



Should highly cited items be excluded in impact factor calculation? The effect of review articles on journal impact factor

Lei Lei¹ · Yunmei Sun¹

Received: 1 October 2019 / Published online: 20 January 2020
© Akadémiai Kiadó, Budapest, Hungary 2020

Abstract

The use of the journal impact factor for evaluation purposes has been widely challenged because of the simplistic algorithm involved in impact factor calculation. As a result, a journal's impact factor may be affected by many variables, such as highly cited review articles. In the present study, we aimed to investigate whether and to what extent review articles affect the impact factors of journals. We examined the citations of original research articles and review articles in the field of physics, and compared the impact factors of the journals *Science*, *Nature*, and *Cell* with and without citation counts of review articles included. The results showed that review articles consistently received higher numbers of citations than did original research articles. In addition, review articles accounted for 3–10% of the impact factors of these top journals. Based on our findings, we suggest that sophisticated methods, in addition to the impact factor, be employed for research evaluation purposes.

Keywords Review articles · Impact factor · Highly cited articles

Introduction

The journal impact factor¹ was first developed to help librarians select quality journals (Garfield 1972, 2006). It has been widely but controversially used as a measure of journal, research, and researcher evaluation (Bornmann and Marx 2018; Lei and Yan 2016; Liu et al. 2016; Waltman 2016). One possible limitation that may challenge the validity of the journal impact factor when used as a measure of evaluation is its relatively simplistic algorithm. A journal's impact factor in a certain year is calculated as the number of citations received for all its items (such as original research articles, review articles, proceedings

¹ Journal Impact Factor is a copyright of Clarivate Analytics, formerly known as The Institute for Scientific Information (ISI).

✉ Yunmei Sun
sunyunmei@hust.edu.cn

¹ School of Foreign Languages, Huazhong University of Science and Technology, Wuhan 430074, China

papers, editorials) published during the previous 2 years, divided by the number of citable items (Clarivate Analytics 2018). For example, a journal's impact factor in 2016 is calculated as shown in Eq. (1):

$$\text{Impact factor in 2016} = \frac{\text{Citations in 2016 to all items published in 2014 and 2015}}{\text{Number of citable items published in 2014 and 2015}} \quad (1)$$

Based on Eq. (1), if one or more items published in 2014 or 2015 received many more citations than other items, then these most highly cited items might help boost the impact factor of the journal in 2016. That is, the journal impact factor may be dramatically affected by the distorted distribution of citations. As one of the reviewers correctly points out, the citation counts are not normally distributed, and thus the mean citation number is a potentially highly skewed or misleading variable. While the median may seem to be a better estimate in the case of non-normal distribution, its use as an index for differentiating journals is not feasible, since most journals would receive an impact factor of 0 or 1. One extreme example of the issue with highly cited items is Sheldrick (2008), which was found to temporarily but dramatically help boost the impact factor of its host journal, from 2.052 in 2008 to 49.926 in 2009 and 54.333 in 2010, dropping back to 2.076 in 2011 (Dimitrov et al. 2010; Foo 2013; Krauskopf 2013). Liu et al. (2018) recently reported another example of such a phenomenon. They found that the “Review of Particle Physics” series articles also helped boost the host journals' impact factor (i.e., the RPP phenomenon). Interestingly, although the RPP phenomenon may substantially affect the impact factor of journals with a lower impact factor via a smaller number of published items, its contribution to the impact factor of journals with a higher impact factor and a larger number of published items is relatively modest. Based on these findings, Liu and colleagues (Liu et al. 2018, p. 3) suggest excluding “the direct effects of publishing the RPP when calculating a journal's impact factor”.

