

The Usage Factor Project

Martin Richardson
Managing Director, Academic Books and Journals, Oxford University Press

Abstract

In 2002, COUNTER (Counting Online Usage of Networked Electronic Resources) was set up to provide a consistent and credible method for reporting journal usage statistics. However, while COUNTER statistics enable reliable comparison of the amount of usage between journals, they do not provide a meaningful usage-based measure of relative quality or value since a large journal will, on average, generate significantly more downloads than a small one. The Usage Factor Project was therefore established in 2007 by the United Kingdom Serials Group (UKSG) to examine whether it would be helpful to develop a metric based on usage data to complement citation-based metrics such as the Institute for Scientific Information’s Impact Factor. The Usage Factor Project was divided into two stages, with the first stage consisting of three phases. This presentation summarizes progress to date and the research that is currently in progress.



Martin Richardson

Managing Director, Academic Books and Journals, Oxford University Press

Martin Richardson has held a variety of positions at Oxford University Press over the past 20 years. He is currently Managing Director of the Academic Books and Journals Divisions, where he is responsible for over 3,000 print and online publications across a broad range of subjects.

As a department of the University of Oxford, OUP is keen to experiment with new ways of increasing dissemination of research and educational material by employing new technologies and business models. Under Martin’s leadership OUP began online journal publishing in the early 1990’s and in 2000 launched the first of many online reference publications - the Oxford English Dictionary. In 2004 Oxford Scholarship Online was launched, providing a platform for the publication of OUP’s extensive monograph programme. Martin has represented OUP on a number of publishing industry organizations, including ALPSP, PLS and CrossRef. He is a regular speaker at conferences and has also published papers on various aspects of online publishing.

Introduction

Most researchers are familiar with the use of the Impact Factor which uses citations as a basis for measuring the quality of research in individual journals. This paper describes our investigation of a usage-based metric to measure research in a way that will complement the Impact Factor.

The standard way of measuring usage in research journals was developed by an organization called COUNTER (Counting Online Usage of Networked Electronic Resources), which was founded by Richard Gedye of Oxford University Press. The purpose of COUNTER was to make usage statistics consistent, credible and compatible. To ensure consistency, COUNTER developed a standard format for measuring usage that has now been adopted by over 100 publishers and database hosts around the world. To achieve credibility, in 2007 the organization established a formal process for publishers to submit their usage data for auditing. Finally, in relation to compatibility, by establishing a standard, COUNTER made it easier to compare usage of one journal to another. However, there still are a number of problems with this metric.

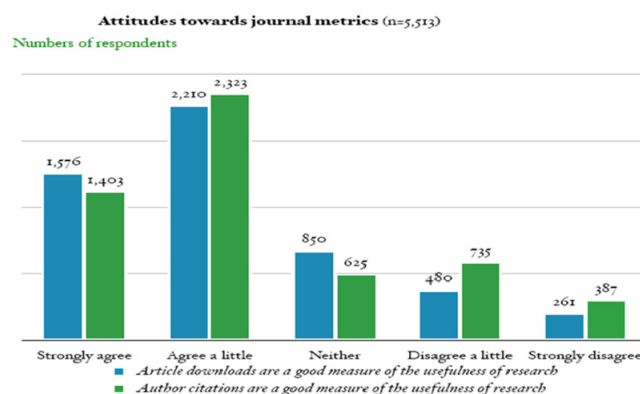
While COUNTER statistics provide a reliable comparison of the amount of usage between journals, they do not really provide a meaningful usage-based measure of relative quality or value. This is because a very large journal will generate more usage than a very small one. For example, a journal publishing 2000 articles a year is naturally likely to get a much higher number of downloads than one publishing only 50 articles per year.

