Web Ecology 10: 23–26.

AGORA

AGORA is a lighter channel of communication between readers and contributors; it aims to stimulate discussion and debate, particularly by presenting new ideas and by suggesting alternative interpretations to the more formal research papers published in WEB ECOLOGY and elsewhere. A lighter prose is encouraged and no summary is required. Formal research papers, however short, will not be considered.

Judging the quality of our research: a self-assessment test

Lonnie W. Aarssen (aarssenl@queensu.ca), Dept of Biology, Queen's Univ., Kingston, ON, Canada, K7L 3N6. – Christopher J. Lortie, Dept of Biology, York Univ., 4700 Keele St, Toronto, ON, Canada, M6S 2E2. – Amber E. Budden, National Center for Ecological Analysis and Synthesis (NCEAS), 735 State Street, Suite 300, Santa Barbara, CA 93101, USA.

Most researchers are interested in submitting their work to the very best possible peer-review journals, usually regarded as those with high 'impact factors'. But how does one know whether the quality of one's work is adequate to compete for the limited page space available within these premium journals? Would a lower-ranking journal be more appropriate? How low? The costs of choosing the 'wrong' journal in this case can be substantial: the wasted time and effort – for both authors as well as referees and editors – that results from a series of rejections when an author repeatedly 'aims too high' (Hochberg et al. 2009); and also the loss of deserved recognition when an author 'aims too low'. It should be possible to minimize these costs to some extent by using a set of core criteria for self-assessment of research quality prior to submission for publication.

With empirical research, the obvious place to begin is with the data, where quality is relatively easy to assess based on widely recognized conventional standards for effective methodology, effective statistical analyses, and effective presentation of the data for the reader (e.g. see Møller and Jennions 2002, Murtaugh 2002, 2007). Similarly, in theoretical research, it is relatively intuitive to assess whether ideas, hypothesis development, and conceptual reasoning are at least plausible, and whether models are structurally logical and computationally correct. Although there is certainly variation, even the lowest ranking journals tend not to publish poor quality data, implausible ideas, or fallacious models – at least not journals that are likely to stay accredited. Hence, if the data are not good, the model is flawed, or the idea is absurd to the intended audience, there is obviously little point in seeking peer-review publication.

Additional criteria, therefore, are often more important in distinguishing relative quality among those studies that have at least some suspected potential for publication. It is these criteria that we address here. For empirical studies, high quality data are of course essential, but not sufficient to define high quality *research.* The latter, most importantly, must also convince the reader that the data are interesting and important in ways that have potential to significantly advance our understanding of the subject matter. Similarly, for theoretical studies, high quality research requires ideas or modeling that are not just plausible, but also inspirational and instructive to the extent that they provide significant guidance and inference for future study and/or further progress in theory development. In truly outstanding research, this potential for advancement may be evident even beyond the immediate subject matter of the study.

In the published literature, therefore, research quality is usually rated by the size of the publication audience or readership that has been inspired or otherwise affected by the work, as judged by citation rate. For papers under review then, quality is commonly judged by their anticipat-

Accepted 28 February 2010

ISSN 1399-1183

ed citation rate, and in some cases (from an editor's view) perhaps also by their potential for maintaining or elevating the impact factor of the journal. It is primarily in this regard that wide variation exists in the 'relative perceived merit' (Lortie et al. 2007) of published studies. Very highly regarded studies commonly end up in the relatively few top-ranking journals, leaving the much larger percentage of apparently less profound studies mostly relegated to a broadly skewed bulge of lower-grade journals (Fig. 1).

Here, we suggest that there are core determinants of variation in research quality that an author can use for self-assessment, before deciding where to submit a paper, and even before embarking on a research program.