In fact, the Sheldrick (2008) case and the RPP phenomenon may not be unique. A closer look reveals that such items are all review articles. Review articles as a genre serve as “a forum for the synthesis of ideas” (Noguchi 2006, p. 16), which are categorised into history reviews, state-of-the-art reviews, theory reviews, and issue reviews. For example, Sheldrick (2008) is a history review, and the RPP articles are state-of-the-art reviews. Since review articles synthesise ideas and the state of the art in a research area, we hypothesise that they attract much more attention from colleagues and hence receive many more citations than academic pieces of other genres such as original research articles. Few studies have provided empirical evidence for this hypothesis, though researchers mention the point in their arguments (e.g., Martin 2016). For example, Amiri and Michel (2018) examined the citation counts of articles in five pharmacology journals, and found that review articles received twice as many citations as original research articles. If the foregoing hypothesis is correct, then the impact factors of the host journals may be boosted not only by the RPP articles but also by the higher numbers of citations to other review articles. Accordingly, the Sheldrick (2008) case and the RPP phenomenon as reported in Liu et al. (2018) may not be unique, but rather pervasive.

To test this hypothesis, we conduct experiments to address two research questions as follows:

RQ1: Do review articles receive more citations than original research articles?

RQ2: Do the citations that review articles receive affect the impact factor of their host journals?

In the remainder of this article, we first introduce our research methods, and then report the findings. We end the article with some discussion.

Data and methods

The data for the experiments were extracted from the Web of Science on the portal of Huazhong University of Science and Technology on October 26–28, 2018. For RQ1, we extracted citations of both original research articles and review articles in the field of physics from 2009 to 2018. Concerning the search techniques, we first used “SU = Physics AND PY = 2009–2018” in the advanced search box to retrieve all items in physics published between 2009 and 2018, and then refined the results with “highly cited in field” and “article” or “review” for document type (i.e., other document types such as editorials and commentaries were excluded). The search code is as follows.

SU = Physics AND PY = 2009–2018

Refined by: ESI Top Papers: (Highly Cited in Field) AND DOCUMENT TYPES: (ARTICLE OR REVIEW)

Time span: 2009–2018. Indexes: SCI-EXPANDED.

Two points should be noted here. First, the information extracted on original research articles and review articles was refined by “highly cited papers”, since the Web of Science only reports citations of publications up to 10,000 items. Second, the data from 2009 to 2018 were used because the Web of Science only reports the citations of highly cited papers (the top 1% highly cited items in a field) in the most recent 10 years (InCites Help 2018).

For RQ2, we investigated the bibliometric information for *Science*, *Nature*, and *Cell*. We chose these journals as the targets for three reasons. First, *Science*, *Nature*, and *Cell* are widely recognised top journals in science. Second, they publish a large number of pieces each year. Third, and most importantly, they publish more than 60 review articles each year. The relatively large number of review articles provides a data set large enough for our experiment. In fact, we hoped to experiment with the journals examined in Liu et al. (2018). However, these journals publish a very limited number of review articles in certain years. For example, *Physical Review D*, the journal with the largest publication size, published only one review article in 2012, and none from 2013 to 2017, according to the records in the Web of Science. With this very limited quantity, the validity of the findings from the journals should be challenged. The search techniques used for the retrieval of the data used in this part were largely similar to those used in the previous part, but with different parameters such as journal titles and years.

Using the bibliometric data extracted from the Web of Science, we first calculated the impact factor of the journals based on Eq. (1) in the most recent 10 years (from 2008 to 2017), and then computed the adjusted impact factor excluding review articles based on Eq. (2) below (taking the impact factor in 2016 as an example). With the values of impact factor and adjusted impact factor, we were able to examine whether the impact factor was affected by the citations of review articles. It is worth noting that we calculated impact

