individual is likely to engage in conspicuous reproductive activities, it is also likely to engage in other conspicuous activities (e.g., foraging). These consistent behavioral differences over time and across contexts are likely to influence the frequency with which individuals with different personality types encounter one another. To date little work has explored the relationship between personality and temporal activity levels, and no work to our knowledge has explored the connection between personality-dependent differences in the timing of activity and reproductive isolation. Likewise, the potential implication of variable habitat use by individuals with different personalities on the evolution of reproductive isolation has not been investigated.

Concluding remarks and future directions

Here, we touch on several ways that personality could influence speciation. The paucity of controlled studies testing for causal links between personality and reproductive isolation points to the need for more deliberate work in this area. Several critical gaps exist in our current understanding of personality as it relates to speciation. First, our understanding of how personality can affect the evolution of traditional pre-mating reproductive isolation barriers (e.g., immigrant inviability and sexual or behavioral isolation) is limited to only a few case studies. Second, we know little about how personality is expressed in hybrids, and whether post-mating reproductive isolation barriers can be driven by maladaptive hybrid personality traits. Finally, although causal relationships between personality and dispersal tendencies have been established, little is known about how personality affects the formation of peripheral isolates and subsequent divergence and speciation within those peripheral isolates. Understanding the interplay between personality, ecology, and evolution offers exciting prospects for future speciation research.

References

1 Wolf, M. and Weissing, F.J. (2012) Animal personalities: consequences for ecology and evolution. *Trends Ecol. Evol.* 27, 452–461
2 Sih, A. *et al.* (2012) Ecological implications of behavioural syndromes. *Ecol. Lett.* 15, 278–289
3 Butlin, R. *et al.* (2012) What do we need to know about speciation? *Trends Ecol. Evol.* 27, 27–39
4 Wilson, D.S. (1998) Adaptive individual differences within single populations. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 353, 199–205
5 Ritchie, M.G. (2007) Sexual selection and speciation. *Annu. Rev. Ecol. Evol. Syst.* 38, 79–102
6 Cote, J. *et al.* (2010) Personality-dependent dispersal: characterization, ontogeny and consequences for spatially structured populations. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 365, 4065–4076
7 Nosil, P. (2012) *Ecological Speciation*, Oxford University Press
8 Kruuk, L.E.B. and Gilchrist, J.S. (1997) Mechanisms maintaining species differentiation: Predator-mediated selection in a Bombina hybrid zone. *Proc. R. Soc. Lond. B: Biol. Sci.* 264, 105–110
9 Both, C. *et al.* (2005) Pairs of extreme avian personalities have highest reproductive success. *J. Anim. Ecol.* 74, 667–674
10 Schuett, W. *et al.* (2010) Sexual selection and animal personality. *Biol. Rev. Camb. Philos. Soc.* 85, 217–246

Errors in science: the role of reviewers

Tamás Székely[1,2], Oliver Krüger[2], and E. Tobias Krause[2,3]

[1] Department of Biology and Biochemistry, University of Bath, Bath, UK
[2] Department of Animal Behaviour, Bielefeld University, Bielefeld, Germany
[3] Behavioural Ecology Group, Department of Animal Sciences, Wageningen University, Wageningen, The Netherlands

Reviewers play a key role in science, although studies suggest the current peer-reviewing system has faults. We propose to introduce a quality control system to evaluate each journal's review process, and produce a Review Quality Index. We propose four schemes that have the potential to reduce errors in a key step in scientific decision making: the reviewing process.

Whilst historically it was enough being a gentleman to publish in top scientific journals, in contemporary science the authors have to run the gauntlet of editors and reviewers. Although pre-peer review decision-making has been critically reflected on by Arnqvist [1], the reviewers' immensely important role also needs to be considered. Reviewers are the gatekeepers of scientific publishing, because they provide the essential quality checks by drawing the editors' attention to inconsistencies, poor methodology, or broad-brush writing [2]. Reviewers, however, are under increasing pressure to focus on their own research, and hence reduce the time they spend on charitable actions such as assessing somebody else's work. Because editors are increasingly relying on specialist reviewers' opinions, the reviewers shoulder immense responsibility in separating the good workmanship from the poor one. However, there are signs that some reviewers are not doing as proper a job as they are supposed to [3,4].

We have two objectives in this article. First, we propose that the process of manuscript reviewing needs to be evaluated and improved by the scientific publishing community. Second, we propose four schemes that have the potential to make improvements. Although each would require some effort (and modest costs), we believe a quality control scheme will ultimately benefit the scientific community by saving people's time, and improving transparency and objectivity.

Informal discussions with academics and researchers suggest that based on the Comments to Authors, a substantial percentage of review statements might contain

Corresponding author: Székely, T. (T.Szekely@bath.ac.uk).
Keywords: peer reviewing; review quality; reviewers; fairness.

0169-5347/
 http://dx.doi.org/10.1016/j.tree.2014.05.001

factual errors, show biased judgments, stipulate to cite the reviewer's own research paper whilst ignoring other important papers in the field, and/or pressurizing the deciding editor to bias the interpretation of the results toward one pet theory. In addition, authors might feel that the confidential comments to the Editor can also lead to biased editorial decisions. It is hard to estimate the extent of these errors, although the low consistency between different reviewers of a given manuscript suggests that one (or several) review(s) might be erroneous leading to disagreement between reviewers [3,4]. We fear that the error rate in reviews might be substantially higher [3] than the 5% (or less) Type I error that scientists commonly stipulate in their own research, and this calls for actions.

