

How Many Papers Should Scientists Be Reviewing? An Analysis Using Verified Peer Review Reports

Vincent Raoult

School of Environment and Life Sciences, University of Newcastle, Ourimbah, NSW 2258, Australia;
Vincent.raoult@newcastle.edu.au

Received: 4 December 2019; Accepted: 14 January 2020; Published: 20 January 2020

Abstract: The current peer review system is under stress from ever increasing numbers of publications, the proliferation of open-access journals and an apparent difficulty in obtaining high-quality reviews in due time. At its core, this issue may be caused by scientists insufficiently prioritising reviewing. Perhaps this low prioritisation is due to a lack of understanding on how many reviews need to be conducted by researchers to balance the peer review process. I obtained verified peer review data from 142 journals across 12 research fields, for a total of over 300,000 reviews and over 100,000 publications, to determine an estimate of the numbers of reviews required per publication per field. I then used this value in relation to the mean numbers of authors per publication per field to highlight a ‘review ratio’: the expected minimum number of publications an author in their field should review to balance their input (publications) into the peer review process. On average, 3.49 ± 1.45 (SD) reviews were required for each scientific publication, and the estimated review ratio across all fields was 0.74 ± 0.46 (SD) reviews per paper published per author. Since these are conservative estimates, I recommend scientists aim to conduct at least one review per publication they produce. This should ensure that the peer review system continues to function as intended.

Keywords: peer review; reviewers; publication; journal; authors; manuscript

1. Introduction

The peer review system underpins modern science [1]. However, the number of scientific publications published every year is increasing exponentially at a rate of ~10% per year [2]. It follows that the number of reviewers and editors required for the peer review system to function is also growing exponentially. Since each review takes approximately 2.6 hours to complete [3], this accounts for a huge number of person-hours for academics. Understanding the cost of reviewing to academia is, therefore, in the best interest of the scientific endeavour. While other aspects of peer review such as its transparency or fairness have been questioned [4], one aspect that has not been examined is the number of reviews that are necessary to balance the peer review process. There are likely large disparities in the number of reviews conducted by various academics, and it is likely that in effect some academics are ‘carrying the weight’ of others that conduct few reviews. Some academics may not believe that they are required to conduct reviews at all. In this context, if the number of reviews required by the peer review system increases faster than the number of reviews being conducted, one would expect a lengthening of the publication process since there would be more publications requiring reviewers than the reviewer pool available to review.

One of the main stressors of the peer review system is the ‘publish or perish’ mentality, which is rife throughout the scientific community. Publish or perish encourages high publication rates at the cost of other aspects of academia such as creative thought and long-term studies [5] and the well-being of researchers [6]. The incentive to publish often is especially high for early career researchers whose career progression is often tied to publication counts [7]. This requirement is in opposition to

the traditional journal peer review process that is often affected by lengthy delays [8]. These delays are believed to be caused by an overtaxing of reviewers [9,10] who identify high workloads as a reason to decline reviews [11]. Editors are also thought to be overtaxed, which by association reduces their selectivity of appropriate studies [12]. Although researchers can benefit from reviewing the work of other scientists [13], increasing numbers of publications are also believed to be responsible for a decline in the general quality of reviews [14]. Perhaps one of the reasons that researchers may not prioritise conducting reviews is that they are not aware how many publications they should be reviewing in the current environment.

There are two aspects to the question ‘how many papers should scientists review’. The first is a practical one, in that if the number of reviews conducted is lower than the number of reviews required by the system, the peer review system should not function as intended. The other is more difficult to quantify and relates to a researcher’s position, experience, career stage, and various circumstances that may lead them to review fewer or more publications. Scientists with significant teaching roles, for example, may not be expected to review as many publications as those that have solely research-focused positions. Researchers who are able and willing to review more papers may not receive many invitations to review. Early career researchers with less experience may not be appropriate reviewers for specific research questions, and late career researchers may be over-targeted to conduct reviews or biased [15]. Gender bias issues have also been found to affect peer review rates [16]. These are factors that may be hard to quantify, which is why focusing on determining the number of reviews required by the peer review system may help researchers by offering a means to improve the peer review system, and provide a basis for research on the second more complex aspect discussed here.

