

Are scientific editors reliable gatekeepers of the publication process?



ABSTRACT

Editors are gatekeepers of the scientific publication process and many manuscripts never make it past them. For the journal *Biological Conservation*, editors reject about half of the manuscripts assigned to them (known as “desk reject”) before sending the remaining manuscripts out for review. Given the importance of the editor's role in the publication process we ask: how consistent are editors' decisions in *Biological Conservation*? To answer this question, in 2018 we asked 10 editors of the journal to evaluate 40 manuscripts that had been previously evaluated in 2017 as part of the regular review process. The editors did not know which manuscripts had been sent out for review (20 total) and which had been desk rejected (20 total) in 2017. We found that in 2018, agreement among editors was reasonable in their decisions, even when papers that editors were asked to evaluate fell outside their area of expertise, and decisions did not differ with editorial experience. Of the 40 manuscripts, at least seven editors agreed on the decisions for 73% (29) of them. On average, for the 20 manuscripts sent out for review in 2017, 70% of editors in 2018 agreed with the decisions to review them. For the 20 manuscripts that were desk rejected in 2017, on average 67% of editors in 2018 agreed with the decisions to desk reject them. While variation in editors' decisions on manuscripts exists, most decisions are repeatable and reliable. The journal can potentially improve the reliability of the review process by encouraging discussion among editors and by providing more detailed guidelines to editors. We urge other journals to carry out similar evaluations of this important desk rejection process.

1. Introduction

The goal of the peer-review process of journals is to make well-founded and consistent decisions and publish high-quality science matched to a journal's subject area. The review process usually involves an initial decision by an editor to desk reject or send out a manuscript for review, followed by two or three anonymous reviewers assessing the manuscript and a final decision made by an editor. Many parts of the review process have been well-studied to identify their strengths and shortcomings (Hojat et al., 2003; Smith, 2006; Resnik and Elmore, 2016). Researchers have examined reviewer biases and the extent to which different models of review might benefit the process (Primack et al., 2009; Darling, 2015; Tomkins et al., 2017; Helmer et al., 2017). However, few studies have considered editors' initial decisions on manuscripts before they go out for review, a step when many, and often most, manuscripts are rejected.

Editors desk reject manuscripts, acting as “gatekeepers” to the journal, if the subject material is poorly suited to their journals, the studies are not sufficiently rigorous, or the topics are not sufficiently novel (Hojat et al., 2003; Primack, 2009; Arnqvist, 2013; Traniello and Bakker, 2014). Editors appreciate the efficiency of desk rejections because many manuscripts will not pass the review process. Desk rejections can save reviewers unnecessary work and editors the trouble of finding reviewers, and provide authors with more rapid decisions. However, desk rejections often frustrate authors because they lack feedback from peer reviewers that can help authors improve their manuscripts (Primack, 2009). These initial desk rejection decisions made by one person may appear to be inconsistent and unfair to authors. Some authors assert that the initial decisions on whether to review a manuscript may depend heavily on the arbitrary decision of an individual editor.

At *Biological Conservation* we are particularly interested in

understanding the dynamics of our peer review process, including consistency in decisions and potential sources of bias. In another study, for example, we have found that the ratings given to the same manuscript by different reviewers tend to be correlated more than expected by chance, with reviewers agreeing on recommendations for about 36% of manuscripts (relative to an expectation of 25% agreement by chance; possible recommendations include accept, minor revision, major revision, and reject) (Campos-Arceiz et al., 2015), although this is not the case at all journals (e.g., Kravitz et al., 2010). Further, the *Biological Conservation* review process shows no detectable gender bias, but shows great variation in acceptance rates for manuscripts submitted by authors from different countries (Primack and Zipf, 2016). As might be expected, editors also appear to strongly consider reviewers' evaluations when making decisions on whether or not to accept a manuscript for publication (Campos-Arceiz et al., 2015). We therefore wanted to explore further the role of the editors in other aspects of the editorial process and assess the consistency of editors as the gatekeepers of our journal (Fig. 1). And as far as we know, this is the first such evaluation of the desk rejection process, despite its importance in the review process. Such an evaluation is an important part of making the editorial process more transparent.