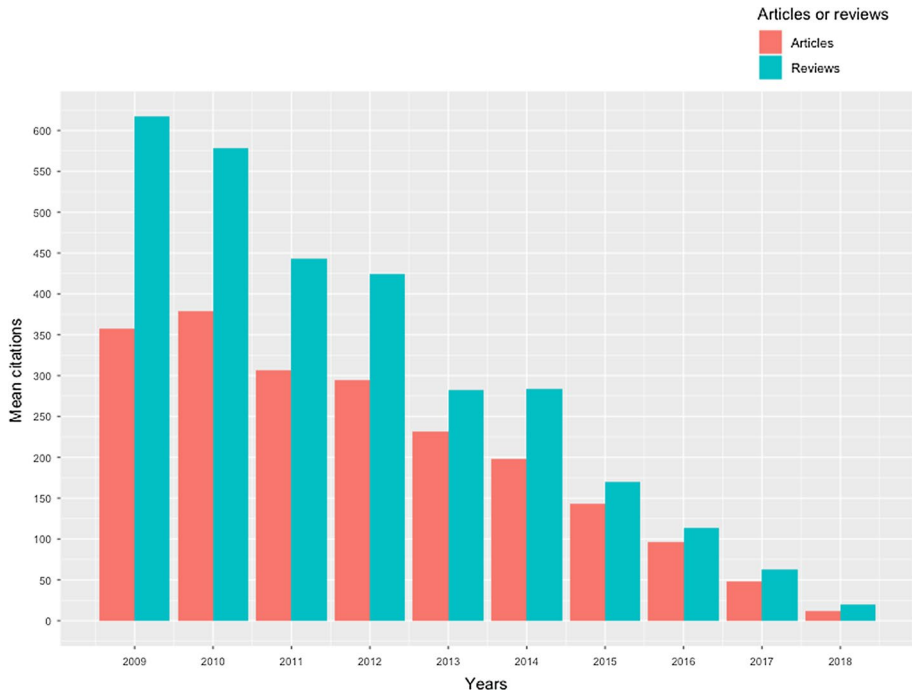


Fig. 1 Mean numbers of citations of highly cited original research articles and review articles

factors for the most recent 10 years, i.e., from 2008 to 2017, in that the data for 2018 was incomplete at the time of this study. Accordingly, the end year should be 2017.

Adjusted impact factor in 2016

$$= \frac{\text{Citations in 2016 to all items excluding review articles published in 2014 and 2015}}{\text{Number of citable items excluding review articles published in 2014 and 2015}}. \quad (2)$$

Results

Do review articles receive more citations than original research articles?

The statistics for the mean citation numbers for the highly cited original research articles and the review articles from 2009 to 2018 are plotted in Fig. 1. It is obvious that the mean citation numbers for the review articles are always higher than those for original research articles. These findings are in accordance with our hypothesis that review articles may receive more citations than original research articles. This finding may also provide empirical evidence for such intuitive arguments in earlier studies, such as Martin (2016). As previously mentioned, review articles integrate ideas, findings, and state-of-the-art development in a research area, and may offer insights into research trends in