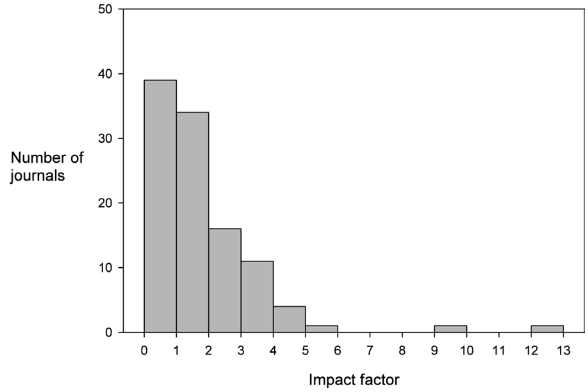

Figure 1. Frequency distribution of journal impact factors from 2004 for 107 journals listed in the 'Ecology' category by ISI Web of Science (<www.isiwebofknowledge.com>).

A self-assessment test

Based on our experiences as editors and reviewers, we suggest six criteria that might be used as a base-line guide in assessing how significantly one's research, or research proposal, is likely to be regarded in the peer-review process. These criteria involve qualitative aspects of the study relative to the published literature, particularly with regard to the extent that the study has potential to significantly advance our understanding of the field of study that is represented by the research. The primary objectives of such a test are to achieve self-identification of the relative merit of one's work, to streamline the review process by minimizing 'aiming too high' behavior, to help guard against unintended 'aiming too low' behaviour, and thus to encourage more direct and efficient reporting and publishing in ecology.

We encourage authors, therefore, to consider whether the following statements apply to their research:

> (1) *My research asks questions that have already been addressed several times in previous studies, e.g. using different species or different habitat types, and the study provides no basis for predicting, at the outset, that there was potential for the results to deviate in an interesting way from these earlier publications.*

> (2) *My research tests a hypothesis that is already widely accepted as part of established theory, and the study provides no basis for predicting, at the outset, that the results might potentially conflict in an interesting way with this theory.*

If neither (1) nor (2) applies to your study, then the research may have a good chance of being considered high quality. However, if either statement applies, then the study will be of interest to a relatively narrow audience, with a relatively limited scope of potential contribution to the literature. Research quality, therefore, will probably be regarded as moderate at best, even if the experiment was well designed and yielded conclusive results. Such studies may add additional examples to an already substantial body of supporting data (and so are often publishable somewhere), but unless the new example is uniquely interesting or remarkable in some way (e.g. because the study species or habitat has recently become rare and threatened with extinction), then these studies are largely just confirmatory. Replications of previous studies often have some publishable value, particularly when they have local applied implications, but they don't usually lead to any real advance in our general understanding of the subject area.

If your research fares well after the above considerations, then check to see if the following statements apply:

> (3) *My research investigates a question or hypothesis using methodology that, if conducted as prescribed, is guaranteed, in advance, to answer the question in a particular way, or is guaranteed, in advance, to support the hypothesis. In other words, if these 'expected' results had not been obtained, then the only plausible explanation would be that something must have gone wrong with the methodology.*

> (4) *The results of my research support the stated hypothesis or prediction, but if instead the results had been 'negative' (i.e. not supporting the hypothesis or prediction), then this study would not have been sufficiently interesting or informative to be publishable.*

If neither (3) nor (4) applies to your study, then the research retains high quality potential. However, if either statement applies, then the research quality rating will probably be more modest. Such a study allows little or no possibility for discovery of anything new. Typically in this type of research there is only one interesting or acceptable possible answer – the one that is already established in the mind of the researcher, and therefore the one that the methodology is designed to produce. When the data are inconsistent with this answer, there is no alternative but to reject the data altogether. It is not possible to revise the 'acceptable' answer *post-hoc*, because in a study like this, all other answers were literally implausible *a-priori* since the study was designed to regard them that way. A 'negative result' then from such a study would be irrelevant and unpublishable. When this is apparent to a referee or an editor, there is a significant risk that the research – even if it yielded a 'positive result' – may be regarded as largely trivial or redundant, serving no real purpose other than to demonstrate the obvious or the inevitable. "When scientific facts come in they rarely conform exactly to our expectations; if they did, we would not have to do science in the first place" (Pinker 2002, p. 281). In the highest-quality research, studies are designed so that the results – provided that nothing 'went wrong' – are interesting and informative regardless of what answers are produced, or potentially produced, and regardless of whether the hypothesis is accepted or rejected. (In some exceptional cases, however, an important discovery might be made about the methodology itself – regarding how/why it 'went wrong', or turned out to be ineffective – thus providing valuable guidance contributing to the development of new, improved methodology).