Until recently, data on review numbers per publication were only available to editors managing their journals, who generally (rightfully) wanted to protect the anonymity of their reviewers. This meant that obtaining estimates of the number of reviews required for the peer review process was not possible. Recently, however, review verification depositories such as Publons (www.publons.com) are becoming more common. Originally, verified reviews were valued for *curriculum vitae* during job applications, as they demonstrate a researcher’s dedication to their profession and that they are valued by their peers through invitations to review for more prestigious journals. The number of reviews a scientist conducted could also be used to highlight the impact on workload to their employer by highlighting the hours that reviewing requires. Since these data are now freely available, they also offer the possibility to conduct analyses on verified reviews and review rates per contributions in various fields.

Here, I determined the “review ratio”: the expected minimum number of reviews an author should conduct in order to balance their contribution (e.g., publications) to their respective fields, using an analysis of verified review data from Publons. The aim was to provide approximate guidelines for scientists to ensure the number of reviews an author conducts correlates with their ‘burden’, through number of publications, on the system. By association the review ratio provides a threshold beyond which higher per-author review rates would reduce the stress on the peer review system.

2. Materials and Methods

Review data were obtained from Publons (www.publons.com) using the ‘Journals’ browsing tab. Publons is a free site that collates verified reviews for authors, and as a result it also collects metadata for journals that have had verified reviews. Results were sorted according to highest number of reviews in the last 12 months, from most to least. This was believed to produce more representative results, since larger numbers of verified reviews would provide larger datasets to work from, and it was believed that journals with higher numbers of verified reviews were more likely to encourage submission of results to Publons. Data were manually copied from the Publons search results onto a spreadsheet. Verified reviews cannot be directly associated with publications using these data, therefore, information on the number of publications these were likely to be associated with needed to be obtained.

In order to determine the review ratio from verified review numbers, it is also necessary to obtain the number of publications and mean number of authors for that publication. Using Web of Science, which has one of the largest collections of publications available [17], each journal with verified reviews identified in the previous step was searched individually to obtain information on the number of publications (research articles, letters and reviews) published in the last 12 months from November 2019, including those available in early access. Early access publications were included because Publons publishes review results only days after they have been submitted, which can be before publications are accepted. Citation information including author number and year for the first 500 publications (less if fewer than 500 were published in the last 12 months) was exported via text file into the Bibliometrix web graphics user interface [18]. Bibliometrix is a publication metadata package for R that can extract and rapidly analyse data from popular citation databases. A maximum of 500 citations is the limit for exporting from Web of Science in a single step, and I judged that 500 was a number of publications large enough to obtain a representative mean number of authors per publication for the journal. These 500 publications were ordered by date of publication, so it is unlikely that these sub-samples were biased. All subsequent analyses were conducted in R statistical software package V. 3.6.6 [19] through RStudio [20]. All graphs were produced using ggplot2 [21].

One of the potential sources of bias that I identified was that, since review certification is voluntary, the number of reviews submitted for each registered publication could be lower than the true number of reviews. This would lead to an underestimation of the number of reviews performed on average per publication. I assumed that a minimum number of reviews per submission was at least 2, therefore, any journals that had on average fewer than 2 reviews per recorded publication were excluded from further analyses as they indicate that not all reviews were verified for the journal. Nevertheless, due to the largely voluntary submission of review reports, it is likely that even in journals where there were more than 2 reviews per publication there are reports missing within the current Publons database. The results presented here should, therefore, generally be considered to be conservative values that are likely to increase with more widespread submission of review reports.

After cleaning the data that did not meet these criteria, this dataset contained review information from 142 journals, 359,399 verified reviews and 105,474 peer reviewed articles across 12 research fields. Publishers such as Elsevier were generally excluded from these data as their publications did not meet the above criteria, while MDPI and Wiley were generally included.