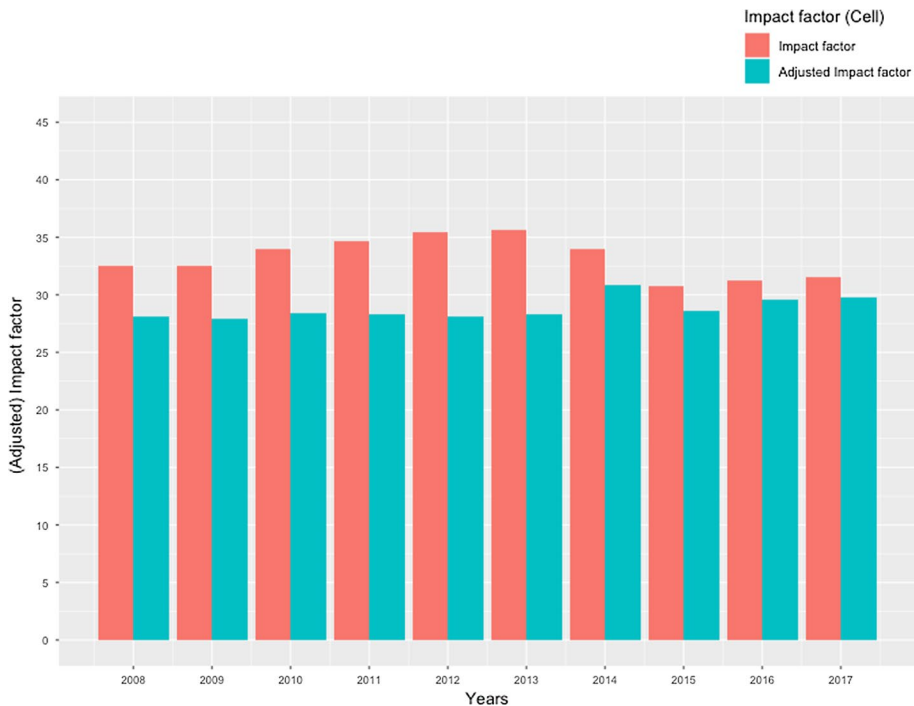


Fig. 2 (Adjusted) impact factors of *Cell* articles and review articles

the field, which as a result attract attention from colleagues and elicit citation in their future publications. In the next section, we will report results on the issue of whether the citations of review articles affect the impact factor of journals as was seen in the Sheldrick (2008) case and the RPP phenomenon (Liu et al. 2018a).

Do the citations that review articles receive affect the impact factor of their host journals?

The impact factors and adjusted impact factors for *Science*, *Nature*, and *Cell* are depicted in Figs. 2, 3, and 4. It is clear that the impact factors of the journals are higher than their adjusted impact factors for all the years examined. That is, when the review articles are removed, the impact factors of the journals will decrease. These findings confirm our hypothesis that citations to review articles affect the impact factor of their host journals. Such findings also seemingly provide partial evidence to support our argument that the Sheldrick (2008) case and the RPP phenomenon as reported in Liu et al. (2018a) are not unique, but rather pervasive, as long as a journal publishes review articles.

It should be noted that the contribution of review articles to the impact factors of journals in the present experiment is not as large as that in the Sheldrick (2008) and RPP cases. The smaller contribution of review articles may be explained by two points. First, the journals examined in the study, i.e., *Science*, *Nature*, and *Cell*, are top