Of the manuscripts submitted to *Biological Conservation*, about 20% are immediately rejected by the Editor-in-Chief (EiC) because they are clearly out of the scope of the journal. For the remaining 80%, the EiC assigns an average of 8–12 manuscripts per month, or around 120–150 manuscripts a year, to each Associate Editor. Associate Editors represent a particular expertise and geographic region. Each Associate Editor, in turn, desk rejects 50–60% of the manuscripts they are assigned and sends out the remainder for peer review. At *Biological Conservation*, Associate Editors accept roughly 20% of the manuscripts they are assigned. Editors generally regard the rejection of many manuscripts prior to sending them out for review as an efficient and



Fig. 1. The *Biological Conservation* team often acts as an editorial gatekeeper, as shown in this 2018 photo, from left to right, in front of an actual gate: Danielle Descoteaux (USA, publisher), Richard Primack (USA, Associate Editor (AE)), Robin Pakeman (UK, AE), Tracey Regan (Australia, AE), Vincent Devictor (France, Editor in Chief), Richard Corlett (China, AE), Liba Pejchar (USA, AE), Bea Maas (Austria, AE), David Johns (Book Review Editor).

effective way to begin the filtering process.

In this study, we ask: are initial decisions on manuscripts made by editors at *Biological Conservation* consistent? Are they acting as reasonable and consistent gatekeepers of the editorial process? We investigate this via five questions:

1. Are decisions to reject or review a manuscript consistent among the team of editors in 2018 (i.e. inter-editor consistency)?
2. How does the original editor decision made in 2017 compare with decisions in 2018 (i.e. temporal consistency)?
3. Are individual editors consistent when assessing the same manuscript in 2017 and 2018 (i.e. intra-editor consistency, in cases when they do not remember previously assessing the manuscript)?
4. Does agreement among editors depend on the level of experience or expertise of the editor?
5. What are the major reasons that editors provide when deciding to either send a manuscript out for review or to desk reject it?

2. Materials and methods

Ten current editors (among 12) of *Biological Conservation* agreed to participate in the study. Of the ten editors, five are experienced and have worked as editors for the journal for at least 3 years each, and five are new to journal editing and have worked at *Biological Conservation* for fewer than 3 years. An editorial staff member selected 40 manuscripts that were submitted to the journal in 2017 and had already been dealt with by a handling editor. Of these 40 manuscripts selected, 20 were desk rejected by the original handling editor in 2017 and 20 were sent out for review. The manuscripts were submitted by authors from a range of countries (among the corresponding authors: 18 different countries), covered a variety of topics and article types (29 research articles, 7 perspective articles, 2 review articles, and 2 short communications) and different taxonomic groups within conservation biology, and included a mixture of men and women authors (among the corresponding authors are 24 men and 16 women). All 40 manuscripts had

already been filtered by the EiC, reflecting that they were all considered a reasonable fit with the aims of the journal and of reasonable quality.

The 10 handling editors were asked to assess each of the 40 manuscripts as they would typically evaluate manuscripts and assign a category of either desk rejection or review. The editors were not given any information on the proportion of manuscripts originally rejected versus sent out for review. Handling editors were asked not to evaluate manuscripts they remembered handling or seeing in print (Editors handle 120–150 manuscripts each year, so it is common that they do not remember individual manuscripts). Each handling editor also noted whether they had expertise in the topic of the manuscript and were asked to briefly describe why they chose to desk reject or send out for review each manuscript. We then used key words and phrases used by the editors to classify each decision in one of ten categories, five representing desk reject justifications and five review justifications.