To calculate the expected reviews that an author should perform at minimum for the peer review system to function relative to their publication output, I used the function below:

$$Rr_i = \frac{nR_i/nA_i}{maA_i} \quad (1)$$

where for journal i the review ratio Rr is related to the ratio of the total number of reviews for the journal nR_i divided by the total number of articles for the journal nA_i , itself divided by the mean author number per article maA_i . The resulting value is in reviews per author per published article. The ratio was designed this way as a guideline so that an author that publishes x publications per year should review at least y publications to match their output, in their given field of research given the number of reviews usually required and the number of authors usually on publications in their field. The review ratio takes into account the mean number of authors per publication in a field, as publications with larger numbers of co-authors should require fewer reviews per author, and vice-versa.

3. Results

The mean number of reviews per article per journal across all fields was 3.59 ± 1.45 (SD), and was highest for computer science (5.50 ± 3.67) and lowest for material science (2.60 ± 0.83 ; Figure 1). Mathematics and multidisciplinary sciences had the most consistent number of reviews per publication, while computer science and engineering had a highly variable mean reviews per publication. The mean authors per publication per journal across all fields was 5.53 ± 1.92 and was highest for medicine (6.64 ± 1.53) and lowest for mathematics (2.86 ± 0.33 ; Figure 2). Physics and medicine had the most varied number of authors per publication per journal, while environmental

science and multidisciplinary journals had the most consistent number of authors. The mean review ratio across all fields was 0.74 ± 0.46 (SD) and was highest for computer science and engineering (1.52 ± 1.20) and lowest for radiology (0.54 ± 0.14 ; Figure 3). The review ratio for engineering and computer science were the most varied, while multidisciplinary and environmental sciences had the most constrained ratios.