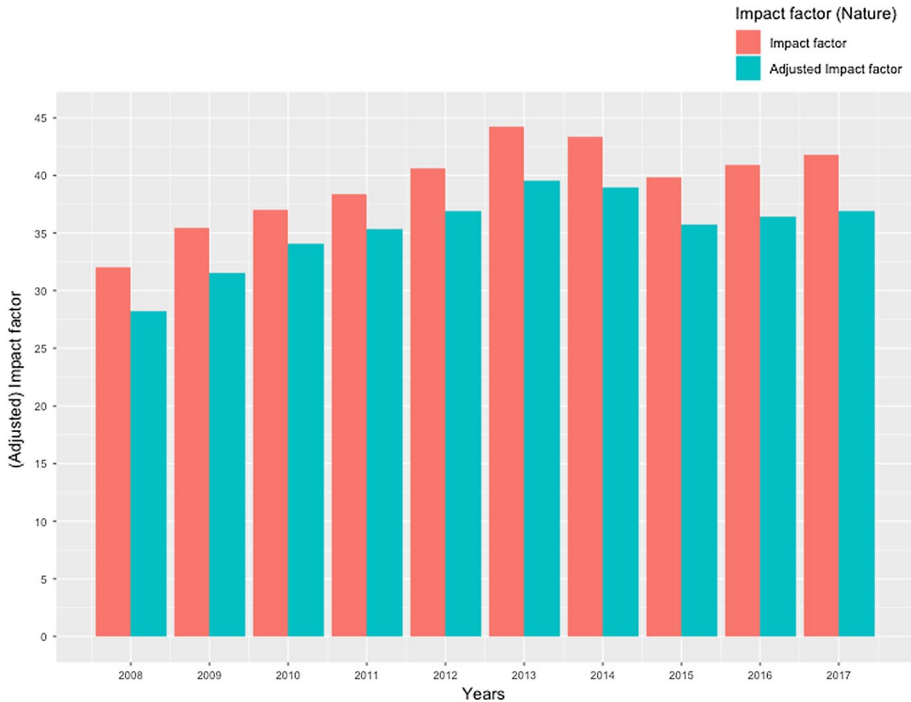


Fig. 3 (Adjusted) impact factors of *Nature* articles and review articles

journals with high impact factors and large numbers of published pieces. Accordingly, the higher citations of review articles may contribute relatively modestly to the impact factor of these journals. Such a phenomenon was also seen by Liu et al. (2018a). In their study, the RPP effect on the impact factor of a journal with a high impact factor and a large number of published items, i.e., *Physical Review D*, was much smaller than that for other journals with lower impact factors and a smaller number of published pieces. Second, the citation gap between review articles and original research articles in *Science*, *Nature*, and *Cell* is much smaller than that between the Sheldrick (2008) and RPP cases and the original research articles in the journals examined in Liu et al. (2018a). The smaller citation gap may also explain the relatively modest contribution of review articles to the impact factor of journals in our experiment.

Despite the foregoing points, we cannot ignore the contributions of review articles to the impact factors of the journals. As the statistics presented in Table 1 show, the review articles account for an average of 10% of the impact factor of *Cell*, and 3% of that of *Nature* and *Science*. Indeed, in certain years, the contribution of review articles is nearly 20% of the impact factor. For example, in 2012 and 2013, 18% of the impact factor is attributable to citations of review articles in *Cell*. Without the review articles, the impact factor would decrease from 35.45 to 29.07 in 2012 and from 35.59 to 29.08 in 2013. In addition, some extremely highly cited review articles are found in these top

