



The Journal Impact Factor and its discontents: steps toward responsible metrics and better research assessment

José Roberto F. Arruda, Robin Champieux, Colleen Cook, Mary Ellen K. Davis, Richard Gedye, Laurie Goodman, Neil Jacobs, David Ross, Stuart Taylor

Abstract

A small, self-selected workgroup was convened to consider issues surrounding impact factors at the first meeting of the Open Scholarship Initiative in Fairfax, Virginia, USA, in April 2016, and focused on the uses and misuses of the Journal Impact Factor (JIF), with a particular focus on research assessment. The workgroup's report notes that the widespread use, or perceived use, of the JIF in research assessment processes lends the metric a degree of influence that is not justified on the basis of its validity for those purposes, and retards moves to open scholarship in a number of ways. The report concludes that indicators, including those based on citation counts, can be combined with peer review to inform research assessment, but that the JIF is not one of those indicators. It also concludes that there is already sufficient information about the shortcomings of the JIF, and that instead actions should be pursued to build broad momentum away from its use in research assessment. These actions include practical support for the San Francisco Declaration on Research Assessment (DORA) by research funders, higher education institutions, national academies, publishers and learned societies. They include the creation of an international "metrics lab" to explore the potential of new indicators and the wide sharing of information on this topic among stakeholders. Finally, the report acknowledges that the JIF may continue to be used as one indicator of the quality of journals and makes recommendations how this can be improved.

OSI2016 Workgroup Question: Impact Factors

Tracking the metrics of a more open publishing world will be key to selling "open" and encouraging broader adoption of open solutions. Will more openness mean lower impact, though (for whatever reason—less visibility, less readability, less press, etc.)? Why or why not? Perhaps more fundamentally, how useful are impact factors anyway? What are they really tracking, and what do they mean? What are the pros and cons of our current reliance on these measures? Would faculty be satisfied with an alternative system as long as it is recognized as reflecting meaningfully on the quality of their scholarship? What might such an alternative system look like?

Introduction

This short report describes the outcomes of a small, self-selected workgroup convened at the first meeting of the Open Scholarship Initiative in Fairfax, Virginia, USA, in April 2016. It is made available as an aid for further discussion, rather than with any claims to being an authoritative text.

Background

The Journal Impact Factor (JIF) is a score based on the ratio of citations to papers published in a journal over a defined period, to the number of papers published in that journal over that period. It is calculated over the dataset provided by the *Journal Citation Report* (JCR) (Thomson Reuters). The JIF is widely used, and misused. The factors influencing it, and their implications have been well documented elsewhere.¹

How does the existence and use of the JIF affect moves toward open scholarship?

Scholarly communication is a complicated system, with subtle relationships between components and some unexpected feedback loops. As a result, it is rather difficult to pin down a direct causal relationship between the existence and use of the JIF, and moves toward open scholarship. In particular, the relationship between a journal's JIF (or lack of one) and its perceived prestige can be subtle. There is probably enough evidence, though, to justify the claim that the JIF inhibits openness and that action should be taken to reduce its influence.

The power of the JIF stems largely from its misuse in research assessment and, especially, in funding, recruitment, tenure and promotion processes. There is both a perception and a reality that such processes are influenced by the JIF, and so researchers who are subject to those processes understandably adjust their publishing behaviour based on the JIF. It would be hard to over-state the power this gives the JIF. So, given the JIF's influence, what are the effects of its use and misuse? We focus here on those effects related specifically to open scholarship.

The influence of the JIF can retard uptake of open practices. For example, whereas hybrid journals are usually well-established titles that have had time to build an impact factor and so attract good authors, wholly Open Access (OA) journals are often new titles, and therefore not in so strong a position. There are a few high-profile exceptions to this, notably:

- eLife, a very new OA journal with a high impact factor, though it is unusual in several ways;
- PLOS Biology, an OA journal that has built up a high impact factor;
- Nucleic Acids Research, a well-established journal, successfully flipped to OA by Oxford University Press in part because its high prestige (JIF) protected it against author concerns about its quality.