We used a Fleiss Kappa statistic to measure the agreement among editors given their decisions on manuscripts in 2018, where a value of 1 is perfect agreement and 0 is the agreement expected by chance (Landis and Koch, 1977). Only manuscripts that all 10 editors assessed were included in this statistical analysis. Intra-editor consistency was examined by calculating the proportion of manuscripts assessed in 2017 and 2018 by the same editor where the decision to reject or review was the same in both years. To examine if agreement depends on level of experience, we split the data into two groups according to whether an editor had > 3 years' experience or not and recalculated the Fleiss Kappa Statistics for each group. To examine whether agreement depends on an editor's reported expertise we calculated for each manuscript the proportion of editors that decided reject versus review and whether the editors reported that they had expertise in the subject area and calculated summary statistics. Lastly, we compared the agreement among editors' decisions in 2018 with the actual decisions in 2017.

3. Results

Of the 400 total decisions (40 manuscripts × 10 handling editors)

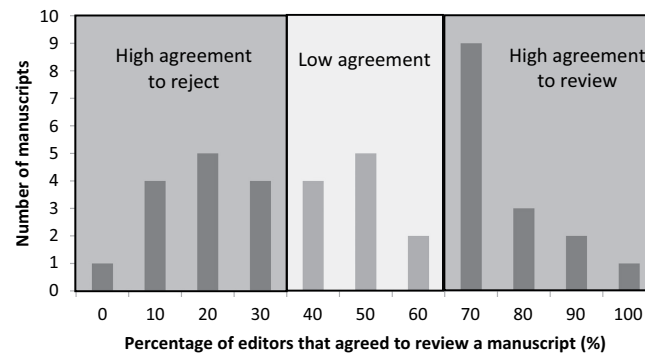


Fig. 2. Frequency of agreement among editors to send to manuscripts for review (as a %). Dark shaded areas represent high agreement (70–100% agreement) to either reject or review a manuscript. Light shaded areas represent low agreement (40–60% agreement) among editors.

in 2018, handling editors' decisions were split evenly between assigning manuscripts to be reviewed and rejecting them without review (desk rejection); individual editors rated between 42 and 55% of manuscripts as appropriate for review. In 14 cases (< 4% of total decisions), editors did not assess manuscripts because they recalled handling a manuscript in 2017 or seeing it in print, resulting in 29 manuscripts being assessed by all 10 editors. The even split between manuscripts to be sent out for review or rejected, is consistent with the original decisions for these manuscripts in 2017, when 50% were sent out for review and 50% rejected without review.

Complete agreement among editors was rare; all 10 editors agreed on only two manuscripts (one reject and one review). However, for 73% (29) of manuscripts, at least 7 of the 10 editors agreed on a decision (whether to review or desk reject). Specifically, for 14 manuscripts at least 7 editors agreed the manuscripts should be rejected, and for 15 manuscripts at least 7 editors agreed the manuscripts should be reviewed. For 11 manuscripts editors were evenly split between recommending desk rejection or review (Fig. 2).

Using Kappa statistics, agreement among editors was moderate ranging from 0.209 to 0.232, values that are statistically significantly different than zero (which would reflect no agreement). Agreement among editors did not depend on the number of years of experience (Table 1); that is, more experienced editors were not more likely to make similar decisions.

In 2018, editors tended to recommend decisions that matched the actual 2017 decisions (Table 2). Of the manuscripts rejected in 2017, on average 70% of the editors agreed it should be rejected in 2018. Similarly, of the manuscripts sent out to review, on average 67% of editors agreed with the 2017 decision (Table 2).