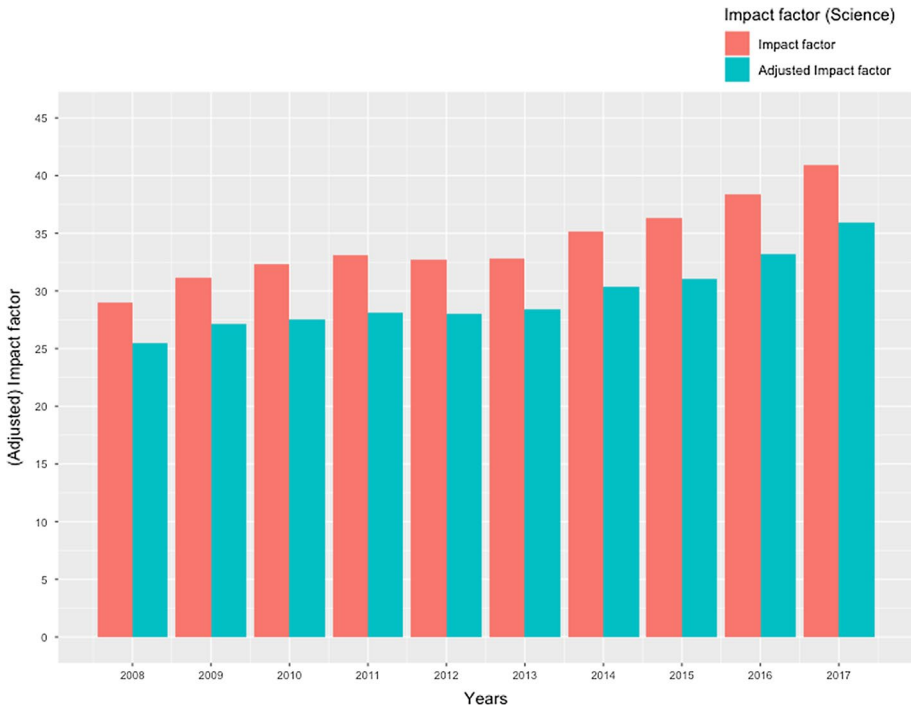


Fig. 4 (Adjusted) impact factors of *Science* articles and review articles

Table 1 Statistics for impact factors (IF) and adjusted impact factors

Year	<i>Cell</i>			<i>Nature</i>			<i>Science</i>		
	IF	Adjusted IF	Contribution of reviews	IF	Adjusted IF	Contribution of reviews	IF	Adjusted IF	Contribution of reviews
2008	32.52	30.14	0.07	32.03	30.43	0.05	29.02	28.72	0.01
2009	32.53	29.34	0.10	35.43	33.83	0.05	31.18	31.04	0.00
2010	33.95	29.62	0.13	37.03	36.19	0.02	32.35	31.54	0.03
2011	34.64	29.53	0.15	38.34	37.73	0.02	33.05	31.86	0.04
2012	35.45	29.07	0.18	40.58	39.71	0.02	32.69	31.48	0.04
2013	35.59	29.08	0.18	44.22	42.57	0.04	32.78	31.76	0.03
2014	33.99	31.49	0.07	43.35	41.74	0.04	35.11	34.04	0.03
2015	30.75	29.17	0.05	39.85	38.99	0.02	36.34	34.94	0.04
2016	31.28	30.41	0.03	40.95	40.10	0.02	38.37	37.37	0.03
2017	31.56	30.91	0.02	41.81	40.59	0.03	40.95	39.70	0.03
Mean	33.22	29.88	0.10	39.36	38.19	0.03	34.19	33.25	0.03

journals with high impact factors. For example, a review article published in *Cell* in 2011 (Hanahan and Weinberg 2011) had received a total of 5313 citations 2 years after its publication (2415 in 2012 and 2898 in 2013). This review article was responsible for 73.44% ($5313/7234 = 73.44\%$) of all citations to review articles in 2013 and 19.82% ($5313/26,802 = 19.82\%$) of all citations to all items published in *Cell* in 2013. This example also offers evidence to our point that the Sheldrick (2008) case and the RPP effect are not isolated phenomena.

Discussion

The present study finds that review articles consistently receive more citations than original research articles. This result provides empirical evidence for such intuitive arguments in previous studies (e.g., Martin 2016). The study also reveals that review articles affect the impact factor of journals. Specifically, review articles have been responsible for an average of 3–10% of the impact factors of the top journals *Cell*, *Nature*, and *Science* over the past 10 years. In addition, certain review articles are responsible for a relatively large proportion of citations to these top journals. One extreme example is a review article published in *Cell* in 2011 (Hanahan and Weinberg 2011), which accounted for approximately 20% of all citations to items published by the journal in 2013. Accordingly, the findings of the study show that the Sheldrick (2008) case and the RPP phenomenon (Liu et al. 2018a) are not unique. In fact, review articles play an important role in journal impact factors.

The impact factor of a journal is affected by many variables. For example, in a recent study, Shi et al. (2017) found that the impact factor was affected by the delay of publications in citing journals. Another example is the influence of documents such as editorials and commentaries on journal impact factors. While citations received by editorials and commentaries will be included only in the nominator and not the denominator in impact factor calculations, the impact factors of journals such as *Nature* and *Science* that publish a large number of editorials and commentaries may be skewed because of the citations from such uncitable items. Future research may focus on the influence of such items on journal impact factors. In addition, in Martin (2016), the editor of *Research Policy* lists many seemingly legitimate strategies that people use to manipulate the journal impact factor, one of which involves two points concerning review articles that are of special interest to the present study. First, journals may invite higher numbers of review articles in order to improve their impact factors. Second, review journals often sit at the top of impact factor rankings. Such an important role of review articles may motivate researchers' proposals to exclude review articles in impact factor calculation (e.g., Liu et al. 2018b), particularly those extremely highly cited reviews such as the Sheldrick (2008) case and the RPP series (Liu et al. 2018a).