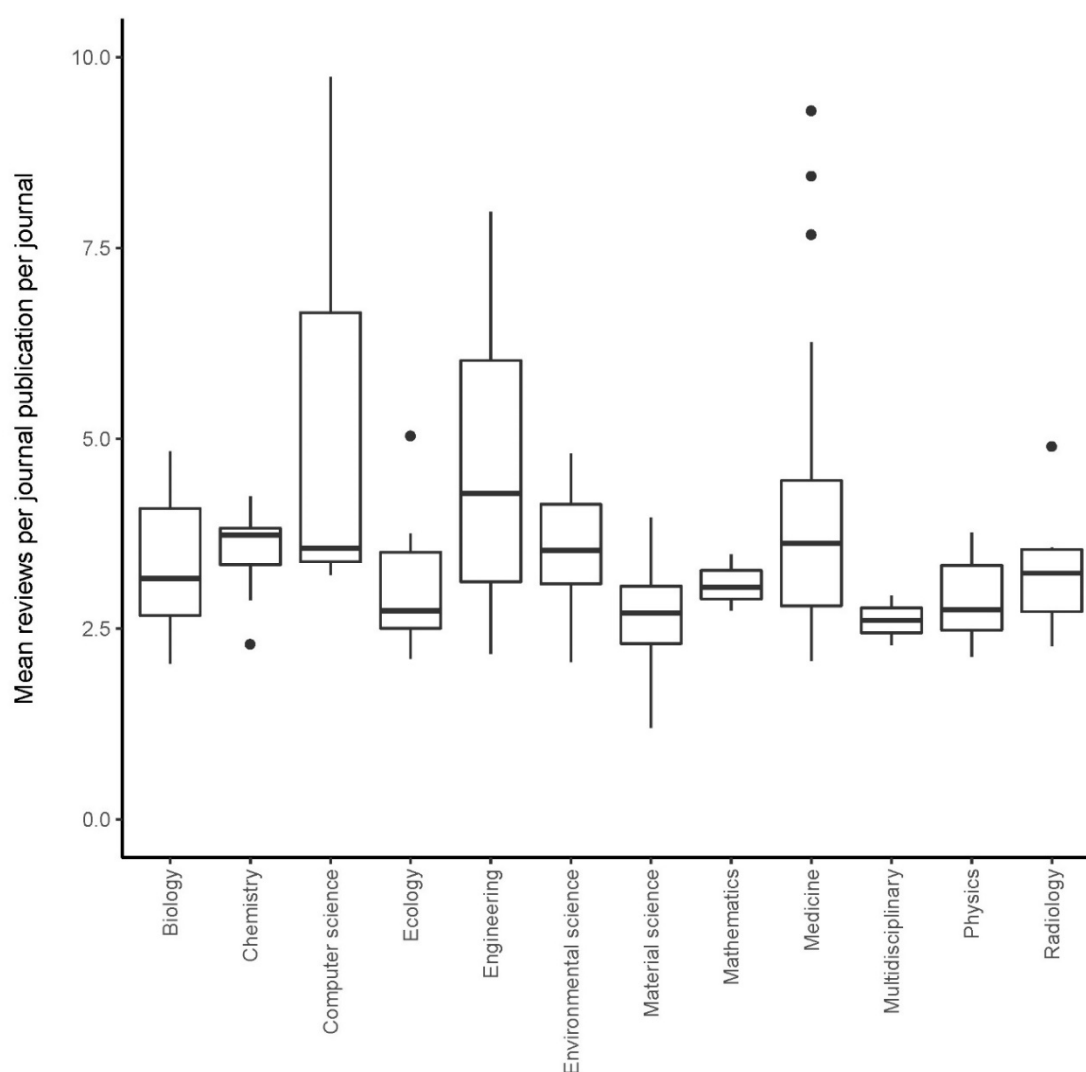


Figure 1. Boxplot of mean reviews per publication per journal across scientific research fields.