Although all editors combined rated 51% of the manuscripts as within their area of expertise, there was variation between individual editors. For example, one editor rated only 6% of manuscripts as in their area of expertise, whereas another editor rated 78% as in their area of expertise. Five handling editors rated a higher percentage of manuscripts in their area of expertise as appropriate to send out for review in comparison with manuscripts outside their area of expertise, but the other five editors had the opposite tendency. An editor's stated expertise on the topic of a manuscript had no apparent influence on the editor's decision with mean number of editors agreeing to review and reject based on expertise being indistinguishable (mean range

Table 2

The average number of editors that agreed in 2018 with the original decisions for each manuscript in 2017. For example, for papers that were desk rejected in 2017, an average to 7 editors agreed that the paper should be rejected in 2018 and 3 editors thought that the article should be reviewed.

		2018 decision	
		Reject	Review
2017 decision	Reject	7	3
	Review	3	6

Table 3

The mean number of 10 editors agreeing to reject and review manuscripts based on their individual expertise, using the 29 papers reviewed by all 10 editors.

	Reject and expert	Reject and non-expert	Review and expert	Review and non-expert
Mean	2.69	2.69	2.72	1.90
Standard dev	1.89	2.22	2.25	1.65

Table 4

Justifications editors gave for their decisions. NA reflects cases in which reviewers did not evaluate manuscripts in 2018 because they remembered evaluating them in 2017. The number of comments is based on 10 editors each evaluating 40 papers.

Editor decision	Justification	Number of comments
Desk reject	Too specific	67
	No conservation implications	44
	Subject material not appropriate for <i>Biological Conservation</i>	41
	Manuscript not of sufficient quality	28
	Not novel	25
NA		14
Review	Broad conservation implications	69
	Interesting perspective	45
	Novel	38
	Topical; dealing with current issues	21
	Under-studied system	8

Table 1

Kappa statistics for multi-rater agreement with all editors combined and then grouped according to experience.

	All editors	Editors > 3 years' experience	Editors < 3 years' experience
Number of editors	10	5	5
Number of manuscripts	29	29	29
Fleiss Kappa statistic	0.232	0.209	0.217
p-Value	< 0.0000	< 0.0004	< 0.0002

1.90–2.72) (Table 3).

The major reasons given for deciding to send manuscripts out for review were that the manuscripts had broad conservation implications, had interesting perspectives, or were novel (Table 4). The main reasons given for deciding to reject a manuscript were the manuscript had mainly local implications and was not appropriate for an international journal, the manuscript did not have strong conservation implications, or the subject material was not appropriate in some other way for *Biological Conservation* (Table 4).

4. Discussion

Our results demonstrate that handling editors make relatively consistent decisions when evaluating manuscripts submitted to the journal *Biological Conservation*. Our editors tend to make decisions consistent with each other and tend to make consistent decisions when faced with the same manuscript a second time (when they do not remember handling it the first time). Although our system of evaluating manuscripts is fairly reliable, there still is variation among editors and some degree of uncertainty in whether a manuscript is desk rejected or sent out for review.

The novelty or importance of a given manuscript can sometimes be perceived differently by the same editor when evaluated in a later year. We found that editors changed their minds on decisions for 33% of the manuscripts they reviewed in both 2017 and 2018. Part of this fluctuation is likely caused by manuscripts that generate borderline decisions (i.e., fall in the low agreement area of Fig. 2), and part can be attributed to changing perceptions of what constitutes novel and high-quality science or the possibility of editor is unconsciously influenced by having reviewed the paper previously. We do not necessarily expect a given manuscript to be assessed similarly by the same editor from year to year. Indeed, the vision each editor has about the novelty and importance of a given manuscript is influenced by several factors, including what was already received and published by the journal and other related journals in the field.

One problem with the current system at our journal, and most other scientific journals, is that most decisions on whether to desk reject manuscripts or send them out for review are made by just one person acting as the gatekeeper. As we have seen, there are many manuscripts in which editors disagree on the scientific value of the manuscript, and decisions to reject or review manuscripts are a matter of chance for such manuscripts. A different system, which is already in place at some journals, would involve two or three editors evaluating each manuscript independently, and only making a final decision if there was agreement between the editors. While such a system might be an improvement in terms of improving the consistency of decisions, it would add considerably to the workload of editors and slow the handling time for each manuscript, a topic that already frustrates many authors (Powell, 2016). A compromise might involve a handling editor assigning manuscripts to one of three categories—reject, review, and borderline—and asking input from other editors for the borderline manuscripts. Some editors already ask the EiC for an additional opinion on specific papers. A more systematic and collective assessment on borderline papers could be an effective option.