However, the discussion should not pertain only to the exclusion of highly cited items for impact factor calculation. First, it is no fault of the highly cited items that they attract higher numbers of citations. They attract citations because of their academic value. Second, although highly cited pieces may affect the journal impact factor, they are not necessarily pertinent to the “manipulation” of the impact factor (Falagas and Alexiou 2008; Yu et al. 2010). Third, if the impact factor is affected or distorted by the highly cited items such as the Sheldrick (2008) case and the RPP series (Liu et al. 2018a) and hence

should be removed, should other highly cited pieces such as that published in *Cell* in 2011 (Hanahan and Weinberg 2011) also be excluded? In fact, the highly cited pieces permeate journals. In addition, a quick search of the most highly cited items published in 2017 in physics reveals that the top four pieces are all published in the same journal (Abbott et al. 2017a, b, c; Akerib et al. 2017), and they will surely increase the impact factor of the journal in 2018 and 2019. But it may not be wise to exclude them from the calculation of the impact factor, since if that is the case, then all of the most highly cited items should be excluded. In fact, as one of the reviewers points out, the use of any single index, whether the impact factor or another measure, is inherently limited as the sole means of evaluating journals, individuals, or institutions.

Lastly, at least two points should be considered concerning the journal impact factor issue. On one hand, rather than simply excluding the highly cited items, other sophisticated methods can be employed to prevent manipulation of the impact factor (e.g., Yang et al. 2016; Yu et al. 2011). On the other hand, the key point concerning the impact factor issue may not be the exclusion or inclusion of highly cited items, but the very use of the impact factor for evaluation purposes, which has been widely challenged (Paulus et al. 2018; Stephan et al. 2017). Thus, in addition to simplistic algorithms such as the journal impact factor, many other factors and methods should be considered for use in research evaluation (e.g., Hicks et al. 2015).

Acknowledgements We very much appreciate the reviewers' comments and suggestions. The research work reported in this article was supported in part by the National Social Science Fund of China (Grant No. 17BYY097).

References

- Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., et al. (2017a). GW170104: Observation of a 50-solar-mass binary black hole coalescence at redshift 0.2. *Physical Review Letters*, *118*(22), 221101. <https://doi.org/10.1103/PhysRevLett.118.221101>.
- Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., et al. (2017b). GW170814: A three-detector observation of gravitational waves from a binary black hole coalescence. *Physical Review Letters*, *119*(14), 141101. <https://doi.org/10.1103/PhysRevLett.119.141101>.
- Abbott, B. P., Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., et al. (2017c). GW170817: Observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, *119*(16), 161101. <https://doi.org/10.1103/PhysRevLett.119.161101>.
- Akerib, D. S., Alsum, S., Araújo, H. M., Bai, X., Bailey, A. J., Balajthy, J., et al. (2017). Results from a search for dark matter in the complete LUX exposure. *Physical Review Letters*, *118*(2), 21303. <https://doi.org/10.1103/PhysRevLett.118.021303>.
- Amiri, M., & Michel, M. C. (2018). Do review articles boost journal impact factors? A longitudinal analysis for five pharmacology journals. *Naunyn-Schmiedeberg's Archives of Pharmacology*, *391*(9), 1033–1035. <https://doi.org/10.1007/s00210-018-1529-x>.
- Bornmann, L., & Marx, W. (2018). Critical rationalism and the search for standard (field-normalized) indicators in bibliometrics. *Journal of Informetrics*, *12*(3), 598–604. <https://doi.org/10.1016/j.joi.2018.05.002>.
- Clarivate Analytics. (2018). 2017 Journal impact factor. *Journal Citation Reports Science Edition*. Retrieved from: http://ipscience-help.thomsonreuters.com/incitesLiveJCR/JCRGro_up/jcrOverview.html.
- Dimitrov, J. D., Kaveri, S. V., & Bayry, J. (2010). Metrics: Journal's impact factor skewed by a single paper. *Nature*, *466*, 179. <https://doi.org/10.1038/466179b>.
- Falagas, M. E., & Alexiou, V. G. (2008). The top-ten in journal impact factor manipulation. *Archivum Immunologiae Et Therapiae Experimentalis*, *56*(4), 223–226. <https://doi.org/10.1007/s00005-008-0024-5>.