If your research involves quantitative modeling, then consider whether these two statements apply:

(5) *The results from my modeling exercise confirm something that literally must be true – based on anticipation from conceptual reasoning – or something that is readily predictable from a simpler qualitative exercise or thought experiment.*

(6) *The results from my modeling exercise support the stated hypothesis or prediction, but if instead the model had produced different results, this study provides no basis for expecting that these different results would have been equally interesting or informative.*

If neither (5) nor (6) applies to your study, then it is reasonable to anticipate that reviewers might reward the study with high relative perceived merit. If either statement applies, however, then there is likely to be some doubt about how valuable the work really is. In this case, we don't really learn anything new from the modeling that is not already apparent, often intuitively. A simple example illustrates this: The multiplication of $2 \times 8 = 16$ is a mathematical model for predicting the outcome of adding $2 + 2 + 2 + 2 + 2 + 2 + 2 + 2$. This model of course provides practical value because it predicts the result efficiently (especially when the required addition involves many digits), but we don't really acquire any new understanding from the result of this model beyond what is already apparent from the conceptually simpler exercise of addition. This kind of modeling runs the risk, for some, of becoming merely an esoteric exercise, presumably because it has some kind of intrinsic intellectual value, purely for its own sake – like solving a jigsaw puzzle. The goal of modeling in this case often involves repeated revision of the model until it produces the result that the researcher has anticipated should be 'correct'. This may provide economy or functional value – e.g. by distilling a paragraph of description into a single equation – but if this is all it does, then there is no real conceptual advance. High quality modeling produces a new progressive and inspirational paragraph of description, and in our view, has at least the potential to produce unanticipated results – and importantly, results when unanticipated, that are interesting and informative in spite of this.

Conclusions

Perhaps only a small percentage of ecology research can be expected to emerge with top scores in the above assessment, and presumably all journals are interested in publishing these papers. Yet very high quality papers and/or very low quality papers apparently do not end up published evenly across the pool of available journals within a field of study – at least this is what is suggested by the distribution of impact factors for ecology journals in Fig. 1. Obviously, different journals, different editors, and different reviewers will place differing degrees of emphasis on the assessment criteria discussed above. Because of the qualitative nature of these criteria, variation in personal views is inevitable, especially in a multi-farious field like ecology. Accordingly, we do not expect unanimous agreement with our evaluation proposed here, any more than one would expect unanimous agreement among the judges of a figure-skating competition. Indeed, from a poll of subject editors at Oikos, there was general agreement that most of the criteria discussed above represent important limitations on the quality of research, and hence potentially significant limitations on suitability for publication in Oikos (Table 1). However, there was considerable variation in views for criteria (4) and (6). (Note that participants, when polled, did not have access to the manuscript of the present paper). We recognize also that there are additional factors – quite apart from merit – including sources of bias and elitism, that can play decisive roles in the editorial/reviewing policies and practices of some journals (Lortie et al. 2007, Aarssen and Lortie 2009). We hope that our suggestions will inspire broader discussion of this topic based on the experiences of others as editors, reviewers, and authors in ecology and in other fields of study.