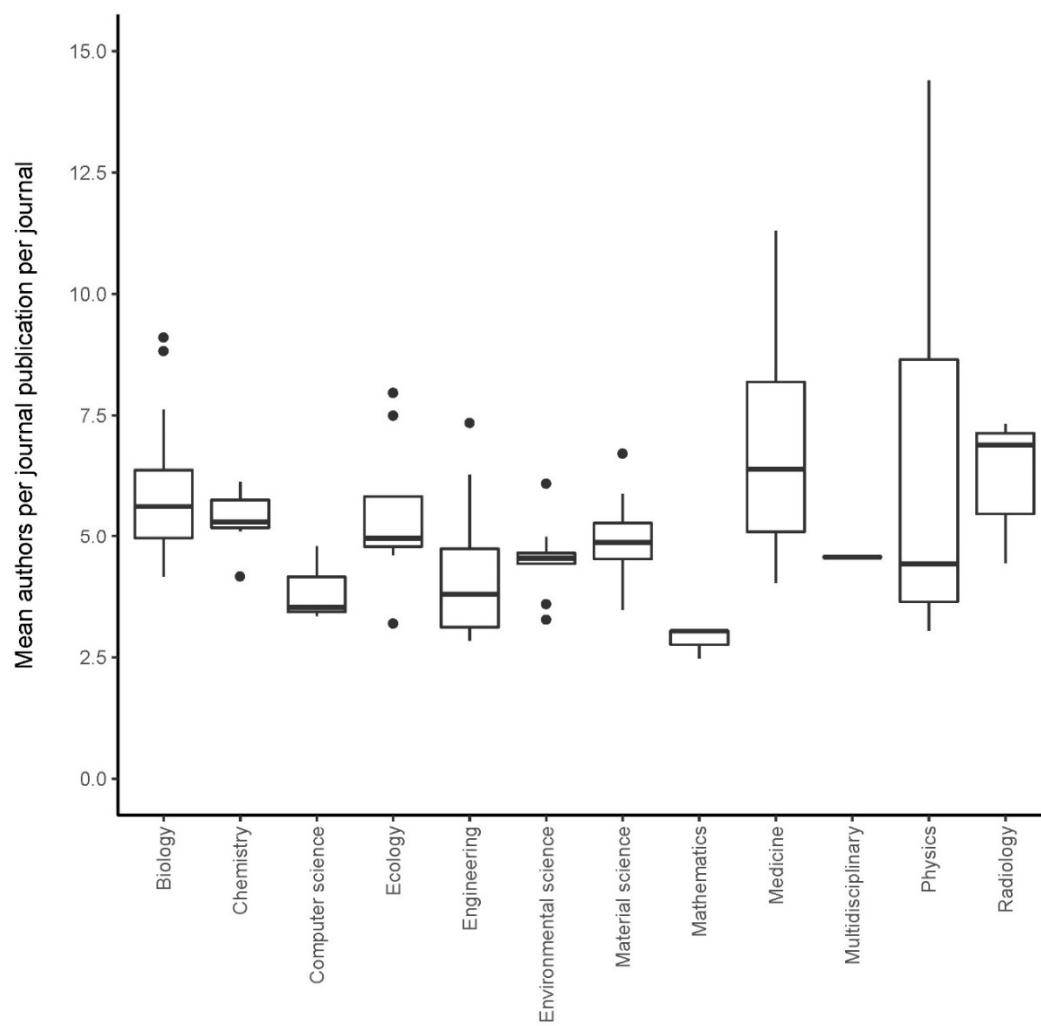


Figure 2. Boxplot of mean authors per journal publication per journal across scientific research fields.

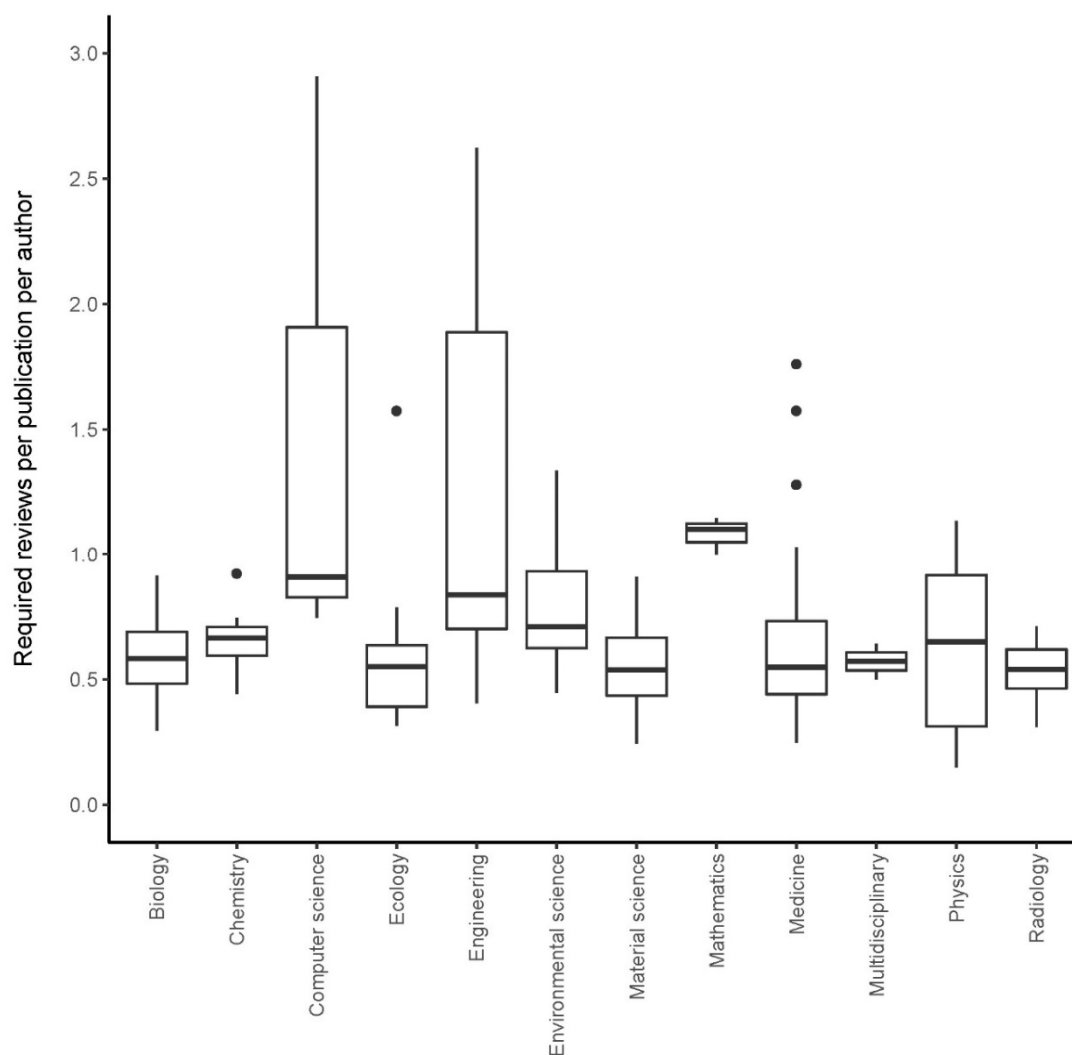


Figure 3. Boxplot of calculated required number of reviews required per author for each of their publications (review ratio) for scientific research fields.

