Directory , and Tenopir and King (2009) conclude from listings in the same database that in 2008 61,620 scholarly journals existed (i.e. were still actively publishing), of which 23,973 were refereed. Placing the number of scholarly, peer-reviewed journals at 24,000 seems a conservative but fair estimate of the number of journals available, which is supported by results of other studies (Mabe, 2003; Morris, 2007; Craig & Ferguson, 2009; Tenopir & King, 2009).

Today the landscape of scholarly journals is characterized by its heterogeneity. Further growth and specialization of science demands specialization of publication venues, which contributes to the foundation of new journals or the splitting up of existing ones (Meier, 2002; Mierzejewska, 2008). The increasing specialization and further development of science has led to a flood of information which is difficult for a single researcher to access or manage (Keller, 2005). As it is not possible for a

2 http://rstl.royalsocietypublishing.org/

researcher to read all the relevant literature, it is now more than ever crucial to select the most important resources so that relevant content is not missed. Journal evaluation can help to identify the most suitable journals.

1.2 Application and Developments of Journal Evaluation

Scientometric analysis of periodicals arose out of the need to identify significant sources of information. Bibliometric indicators were first and foremost developed to cope with the flood of information created by an increasing specialization and differentiation of science and the growth of research output. With the amount of literature available, the scientific landscape has become complex and the chances are that relevant content is missed (Keller, 2005). Quantitative methods provide focus (Craig & Ferguson, 2009) and are applied to help researchers in their role as readers and authors to select the most suitable journals satisfying their information needs, on the one hand, and communicating their results, on the other.

At the beginning, journal evaluation was applied locally by librarians for collection management purposes. Gross and Gross (1927) are considered to be the first describing a reference-based journal analysis for managing library holdings by objective, quantitative methods (Archambault & Larivière, 2009). Bradford (1934) further influenced librarians and collection management through his famous law of scattering stating that the majority of documents on a given subject are concentratively published in a few core journals.

With the development of the SCI by the Institute for Scientific Information (ISI), journal evaluation transcended local library collection management and was applied on a global scale. With the impact factor, the first widely applied indicator for journal evaluation was developed, although, initially it was constructed as a means to identify the most frequently cited journals per discipline for the SCI rather than a journal metric on its own. As a multidisciplinary citation index, the SCI was in the first instance not developed for evaluational purposes but for the retrieval of scientific information (Garfield, 1955, 1998).

The SCI has, however, fostered the culture of research evaluation. The impact factor is no longer a simple metric to identify candidates for database coverage, but has become a synonym for journal prestige powerful enough to influence scholarly communication and even researchers' careers. In these times of journal-based research evaluation, the ranking of periodicals is central and the impact factor has become a cure-all indicator used and misused by authors, readers, editors, publishers and research policy makers alike (Adam, 2002; Rogers, 2002; The PLoS Medicine editors, 2006).

As the limitations of the impact factor and its inability to fulfill various information needs and fully represent a journal's standing in the scholarly community became obvious, many alternative metrics were proposed from within and outside the bibliometric community. Figure 1.1 shows the growth of publications on journal evaluation[3] and their impact. During the 30 years between 1980 and 2009 almost 3,500 documents on

3 3,497 documents retrieved in January 2011 from SCI, SSCI, A&HCI and CPCI-S published between 1980 and 2009, matching the following query: ti=("impact factor*" or (journal$ and (evaluat* or analy* or measur* or indicat* or cited or citing or citation* or rank* or scientometr* or bibliometric* or informetric*))).

journal evaluation were published, 10% of which were published in 2009. While new or newly named indicators emerge regularly within the bibliometric community, in applied journal analyses evaluation methods are mostly limited to the impact factor, which is largely due to its simplicity and availability (Seglen, 1997; Glänzel & Moed, 2002; Moed, 2002; Favaloro, 2008).

> Apart from the impact factor, the use of other journal citation measures is rather occasional. [...] All other attempts to improve, substitute or supplement the impact factor have encountered serious obstacles to wider application. On the one hand, the achieved 'accuracy' is often at the expense of simple interpretation. On the other hand, several alternative journal citation measures could not always be rendered accessible to a broader audience, or at least not regularly be updated like in the case of the IF through the annual updates of the JCR. In lack of regular updates, these indicators could not be readily reproduced by other research centres. (Glänzel & Moed, 2002, p. 177 f.)