It is important to emphasize that *discovery* – not publication, is the primary mission of science. Nevertheless, few research projects are launched without anticipation that the results will make a meaningful contribution, and thus be publishable. We suggest that rigorous self-assessment, as in the example provided above, can enable researchers to arrive at a reasonable estimation of the relative merit of their work, and so to make sensible choices in selecting a suitable journal for submission, and in making revisions to research proposals. Because of their potential for more limited knowledge of the literature, relatively young

Table 1. Responses from a poll of 22 subject editors at Oikos who were asked to consider scenarios in which they had detected six quality assessment criteria (see statements (1) to (6) in the text) as characteristics in submitted papers that they were handling for Oikos. In each case, participants were asked to rate the extent to which they would agree with the following: "*this characteristic should be regarded as an important limitation on the quality of the study, and hence a potentially significant limitation on its suitability for publication in Oikos*" (participants did not have access to the manuscript of the present paper). Table entries show the percentage of responses for each criterion.

Criterion	Frequency (%) of editor responses				
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
(1)	–	–	4.5	68.2	27.3
(2)	4.5	–	–	68.2	27.3
(3)	–	–	13.6	22.7	63.6
(4)	9.1	45.5	18.2	9.1	18.2
(5)	–	9.1	13.6	50.0	27.3
(6)	9.1	18.2	50.0	9.1	13.6

researchers (e.g. graduate students), in seeking feedback from mentors, might be advised to request specific comments on their research proposals and manuscripts that are guided by the assessment criteria proposed here. At the same time, these criteria may guide young researchers in developing their own reviewing skills as they become more directly involved in the peer-review process (Donaldson et al. 2010).

Manuscript rejection rate increases with journal impact factor (Aarssen et al. 2008), and so, for studies that rank lower from self-assessment, it is clearly advisable to submit them to correspondingly lower-ranking journals. Since there are many more of these journals to choose from (Fig. 1), then a lower ranked self-assessment, ironically – and encouragingly for most authors – probably does not mean a lower probability of publication. There is also some evidence, as one might predict, that submission to these lower impact journals can result in publication sooner (Koricheva 2003). In addition, choosing an appropriately ranked journal for one's study in the first submission attempt further expedites the publication process by eliminating some, if not most, of the 'ratcheting down the impact factor ladder', until the 'correct' journal is found. As authors in science, we are not playing a lottery; our aim is to communicate our research results quickly and effectively to the scientific community. Using self-assessment instead of peer-review and/or chance in identifying the 'correct' journal maximizes the efficiency of research time and energy for all concerned – authors, editors and reviewers. Ultimately, speeding up the publication process accelerates the potential rate of progress within science.

Acknowledgements – This research was conducted as part of a working group – 'The role of publication-related biases in ecology' – supported by the National Center for Ecological Analysis and Synthesis (NCEAS), funded by NSF (grant no. DEB-0072909). Helpful comments on the manuscript were provided by Carlos M. Herrera, Tiffany Knight, Wim van der Putten, Eric Seabloom and Diego Vázquez.

References

Aarssen, L. W. and Lortie, C. J. 2009. Ending elitism in peer-review publication. – Ideas Ecol. Evol. 2: 18–20.

Aarssen, L. W. et al. 2008. Bang for your buck: rejection rates and impact factors in ecological journals. – Open Ecol. J. 1: 14–19.

Donaldson, M. R. et al. 2010. Injecting youth into peer-review to increase its sustainability: a case study of ecology journals. – Ideas Ecol. Evol. 3: 1–7.

Hochberg, M. E. et al. 2009. The tragedy of the reviewer commons. – Ecol. Lett. 12: 2–4.

Koricheva, J. 2003. Non-significant results in ecology: a burden or a blessing in disguise? – Oikos 102: 397–401.

Lortie, C. J. et al. 2007. Publication bias and merit in ecology. – Oikos 116: 1247–1253.

Møller, A. P. and Jennions, M. D. 2002. How much variance can be explained by ecologists and evolutionary biologists? – Oecologia 132: 492–500.

Murtaugh, P. A. 2002. Journal quality, effect size, and publication bias in meta-analysis. – Ecology 83: 1162–1166.

Murtaugh, P. A. 2007. Simplicity and complexity in ecological data analysis. – Ecology 88: 56–62.

Pinker, S. 2002. The blank slate: the modern denial of human nature. – Penguin Books.