These are the exceptions, however; in general, the JIF imposes a high barrier to entry for journals, and since OA is an innovation in journal publishing, that barrier is particularly acute for OA journals. As soon as one moves beyond conventional journal publishing (for example, models such as *F1000 Research*² or pre-

print repositories) the influence of the JIF is extremely strong and inhibits take-up by authors. Furthermore, the JIF is based on a largely Anglophone dataset (the JCR), which makes it likely that the JIF particularly disadvantages alternative models of scholarly communication outside the “global north.” There are operational implications here, especially where the JIF is used in research assessment, but there are also implications with respect to research culture and values.

Without going into current debates about the functioning of the Article Processing Charge (APC) market, a high JIF can be used by publishers to justify a high APC level for a journal, despite concerns about whether this is legitimate.

But open scholarship is about more than just OA, it also includes sharing research data, methods and software, the pre-registration of protocols and clinical trials, better sharing of the outcomes of all research including replication studies and studies with negative results, and early sharing of information about research outcomes. The power of the JIF acts against all of these aspects, for example by not counting all the specific kinds of research output, or by focusing on authorship as the sole contribution to a research output. The influence of the JIF can also weaken the position of low JIF journals, which then risk losing authors if the journals put up perceived barriers to submission such as data sharing requirements, while strengthening the position of high JIF journals, which may then prevent early disclosure of research findings for fear of being scooped by the science press. Another key problem is the distortion of the scholarly record that arises from disproportionately incentivising the publication of papers that are likely to be

cited highly early in their life, as opposed to papers that comprise sound research but are of a type (replication studies, or negative results) that are unlikely to be “citation stars.” Given the highly skewed distribution of citations within a journal, editors seeking to maximise their JIF are incentivised to look out for such “citation stars” that will boost the journal’s JIF. *PLOS One* and other similar journals, which focus their acceptance decisions on research method, not outcome, argue that their success is *despite*—not because of—the power of the JIF.

Of course, the JIF is unable to measure the impact of research beyond merely the citation of papers by other papers. Public engagement, impact on policy, and the enabling of commercial innovation, for instance, are all beyond the scope of JIF. These are all important aspects of open scholarship that could be highlighted by other indicators, and it is troubling that use of the JIF is seldom supplemented by the use of such indicators.

Fundamentally, many of these problems result from the fact that the JIF is an indicator (albeit imperfect) of the quality of the container (the journal) rather than of the research itself.

Finally, but by no means less importantly, the JIF is not itself open. Neither the dataset nor the algorithm is truly open, which flies in the face of moves toward a more transparent approach to scholarship. There are moves, such as the forthcoming Crossref Event Data service,³ and various other open citation initiatives,⁴ that might address this problem in due course.

Research assessment and the JIF

As a result of the above and other considerations, our team reached consensus on the following six points:

1. There is a need to assess research and researchers, to allocate funding and to make decisions about tenure and promotion.
2. JIF is not appropriate for these purposes.
3. No single metric would be appropriate for these purposes either.
4. A number of metrics may be developed which can help inform these decisions (including, but not limited

to, “altmetrics”⁵) in addition to peer review.

5. Some of these metrics might be based on citation data.
6. Enough information exists about the issues and shortcomings of the JIF⁶ to render further significant research on this unnecessary.

Action plans

To improve the current situation, and move toward responsible metrics and better research assessment in support of open scholarship, the workgroup proposes the following actions:

#	Intended change	Specific actions
1	The DORA recommendations should be implemented.	<ol style="list-style-type: none"> i. Research funders should only provide funding to higher education institutions that have signed DORA and that have published a recruitment, tenure and promotion framework, which demonstrates their implementation of the DORA recommendations. <ol style="list-style-type: none"> a) Future OSI workgroups focused on indicators or impact factors should assess the initial response of research funders, especially in the biomedical field, to this proposed action and amend the following actions accordingly. b) National academies should gather and present evidence to inform the case for funders to take this action, and should release open invitations to funders to join this conversation via meetings, workshops or other forums. c) National academies, senior institutional representative organisations, and research funders should agree on how this action can be implemented to greatest effect and with the least burden in their particular national context.