There was a negative relationship between the mean number of authors and the review ration (Figure 4). More specifically, the review ratio for journals with lower mean numbers of authors per publication had much higher review ratios. Once the mean number of authors per publication reached ~5, the review ratios plateaued towards an asymptote of zero. Journals with mean number of authors of 10 or more were likely to have review ratios approaching zero.

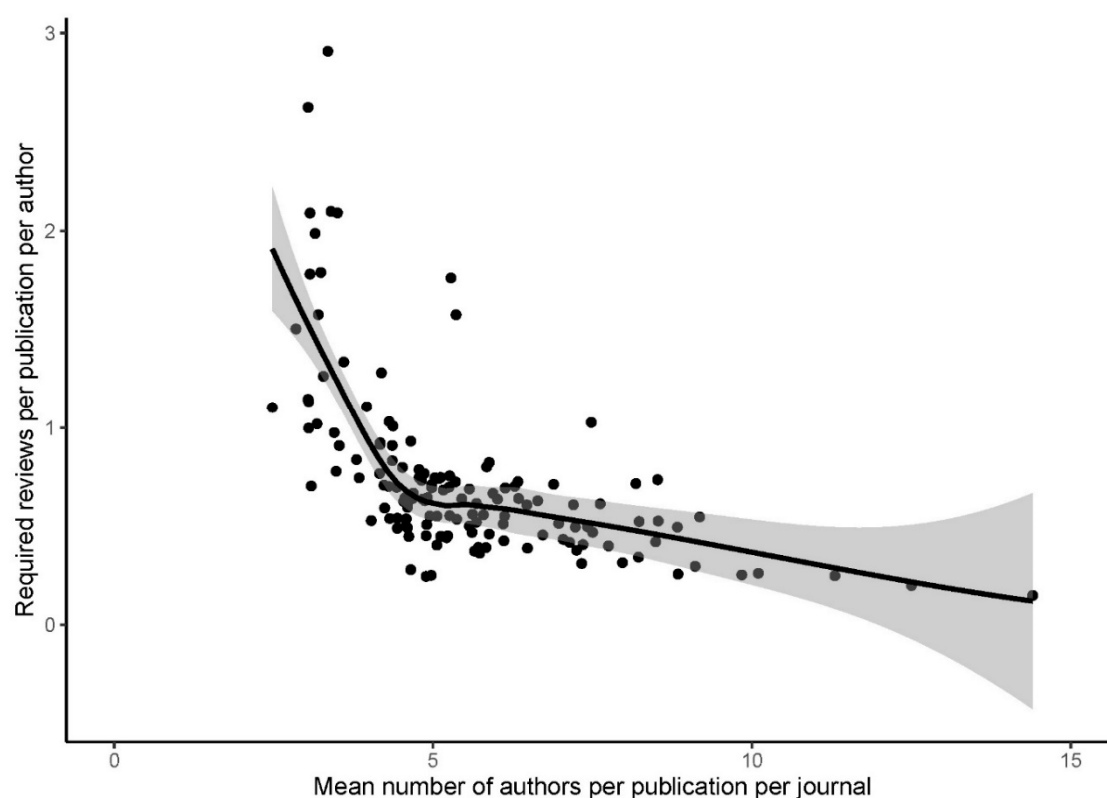


Figure 4. Relationship between mean number of authors per publication per journal and review ratio. Each point is the value obtained from a single journal ($n = 142$). Line is a LOESS (locally estimated scatterplot smoothing) curve with 95% confidence intervals (in grey).