Presubmission inquiries (asking the EiC about the suitability of a paper based on title and abstract before formal submission of a full paper) is also an important tool for authors ensure a good fit with a journal. In this case, the EiC often asks other editors their opinion. Authors are then rapidly provided with feedback and guidelines for their paper, and an indication if the paper is suitable for the journal. Moreover, although a desk rejection almost always comes with a justification, editors can consider an appeal and change their decision if an author has a strong and convincing argument. In this case, an editor sometimes seeks the advice of other editors and researchers.

The editorial process could also be improved by having more explicit guidelines for editors and authors about the standards a

manuscript has to meet for it to be sent out for review. These guidelines, often presented as a Guide to Authors, should be reviewed periodically, perhaps once a year, to make sure that they are current, and editors are using them appropriately. *Biological Conservation* has such guidelines on its website, and new editors are already carefully instructed in how to apply them in practice when reviewing papers.

Although we asked editors to use their regular methods to evaluate these manuscripts, in fact editors rated about half of these manuscripts as outside their area of expertise. This is a higher percentage of such manuscripts than editors would normally handle, as the EiC mostly assigns editors manuscripts in their areas of expertise. The agreement among editors is thus even more heartening—editors appear to evaluate these manuscripts based on inherent quality or objective criteria, rather than relying on their own specialized knowledge. Editors also have about the same average rate of acceptance for manuscripts within and outside their areas of expertise; there is no tendency for editors to reject manuscripts outside their specialized areas.

Overall, we found that editors are reasonably good gatekeepers of the journal review process, even when papers are outside of their area of expertise. Editors in 2018 mostly agreed with decisions made by 2017 editors, and agreed among themselves in deciding on desk rejection or sending manuscripts out for review. Hopefully these decisions have been made on the basis of scientific merit and suitability of the manuscripts for the journal. It is possible that other factors might have affect decisions to desk reject manuscripts, such as biases against certain types of research, certain countries, or studies without significant results, as well as biases in favor of review articles (Hojat et al., 2003; Primack et al., 2009). Our journal takes such considerations seriously, and we make efforts to publish special issues in new areas of research and to encourage authors from countries that are under-represented in our journal. However, based on the results of this study, we believe editors at *Biological Conservation* are applying a reasonably consistent and reliable standard of evaluation in the editorial process. We urge other journals to carry out comparable studies of their own desk rejection process.

Declaration of competing interest

The authors have no conflicts of interest.

Acknowledgements

We thank Danielle Descoteaux of Elsevier for encouraging us to carry out this study, and we thank Abe Miller-Rushing for helpful comments on the manuscript.