#	Intended change	Specific actions
1	DORA recommendations (cont).	<ul style="list-style-type: none"> d) National academies, learned societies, and institutional representative organisations should work with senior academics in universities to ensure that this action finds support in the academic community. e) Supportive funders should recommend this action to their peers, e.g. through the Global Research Council. ii. OSI workgroups focused on indicators or impact factors should support DORA's publicity and marketing efforts, including gathering testimonials from those who have signed it, and investigating why others have not. iii. Funders or institutions that are already implementing the DORA recommendations in their internal evaluation processes should be asked to declare this publicly.⁷ iv. The meetings recommended above should be used by all stakeholders as an opportunity for discussion of the wider issues associated with metrics, research assessment and open scholarship.
2	Disciplines take ownership over the assessment of research in their area, through the development and use of tools, checklists and codes of conduct.	<ul style="list-style-type: none"> i. Create templates for universities / disciplines, to facilitate the development of appropriate tenure and promotion frameworks to implement DORA (see 1, above). Relevant learned societies should create discipline-specific outline templates based on DORA and existing evidence on good practice in using evidence in research evaluation. These efforts should be informed iteratively as further evidence becomes available on the potential of indicators, e.g., from the metrics lab (see 3, below). This work should be done in consultation with relevant funders and university representatives; some limited international coordination may be beneficial and practical. ii. OSI workgroups focused on indicators or impact factors should discuss with learned societies whether author-publishing practices (in particular avoiding reference to the JIF in publishing decisions) should be part of the scope of their codes of practice.

#	Intended change	Specific actions
3	Create an international metrics lab, learning from prior attempts to do this. This would include: data sources; developers to explore and propose indicators; incentives to participate; and tests for reliability, validity, and acceptability of proposed indicators.	<ul style="list-style-type: none"> i. OSI workgroups focused on indicators or impact factors should build a coalition of parties willing to undertake this effort. At a first pass, this coalition might include Force11, Crossref Labs, Association of Research Libraries, Jisc, Snowball Metrics, NISO, COUNTER and other standards bodies, representatives of publishers (e.g., STM futures lab), and funders. ii. This coalition should identify a trusted organisation to lead the metrics lab initiative or, at least, to coordinate it. iii. The coalition should define the terms of reference for the metrics lab. iv. The coalition should identify funding, governance and operational options. v. The coalition should commission work to create and maintain a register of open data sources that could underpin useful indicators, e.g. OpenURL, Crossref Event Data.
4	Share information about the JIF, metrics, their use and misuse	<ul style="list-style-type: none"> i. OSI should add a resources page on its website to bring this information together and publicise it. The Metrics Dashboard, a pilot project recently funded by FORCE11, which aims to provide actionable information on research metrics use and misuse could be leveraged as a data source. Additionally, the page should include the NISO use cases for altmetrics,⁸ Crossref Event Data,⁹ the UK Metric Tide report,¹⁰ DORA,¹¹ the Leiden Manifesto,¹² the NIH Biosketch,¹³ CRediT,¹⁴ etc.

In addition to the above actions, which are specifically about the use of metrics in research assessment (where the JIF is not appropriate), the following actions are

proposed to improve how journals are compared. This is a different and entirely separate use case to research assessment, and the JIF may be a useful indicator here.

#	Intended change	Specific actions
1	Improve the validity of the JIF as one indicator of journal quality	<ul style="list-style-type: none"> i. OSI workgroups focused on indicators or impact factors should draft a list of improvements required to the JIF to improve its validity and openness. ii. OSI workgroups focused on indicators or impact factors should gather support for this list and present it to the owners of the JIF.
2	Investigate whether best practice or standards can	<ul style="list-style-type: none"> i. OSI workgroups focused on indicators or impact factors should identify a willing partner to commis-

	<p>be agreed to describe and measure aspects of journal publishing services, e.g. to inform the operation of journal comparison sites</p>	<p>sion a landscape review and analysis of how journal publishing services for authors are already being compared, the criteria used, the rigour of the assessment, etc.</p> <p>ii. OSI workgroups focused on indicators or impact factors should identify a willing partner to commission landscape review and analysis of how journal publishing services for readers (and librarians) are already being compared, the criteria used, the rigour of the assessment, etc.</p> <p>iii. OSI workgroups focused on indicators or impact factors should consider the findings of these two studies and recommend next steps.</p>
--	---	---

Challenges

Some significant challenges and questions toward the implementation of these actions exist that are not specific to this workgroup but are general to OSI. They include:

1. How to continue to engage the OSI participants in this activity, to ensure we remain active and effective?
2. What channels and methods should be used to effectively extend the participation to represent fully all stakeholders from around the world?
3. Given limited resources, how should the work that we have proposed be prioritized?