In order to be used, alternative metrics have to be understood. This work aims to provide a systematic overview of possibilities of multidimensional journal evaluation with a focus on the applicability and limitations of existing indicators. In addition, new data sources such as bookmarks of and tags assigned to journal articles by users of specific web 2.0 platforms are introduced as alternatives to represent the readers' perspective (Haustein & Siebenlist, 2011). It will be shown that one journal metric, such as the impact factor, is not able to fulfill the various requirements for journal evaluation of the different user groups.

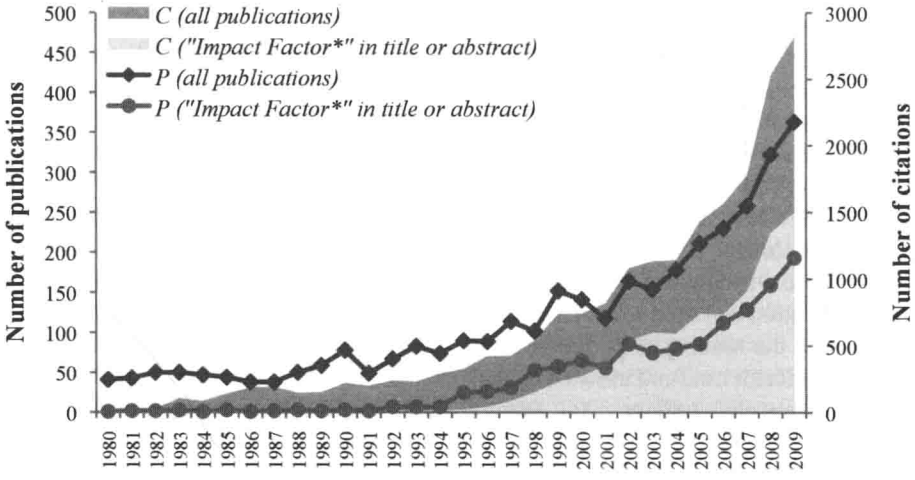

Figure 1.1 Number of publications (*P*) and citations (*C*) of journal evaluation literature 1980 to 2009.

In the journal publication landscape, five groups of participants can be identified: researchers (i.e. authors), gatekeepers (i.e. editors), publishers, libraries and users (i.e. readers). These groups "may significantly influence the functioning and success of

academic journal publishing" (Mierzejewska, 2008, p. 7). These participants represent actual and potential users of journal evaluation. In the process of evaluating scholarly journals, four stakeholders can be identified, i.e. academic authors selecting a publication venue, readers choosing a suitable source of information, librarians managing a collection, and editors and publishers evaluating journal performance and observing competitors (Garfield, 1972; Todorov & Glänzel, 1988; Nisonger, 1999; Glänzel & Moed, 2002; Rousseau, 2002). Research evaluators at universities, government offices and funding agencies can be identified as a fifth stakeholder. However, their purpose in evaluating journals is secondary in so far that they rank journals as a means of evaluating the researchers and institutions publishing in them. Thus, policy makers are not treated as a distinct user group in this study.

Each of the four user groups has different requirements for identifying and ranking scholarly journals (Moed, Van Leeuwen, & Reedijk, 1998). Depending on these needs, one indicator might be more suitable for one stakeholder than another. The same set of periodicals might even be ranked differently from the readers' and authors' perspectives.

> 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. (Rousseau, 2002, p. 418)

Journal evaluation depends on the individual purpose and requirements of the user. The four stakeholders are introduced in the following together with the differences and similarities of their positions in journal evaluation.

Readers

Readers of scholarly literature can be divided into two groups, namely publishing and pure readers (Rowlands & Nicholas, 2007). The former can be considered as active researchers who use the scholarly journal as a means to communicate their research by publishing results which then influence and improve further research and technological development. These publishing readers need to select a set of journals to be informed about the research front. In journal selection they give priority to the topicality, scholarliness and high quality of journal content guaranteed through peer review. The group of pure readers includes those readers who do not actively participate in research but read scholarly literature to apply in their everyday worklife and teaching. Students make up a large part of pure readers. Pure readers primarily select review journals to obtain an overview and and be generally informed (Rousseau, 2002).

In general, researchers in their role as readers wish to read as little as possible to fulfill their information needs. Due to the evaluation culture, in their role as authors the same researchers, however, aim to publish as much as possible and to be read by as many people as possible. This schizophrenic behavior (Meier, 2002), described by Mabe and Amin (2002) as 'Dr Jekyll and Dr Hyde', is also reflected in journal selection.