4. Discussion

Across all fields, these results suggest authors should review at least one publication per article they produce to balance the peer review system. Reviewing rates higher than that on a broad scale may have the benefit of reducing publication times across fields by increasing the available pool of reviewers that are willing to review. Review ratios are field-dependent, however, and researchers should be aware their field may require more or fewer reviews per publication. In particular, there appears to be an inflection point for journals with approximately five mean authors per publication, above which review ratios approached zero, while lower values increased rapidly towards and intersection of approximately three per single author publications. I underline that these are conservative guidelines due to the nature of these data, and researchers should view these guidelines as minimums and aim to review more. In addition, this study does not extend to arts and humanities that are generally agreed to have few verified reviews. The usefulness of verified reviews for the wider research community for producing studies like this one should encourage more researchers to verify their reviews in databases such as Publons.

Reviews require significant time and commitment from researchers. Providing recommendations for minimum numbers of reviews per author per publication could be used as justification to increase already overloaded workloads. However, while these recommendations may seem unachievable for many, there are many very prolific reviewers. For example, over 250 reviewers in the Publons system have more than 100 verified reviews in the last year. In contrast, some of the most highly-cited researchers produce approximately 50 publications in the past year (Webometrics). It follows that it is possible for successful researchers to review more than the number required by the peer review system. Despite this, there are few reviewers available to review and a general impression of low review rates [12], a pattern that is not explained by reviewer fatigue [10]. It is widely believed that a lack of incentives to review, coupled with the high workloads of academics is responsible for high rates of declines to review [11,22]. Whether they be financial in the form of direct

payment or publication vouchers for open access journals [23], or with express allowance for reviewing in workload models [24–26], employers and publishers alike should facilitate reviews to rectify this. Alternatively, chastising authors whose review rate is much lower than the review ratio for their field is an option that has been suggested [27]. Either way, examining review ratios can allow researchers to identify whether their review rate is appropriate for their field, and if not, provide a justification to their employer or to journals to further incentivise peer review.

There are some methodological aspects to be considered when interpreting these results. Verified reviews are mostly submitted on a voluntary basis, and journals that do not have explicit partnerships with Publons are less likely to encourage verification of reviews (this is likely why journals owned by Wiley and MDPI, which have had long partnerships with Publons, met the data selection criteria while Elsevier did not). As a result, the number of verified reviews used in these calculations are likely fewer than the true number. This would by association suggest that the review ratios presented here are lower than the ‘true’ value. However, in some cases it appears that this review ratio is likely accurate, as internal reviews conducted by a medical journal recorded review rates of 4.1 reviews per publication [28], which overlaps with the values I calculated. I also only sampled author numbers per publication based off the first 500 publications for a given journal, however, since this number is fairly large, I believe it would lead to representative samples.

Not knowing how many papers academics need to review is a question that is asked frequently by early career researchers, and it is likely that many academics review a number of publications they feel is acceptable, whether that number is sufficient or not. The guidelines provided here should be used to inform them, and to improve the already stressful peer review system.

Acknowledgments: The idea for this study arose from conversations that occurred on Twitter. This highlights one of the numerous benefits of Twitter for academics and research. To #AcademicChatter, #ECRchat and the many associated senior and junior academics alike who maintain open and honest conversations about science, this one is for you.