The OSI Impact Factor Workgroup

Workgroup delegates comprised a wide mix of stakeholders, with representatives from Brazil, Canada, the United Kingdom, and the United States:

José Roberto F. Arruda, São Paulo State Foundation (FAPESP), Brazil

Robin Champieux, Scholarly Communication Librarian, Oregon Health and Science University, USA. ORCID: 0000-0001-7023-9832

Dr Colleen Cook, Trenholme Dean of the McGill University Library, Canada

Mary Ellen K. Davis, Executive Director, Association of College & Research Libraries, USA

Richard Gedye, Director of Outreach Programmes, International Association of Scientific, Technical & Medical Publishers (STM). ORCID: 0000-0003-3047-543X

Laurie Goodman, Editor-in-Chief, GigaScience. ORCID: 0000-0001-9724-5976

Dr Neil Jacobs, head of scholarly communications support, Jisc, UK. ORCID: 0000-0002-8050-8175

David Ross, Executive Director, Open Access, SAGE Publishing. ORCID: 0000-0001-6339-8413

Dr Stuart Taylor, Publishing Director, The Royal Society, UK. ORCID: 0000-0003-0862-163X

Notes:

¹ For example: the Metric Tide report: as of May 24, 2016:

<http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title,104463,en.html>;

San Francisco Declaration on Research Assessment, DORA, as of May 24, 2016:

<http://www.ascb.org/dora/>; Leiden Manifesto, as of May 24, 2016:

<http://www.leidenmanifesto.org/>

² F1000 Research, as of May 24, 2016: <http://f1000.com/>

³ Crossref DOI event data service, as of May 24, 2016: <http://eventdata.Crossref.org/>

⁴ For example, CORE semantometrics experiment, as of May 24, 2016:

<http://www.slideshare.net/JISC/introducing-the-open-citation-experiment-jisc-digifest-2016-58968840>; Open Citation Corpus, as of May 24, 2016:

<https://is4oa.org/services/open-citations-corpus/>; CiteSeerX, as of May 24, 2016:

<http://citeseerx.ist.psu.edu/index.jsessionid=9C3F9DA06548EACB52B7E8D50E9009F2>

⁵ See NISO altmetrics initiative, as of May 24, 2016:

http://www.niso.org/topics/tl/altmetrics_initiative/#phase2

⁶ E.g., HEFCE (2015) The Metric Tide, as of May 24, 2016:

<http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title,104463,en.html>; and

studies such as Kiesslich T, Weineck SB, Koelblinger D (2016) Reasons for Journal Impact Factor Changes: Influence of Changing Source Items. PLoS ONE 11(4): e0154199.

doi:10.1371/journal.pone.0154199

⁷ Indiana University Bloomington has recently made a strong statement in this direction, as of May 24, 2016: <http://inside.indiana.edu/editors-picks/campus-life/2016-05-04-from-the-desk.shtml>

⁸ NISO altmetric initiative, as of May 24, 2016:

http://www.niso.org/topics/tl/altmetrics_initiative/#phase2

⁹ Crossref Event Data, as of May 24, 2016: <http://eventdata.Crossref.org/>

¹⁰ Metric Tide report, as of May 24, 2016:

<http://www.hefce.ac.uk/pubs/rereports/Year/2015/metrictide/Title,104463,en.html>

¹¹ DORA, as of May 24, 2016: <http://www.ascb.org/dora/>

¹² Leiden Manifesto, as of May 24, 2016: <http://www.leidenmanifesto.org/>

¹³ Example of NIH Biosketch, as of May 24, 2016:

<https://grants.nih.gov/grants/funding/2590/biosketchsample.pdf>

¹⁴ CASRAI CRediT, as of May 24, 2016: <http://casrai.org/credit>