The Institute for Scientific Information’s Impact Factor deals with this problem by taking the number of articles into account when performing a calculation. We considered whether a similar approach could be taken to usage, for example, by developing a usage factor that took into account the size of the journal. A very simple calculation such as the following could be used:

Usage Factor =

$$\frac{\text{Total usage over period 'x' of articles published during period 'y'}}{\text{Total articles published during period 'y'}}$$

But what is the evidence that researchers might accept usage as a metric for measuring quality? Nicholas and Rowlands of the CIBER Institute for Research at University College London reported the results of a survey (2005) based on emailing a questionnaire to over 30,000 researchers to ask them a number of questions about how they used online information and their perceptions of several different areas. Figure 1 below, which is taken from their report[1], compares the answers from the 5513 responses received to two questions, one about the use of citations and the second about the use of downloads as a measure of the usefulness of research.



From *New journal publishing models: an international survey of senior researchers*; Ian Rowlands and Dave Nicholas, A CIBER report for the Publishers Association and the International Association of STM Publishers, 22 September 2005

Figure 1

As the figure shows, there is a very strong correlation in the responses to these questions. Rowlands and Nicholas concluded that researchers believe that article downloads offer a measure of the usefulness of research which is as good as, if not slightly better than, author citations. They also concluded that download metrics would have considerable credibility amongst the author

community and would certainly be of great appeal to librarians and many publishers as well as to researchers.

Investigation of Usage Factor

As a result of this positive feedback from the research community, in 2007 the United Kingdom Serials Group launched a project in collaboration with COUNTER to examine the potential usefulness of a Usage Factor if it were to be developed. The Usage Factor Project was divided into two stages, with the first stage consisting of three phases. In the first phase of stage 1, a series of interviews was conducted with the following groups: 7 authors, 13 publishers, and 9 librarians.

Phase 1 - Survey results

In answer to the question, ‘Would Journal Usage Factors be helpful to you in assessing the value, status and relevance of a journal’, 100% of librarians and 100% of authors replied positively.

Interview participants indicated that a Usage Factor would have several advantages.

Firstly, it would provide a counterweight to the Impact Factor, giving a useful additional perception of journal value. Secondly, the Usage Factor would be particularly helpful for those journals and subject areas not so comprehensively covered by ISI, particularly in the humanities and social sciences, and thirdly, the Usage Factor would be especially helpful for journals that had a high readership outside of the research community, for example, among undergraduates or practitioners. It would also be especially helpful for journals publishing a relatively small number of articles per year. Lastly, the data could be produced much more quickly than the Impact Factor.

Below are two quotes from participants in this first phase of the research. The first quote is from a well-known professor of information science at the University of Tennessee, Carol Tenopir, who has

published 5 books and over 200 research articles on issues relating to the scientific and academic information industry and the ways in which technology impacts on authors and readers. In summarizing the difference between the Impact Factor and Usage Factor she said, “Authors select journals that will give their articles prestige and reach. The Impact Factor is a widely used surrogate for the former, while perceived circulation and readership reflect the latter. But usage is becoming more important as a measure of reach.” The second quotation is typical of the responses from authors who were interviewed: “Many of the publications in which I publish and in which I would like to publish do not have Impact Factors and the current system almost requires serious authors to publish in journals that have Impact Factors.”

Phase 2 – Results of broader survey

The second phase of the research broadened the number of people surveyed by undertaking a web-based survey of 155 librarians and 1400 authors to give a more quantitative approach. Figure 2 shows the responses to the question, ‘Would you welcome the development of new quantitative measures to help assess the value of scholarly journals based upon verifiable data which describes the number of times articles from those journals have been downloaded?’ The results presented here correlate very strongly to the CIBER research mentioned earlier.

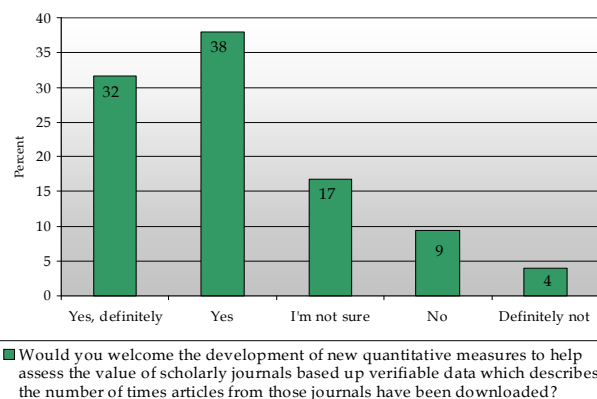


Figure 2: Results of author survey to assess support for new usage based measure.