References

1. Bornmann, L. Scientific peer review. *Annu. Rev. Inf. Sci. Technol.* **2011**, *45*, 197–245.
2. Bornmann, L.; Mutz, R. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *J. Assoc. Inf. Sci. Technol.* **2015**, *66*, 2215–2222.
3. Yankauer, A. Who are the peer reviewers and how much do they review? *Jama* **1990**, *263*, 1338–1340.
4. Ho, R.C.-M.; Mak, K.-K.; Tao, R.; Lu, Y.; Day, J.R.; Pan, F. Views on the peer review system of biomedical journals: An online survey of academics from high-ranking universities. *BMC Med Res. Methodol.* **2013**, *13*, 74.
5. Backes-Gellner, U.; Schlinghoff, A. Career incentives and publish or perish in German and US universities. *Eur. Educ.* **2010**, *42*, 26–52.
6. Herbert, D.L.; Coveney, J.; Clarke, P.; Graves, N.; Barnett, A.G. The impact of funding deadlines on personal workloads, stress and family relationships: A qualitative study of Australian researchers. *BMJ Open* **2014**, *4*, e004462.
7. Moosa, I.A. *Publish or Perish: Perceived Benefits Versus Unintended Consequences*; Edward Elgar Publishing: Cheltenham, UK, 2018.
8. Sarabipour, S.; Debat, H.J.; Emmott, E.; Burgess, S.J.; Schwessinger, B.; Hensel, Z. On the value of preprints: An early career researcher perspective. *PLoS Biol.* **2019**, *17*, e3000151.
9. McCook, A. Is peer review broken? Submissions are up, reviewers are overtaxed, and authors are lodging complaint after complaint about the process at top-tier journals. What's wrong with peer review? *Scientist* **2006**, *20*, 26–35.
10. Fox, C.W.; Albert, A.Y.; Vines, T.H. Recruitment of reviewers is becoming harder at some journals: A test of the influence of reviewer fatigue at six journals in ecology and evolution. *Res. Integr. Peer Rev.* **2017**, *2*, 3.
11. Tite, L.; Schroter, S. Why do peer reviewers decline to review? A survey. *J. Epidemiol. Community Health* **2007**, *61*, 9–12.
12. Nguyen, V.M.; Haddaway, N.R.; Gutowsky, L.F.; Wilson, A.D.; Gallagher, A.J.; Donaldson, M.R.; Hammerschlag, N.; Cooke, S.J. How long is too long in contemporary peer review? Perspectives from authors publishing in conservation biology journals. *PLoS ONE* **2015**, *10*, e0132557.

13. Cho, Y.H.; Cho, K. Peer reviewers learn from giving comments. *Instr. Sci.* **2011**, *39*, 629–643.
14. Arns, M. Open access is tiring out peer reviewers. *Nat. News* **2014**, *515*, 467.
15. Kassirer, J.P.; Campion, E.W. Peer review: Crude and understudied, but indispensable. *Jama* **1994**, *272*, 96–97.
16. Felicitas Domínguez-Berjón, M.; Godoy, P.; Ruano-Ravina, A.; Negrín, M.Á.; Vives-Cases, C.; Álvarez-Dardet, C.; Bermúdez-Tamayo, C.; López, M.J.; Pérez, G.; Borrell, C. Acceptance or decline of requests to review manuscripts: A gender-based approach from a public health journal. *Account. Res.* **2018**, *25*, 94–108.
17. Waltman, L.; Noyons, E. Bibliometric and research management and research evaluation. In *A Brief Introduction*; Leiden Universiteit: Leiden, The Netherlands, 2018.
18. Aria, M.; Cuccurullo, C. Bibliometrix: An r-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975.
19. Team, R.C. Development Core Team. R: A language and environment for statistical computing. In *R Foundation Statistical Computing*; Vienna Austria: Vienna, Austria, 2013.
20. Allaire, J. *Rstudio: Integrated Development for R*; RStudio, Inc.: Boston, MA, USA, 2015.
21. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2009.
22. Willis, M. Why do peer reviewers decline to review manuscripts? A study of reviewer invitation responses. *Learn. Publ.* **2016**, *29*, 5–7.
23. Pasternak, J.; Glina, S. Paying reviewers for scientific papers and ethical committees. *Einstein (Sao Paulo)* **2014**, *12*, 7–15.
24. Jonnalagadda, S.; Petitti, D. A new iterative method to reduce workload in the systematic review process. *Int. J. Comput. Biol. Drug Des.* **2013**, *6*, 5.
25. Miwa, M.; Thomas, J.; O'Mara-Eves, A.; Ananiadou, S. Reducing systematic review workload through certainty-based screening. *J. Biomed. Inf.* **2014**, *51*, 242–253.
26. Warne, V. Rewarding reviewers—sense or sensibility? A wiley study explained. *Learn. Publ.* **2016**, *29*, 41–50.
27. Hauser, M.; Fehr, E. An incentive solution to the peer review problem. *PLoS Biol.* **2007**, *5*, e107.
28. Bordage, G. Reasons reviewers reject and accept manuscripts: The strengths and weaknesses in medical education reports. *Acad. Med.* **2001**, *76*, 889–896.