References

- Arnqvist, G., 2013. Editorial rejects? Novelty, schnovelty!. *Trends Ecol. Evol.* 28, 448–449. <https://doi.org/10.1016/j.tree.2013.05.007>.
- Campos-Arceiz, A., Primack, R.B., Koh, L.P., 2015. Reviewer recommendations and editors' decisions for a conservation journal: is it just a crapshoot? And do Chinese authors get a fair shot? *Biol. Conserv.* <https://doi.org/10.1016/j.biocon.2015.02.025>.
- Darling, E.S., 2015. Use of double-blind peer review to increase author diversity. *Conserv. Biol.* <https://doi.org/10.1111/cobi.12333>.
- Helmer, M., Schottdorf, M., Neef, A., Battaglia, D., 2017. Gender bias in scholarly peer review. *Elife*. <https://doi.org/10.7554/elifelife.21718>.
- Hojat, M., Gonnella, J.S., Caelleigh, A.S., 2003. Impartial judgment by the “gatekeepers” of science: fallibility and accountability in the peer review process. *Adv. Health Sci. Educ. Theory Pract.* <https://doi.org/10.1023/A:1022670432373>.
- Kravitz, R.L., Franks, P., Feldman, M.D., Gerrity, M., Byrne, C., Tierney, W.M., 2010. Editorial peer reviewers' recommendations at a general medical journal: are they reliable and do editors care? *PLoS One*. <https://doi.org/10.1371/journal.pone.0010072>.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174. <https://doi.org/10.2307/2529310>.
- Powell, K., 2016. Does it take too long to publish research? *Nature*. <https://doi.org/10.1038/530148a>.
- Primack, R.B., 2009. Why did we reject your paper? *Biol. Conserv.* <https://doi.org/10.1016/j.biocon.2009.04.027>.

- Primack, R.B., Ellwood, E., Miller-Rushing, A.J., Marrs, R., Mulligan, A., 2009. Do gender, nationality, or academic age affect review decisions? An analysis of submissions to the journal biological conservation. *Biol. Conserv.* <https://doi.org/10.1016/j.biocon.2009.06.021>.
- Primack, R.B., Zipf, L., 2016. The United States: trends or noise? *Biol. Conserv.* 196. <https://doi.org/10.1016/j.biocon.2016.01.01>.
- Resnik, D.B., Elmore, S.A., 2016. Ensuring the quality, fairness, and integrity of journal peer review: a possible role of editors. *Sci. Eng. Ethics.* <https://doi.org/10.1007/s11948-015-9625-5>.
- Smith, R., 2006. Peer review: a flawed process at the heart of science and journals. *J. R. Soc. Med.* <https://doi.org/10.1258/jrsm.99.4.178>.
- Tomkins, A., Zhang, M., Heavlin, W.D., 2017. Reviewer bias in single- versus double-blind peer review. *Proc. Natl. Acad. Sci.* <https://doi.org/10.1073/pnas.1707323114>.
- Traniello, J.F.A., Bakker, T.C.M., 2014. Editorial fairness: speed, accuracy and awareness benefit authors. *Trends Ecol. Evol.* <https://doi.org/10.1016/j.tree.2013.10.011>.

Richard B. Primack^{a,*}, Tracey J. Regan^b, Vincent Devictor^c, Lucy Zipf^a, Laurent Godet^d, Rafael Loyola^{e,f}, Bea Maas^g, Robin J. Pakeman^h, Graeme S. Cummingⁱ, Amanda E. Bates^j, Liba Pejchar^k, Lian Pin Koh^l

^a *Biology Department, Boston University, 5 Cummington Mall, Boston, MA 02215, USA*

^b *Arthur Rylah Institute for Environmental Research, Department of Environment, Land Water and Planning, 123 Brown Street, Heidelberg,*

Victoria 3084, Australia

^c *ISEM, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France*

^d *CNRS, Université de Nantes, UMR LETG, B.P. 81223, 44312 Nantes Cedex 3, France*

^e *Fundação Brasileira para o Desenvolvimento Sustentável, Rio de Janeiro, Brazil*

^f *Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, Brazil*

^g *Department of Botany and Biodiversity Research, Division of Tropical Ecology and Animal Biodiversity, University of Vienna, Rennweg 14, 1030 Vienna, Austria*

^h *The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, UK*

ⁱ *ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville 4811, Australia*

^j *Department of Ocean Sciences, Memorial University of Newfoundland, St. John's 3ZH, Canada*

^k *Department of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA*

^l *Conservation International, Arlington, VA 22202, USA*

E-mail address: primack@bu.edu (R.B. Primack).

* Corresponding author at: Department Boston University, 5 Cummington Mall, Boston, MA 02215, USA.