Thus good science might be rejected for bad reasons, causing delays in publishing important results. The editors often do not have time (or expertise) to provide fair judgements, and more often than not, they side with the reviewer who might be the editors' choice in the first place: the author(s) are treated as guilty by default. Erroneous review statements can lead to wrong editorial decisions and thus put an additional pressure on the already overstretched scientific publishing system, because subsequently the author(s) can appeal, or submit their manuscript to a different journal. However, increasing the number of reviewers [4], swapping to open peer-reviewing [5] or publishing referee reports together with the research manuscript [6], are impractical or do not seem to improve the quality of the refereeing process [2,5,6].

We propose four potential scenarios. Firstly, journals might select a random set of submitted manuscripts, and upload these manuscripts along with the reviewers' statements and the editorial decision (reject or accept) to an open repository. All of these should be anonymised, that is, the identity of authors, reviewers, and editor(s) would only be known to the journal. Voluntary peers can then evaluate the reviews, with respect to correctness, objectiveness, and fairness, and these peer evaluations would provide a score for each reviewer. These scores can be coarse (e.g., fair, neutral, or erroneous), or finer-ranging, for instance between one and five. The reviewers' scores can then be calculated for each journal, and overall, produce a Review Quality Index for each journal. Scientists are likely to participate in this process that improves quality and fairness in scientific publishing, given that humans have a tendency to contribute altruistically to systems that increase fairness and transparency [7]. Although the cost associated with this scheme would be small, peer reviewers are self-selected and this might not guarantee an entirely error-free assessment.

Secondly, accompanying their decision letters, the journal editor can also send a link to an online query where authors can rate the reviewers' statements with respect to correctness, objectiveness, and fairness [5]. These anonymous ratings would be stored by a web-based review repository that is independent from scientific journals. In addition to the authors' voluntary ratings of the reviewers' performance, journals can also provide the corresponding editorial decision (accept or reject) to the web-based depository. From the authors' voluntary ratings, a Review Quality Index can be calculated. It is conceivable that the rejected versus accepted manuscripts will generate different reviewer quality scores by the authors, although this is not the point of the scheme. The focus is on comparing the scores across different journals separately for accepted and rejected papers. This scheme seems easy to implement, although it crucially depends on the willingness of authors and journal editors.

Thirdly, the editorial board of each journal might select an external panel of senior experts who evaluate a random subset of review statements of both published and rejected manuscripts. Evaluation should be blind: the external quality control panel would neither know the authors of the original study nor the identity of the reviewers. This process again would produce a Review Quality Index separately for each journal. A given scientific field (e.g. evolution, ecology, genetics, neuroscience) might use a shared panel of experts so that the Review Quality Indices will be comparable between journals of a give scientific field. However, it might be difficult to find enough senior experts as their crucial participation would mean an additional workload that would further strain the already time-consuming peer review system.

Finally, journals might provide the details of articles submitted, but rejected, to an external web-based repository. This external repository could then track the fate of rejected manuscripts, and monitor whether they are subsequently published by another journal. Using the citation of these initially rejected versus accepted manuscripts, it would be possible to compare the reliability of decision-making process. Based on these data, a Review Quality Index might be calculated. One potential problem of this scheme that it requires accurate identification of a particular piece of work, although manuscripts often go through substantial changes between submissions to different journals. Also, the manuscripts that are resubmitted to different journals can be a biased subset of all manuscripts that were initially rejected.

We propose that a Review Quality Index for a journal, indicating how reliable and fair their reviewers' statements and editorial decisions are, might prompt reviewers to do more decent jobs, and expedite the editors to discount substandard review statements. Such a Review Quality Index might be displayed by citation databases, such as the Thompson Reuters, and can be established as an accompanying index to the Journal Citation Reports. We believe the Review Quality Index would make some journals more attractive: knowing that a journal has a tough, but well-judged, decision mechanism would improve the journal's reputation. For example, if two top journals have 80% rejection rates, but one has a much larger margin of error, authors who believe their work is important would probably prefer to submit their manuscript to the journal with lower error rate.

Introducing a quality control would therefore benefit the scientific community, and reduce the potentially erroneous decisions that can have severe knock-on effects on the direction of scientific research and on people's careers. The Review Quality Index would also make scientific reviewing more transparent and objective. The costs of the different schemes vary, although given that they only involve a subset of manuscripts and the associated reviewer statements, these costs may be modest.

Given that scientific funding agencies are investing vast sums into quality control for administering and running grant awarding systems, it is striking that few (if any) efforts are invested into assuring high quality in a key step of the scientific process: manuscript reviewing. It is now time to improve the system for the benefit of all involved. We hope that our paper highlights the need for change, and that it will encourage editors, publishers, researchers, and scientometricians to discuss these propositions.

Acknowledgements

We appreciate the suggestions of Dr Paul Craze and an anonymous reviewer on a previous version of this manuscript. T.S. was a Mercator Professor at Bielefeld University. E.T.K. was funded by the Volkswagen Foundation (85994).

References

1 Arnqvist, G. (2013) Editorial rejects? Novelty, schnovelty! *Trends Ecol. Evol.* 28, 448–449
2 Park, I.U. *et al.* (2013) Modelling the effects of subjective and objective decision making in scientific peer review. *Nature* 506, 93–96
3 Kravitz, R.L. *et al.* (2010) Editorial peer reviewers' recommendations at a general medical journal: are they reliable and do editors care? *PLoS ONE* 5, e10072
4 Jackson, J.L. *et al.* (2011) The validity of peer review in a general medicine journal. *PLoS ONE* 6, e22475
5 van Rooyen, S. *et al.* (1999) Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *Br. Med. J.* 318, 23–27
6 Pulverer, B. (2010) Transparency showcases strength of peer review. *Nature* 468, 29–31
7 Fehr, E. and Fischbacher, U. (2003) The nature of human altruism. *Nature* 425, 785–791