Librarian results: existing journals

The next question analyzed the answers to the question in Figure 2 in more detail and found that the responses to the ‘Yes, definitely’ section were highest in the biomedical sciences and lower in the other areas, whereas the other responses were very similar from one subject area to another.

The web survey also asked librarians several questions about the Usage Factor, for example, they were asked to rank their current criteria for selecting new journals to subscribe to. The survey then defined and explained the Usage Factor concept and asked the librarians to re-rank their criteria for new journal selection if a Usage Factor was already in existence. In response, the librarians placed the Usage Factor as the second highest criterion, ahead of Price, Impact Factor, and even the reputation of the publisher.

Ranking without Usage Factor	Ranking with Usage Factor
1. Feedback from library users	1. Feedback from library users
2. Usage	2. Usage
3. Price	3. Usage Factor
4. Cost per Download	4. Price
5. Impact Factor	5. Cost per Download
6. Reputation/status of publisher	6. Impact Factor
	7. Reputation/status of Publisher

This difference in ranking was explained by one of the librarians who said he/she would view the Usage Factor as an aid for collection rather than cancellation decisions, with usage per se being a more suitable tool when considering cancellation.

Development of plan to launch a Usage Factor

Based on the enthusiastic feedback received from researchers and librarians as well as publishers, the project then decided to put in place a second stage to develop a plan for launching a Usage Factor. Under this second stage, journal usage logs from many publishers will be collected together and converted to a uniform standard report format for analysis by an expert third party. An invitation to submit a tender for this work is currently being written to send out to potential organizations that might undertake this work.

However, before this invitation to tender can be issued, there are a number of outstanding items that need to be addressed. Firstly, it will be necessary to agree on the exact format of the standard report that will be produced. Secondly, it will be very important to ensure consistency in data, data integrity, and the fitness for purpose of this data. For example, it will be necessary to decide how to measure the number of qualifying items published in the journal, how to assign a correct publication year for each of those items and how to exclude the use of computers; robots, spiders, etc.

Librarian results: new journals

Ranking without Usage Factor	Ranking with Usage Factor
1. Feedback from library users	1. Feedback from library users
2. Price	2. Usage Factor
3. Reputation/status of publisher	3. Price
4. Impact Factor	4. Impact Factor
	5. Reputation/status of publisher

Librarians were then asked a similar question, but this time looking at journals that they already subscribed to. The first column gives the ranking of the factors that librarians gave without considering the Usage Factor. When they considered the Usage Factor, it was ranked at number 3, below usage but above price, cost per download, Impact Factor, and reputation/status of publisher.

With respect to the question about the number of qualifying items, it has now been decided to include anything with a Digital Object Identifier which is also included in one of a number of the major indexes of journal articles. Deciding on the correct publication year is also an interesting challenge because there can sometimes be a significant gap between the online release of an article and the final, formal version. The project team has also completed a successful search for a third-party classification system from the British Library that will allow each journal to be assigned an agreed subject category.

Conclusion

The project plans to publish a final report at the beginning of next year, which will outline the various metrics assessed. It will also recommend which of the metrics proved consistent and robust enough to be adopted for scaling up. The report will also suggest any ways in which data providers might capture their usage data to make the measurement of Usage Factors easier and more reliable. It will also propose ways to audit Usage Factors for accuracy, just as COUNTER statistics are audited.

(For more information on the project, here is the URL:
www.uksg.org/usagefactors)

Reference

- [1] New journal publishing models: an international survey of senior researchers; Ian Rowlands and Dave Nicholas, A CIBER report for the Publishers Association and the International Association of STM Publishers, 22 September 2005