- Foo, J. Y. A. (2013). Implications of a single highly cited article on a journal and its citation indexes: A tale of two journals. *Accountability in Research*, 20(2), 93–106. <https://doi.org/10.1080/08989621.2013.767124>.
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178(4060), 471–479. https://doi.org/10.1300/J123v20n02_05.
- Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of the American Medical Association*, 295(1), 90–93. <https://doi.org/10.1001/jama.295.1.90>.
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/j.cell.2011.02.013>.
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431. <https://doi.org/10.1038/520429a>.
- InCites Help. (2018). *Highly cited papers*. Retrieved from <http://ipscience-help.thomsonreuters.com/inCites2Live/indicatorsGroup/aboutHandbook/usingCitationIndicatorsWisely/highlyCitedPapers.html>.
- Krauskopf, E. (2013). Deceiving the research community through manipulation of the impact factor. *Journal of the American Society for Information Science and Technology*, 64(11), 2403. <https://doi.org/10.1002/asi.22905>.
- Lei, L., & Yan, S. (2016). Readability and citations in information science: Evidence from abstracts and articles of four journals (2003–2012). *Scientometrics*, 107(3), 1361–1372. <https://doi.org/10.1007/s11192-016-2036-9>.
- Liu, W., Hu, G., & Gu, M. (2016). The probability of publishing in first-quartile journals. *Scientometrics*, 106(3), 1273–1276. <https://doi.org/10.1007/s11192-015-1821-1>.
- Liu, W., Liu, F., Zuo, C., & Zhu, J. (2018a). The effect of publishing a highly cited paper on a journal's impact factor: A case study of the review of particle physics. *Learned Publishing*, 31(3), 261–266. <https://doi.org/10.1002/leap.1156>.
- Liu, W., Zhu, J., Zuo, C., & Wang, H. (2018b). Is it suitable for a journal to bid for publishing a review that is likely to be highly cited? *Science and Engineering Ethics*. <https://doi.org/10.1007/s11948-017-0016-y>.
- Martin, B. R. (2016). Editors' JIF-boosting stratagems—Which are appropriate and which not? *Research Policy*, 45(1), 1–7. <https://doi.org/10.1016/j.respol.2015.09.001>.
- Noguchi, J. (2006). *The science review article: An opportune genre in the construction of science*. Bern: Peter Lang.
- Paulus, F. M., Cruz, N., & Krach, S. (2018). The impact factor fallacy. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2018.01487>.
- Sheldrick, G. M. (2008). A short history of SHELX. *Acta Crystallographica Section A*, 64(1), 112–122. <https://doi.org/10.1107/S0108767307043930>.
- Shi, D., Rousseau, R., Yang, L., & Li, J. (2017). A journal's impact factor is influenced by changes in publication delays of citing journals. *Journal of the Association for Information Science and Technology*, 68(3), 780–789. <https://doi.org/10.1002/asi.23706>.
- Stephan, P., Veugelers, R., & Wang, J. (2017). Reviewers are blinkered by bibliometrics. *Nature*, 544(7651), 411–412. <https://doi.org/10.1038/544411a>.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365–391. <https://doi.org/10.1016/j.joi.2016.02.007>.
- Yang, D.-H., Li, X., Sun, X., & Wan, J. (2016). Detecting impact factor manipulation with data mining techniques. *Scientometrics*, 109(3), 1989–2005. <https://doi.org/10.1007/s11192-016-2144-6>.
- Yu, G., Yang, D.-H., & He, H.-X. (2011). An automatic recognition method of journal impact factor manipulation. *Journal of Information Science*, 37(3), 235–245. <https://doi.org/10.1177/0165551511400954>.
- Yu, G., Yang, D.-H., & Liang, W. (2010). Reliability-based citation impact factor and the manipulation of impact factor. *Scientometrics*, 83(1), 259–270. <https://doi.org/10.1007/s11192-009-0083-1>.