

# Standardising Peer Review in Paleontology journals

Jon Tennant

22 January 2020

## Introduction

“However ill-defined it may be, the peer-review process is still the gold standard that will continue to drive scholarly publication.”

“While it is not a perfect process, traditional peer review remains the gold standard for evaluating and selecting quality scientific publications.”

Mayden, 2012.

That peer review operates as a ‘gold standard’ for scholarly research has become a common, institutionalised mantra. It is widely considered to be fundamental in maintaining the rigour and validity of scholarly research. However, not only is peer review non-standardised, it also fails most basic tests of academic credibility and objectivity by being secretive, exclusive, and irreproducible. The process is often opaque, which can introduce bias into reporting standards for research and impact the overall quality of the published record. Despite being widely criticised, it remains almost ubiquitous as a critical part of scholarly communication systems. At the present, because we typically have almost no understanding of what peer review actually does, even generally, it can only be considered a ‘pseudo-standard’ at best. Training and support is generally lacking, and it is often the case that reviewers, through no fault of their own, are unaware of the critical questions to be asking with respect to research design, methods, reporting, and analysis.

In a modern knowledge production system, peer review is largely coupled to journals. Given that there are over 200 journals that publish palaeontology research Tennant and Lomax, 2019, this can make the process confusing for reviewers, irrespective of their relative expertise. As these journals are mostly owned by commercial entities in paleontology, they have a primary duty to increase financial value, and not to improve scientific quality or legitimacy. Because of this, management of the function of peer review is not in the hands of the research community, which means that our collective responsibility to ensure that it is performed to a high standard of integrity has been compromised. It is our duty as scholars to be critical of all elements of knowledge production, and if peer review is to be considered to be any sort of test of integrity or validation, then we must improve it.

The aim of this project is to formulate a clear set of guidelines explicitly for reviewers in palaeontology journals, or multi-disciplinary journals that include palaeontology submissions. Through this, peer review can be more transparent and objective, representing a valid form of peer review best practice and part of responsible research conduct. It can also help to improve the soundness and reporting standards for palaeontology research as a whole, increasing the field’s overall legitimacy. These guidelines were inspired by Parker et al. 2018, who created a similar checklist for the fields of ecology and evolution. Some of these points are adapted from the TTEE (Tools for Transparency in Ecology and Evolution) guidelines.

The purpose of this operational checklist serves one primary purpose: For an article to ‘pass’ peer review, articles must satisfy a specific quality threshold based on standardised guidelines. Thus, the quality of the peer review process is simultaneously ensured through an open and technical standardisation process. This should be of interest to all stakeholders engaged in the publishing process, including authors, editors, reviewers, and the publishers themselves, who all have a duty to uphold the integrity of the published research record. While it might initially increase the bureaucracy involved in publishing, ultimately it should save time and

effort as it becomes more widely established as an embedded scholarly norm, with integrity a formative part of peer review culture.

## Guidelines

These guidelines are presented in a way that reflects the typical structure for research articles. They are presented as a specific checklist of items or questions to be fulfilled during the review process, and are written from the perspective of the reviewers. Specific sections are provided for articles that either involved quantitative research or the use of fossil specimens. Modern palaeontology research usually falls into two categories:

1. Quantitative, involving data analysis and statistical methods.
2. Qualitative, involving the systematic description of new species or specimens.

Some of the items in this checklist will not be applicable to different types of research paper. Some paleontology journals also only publish on specific types of research, and thus these guidelines remain operationally flexible to accommodate different selective publishing criteria.

## General

- Have any relevant funding sources been disclosed?
- Have any relevant conflicts of interest been disclosed? This can include editorial and institutional positions, any sources of income, and any other positions of authority that each of the authors might have.
- Have any previous versions of this work been indicated (e.g., preprints or pre-registrations)? If so, is this clear and all relevant links provided?
- Is the language used appropriate for a scientific publication?
- Is the manuscript structure appropriate?
- Have all relevant acknowledgements been made? This can include where materials were viewed or loaned from, where any technical analyses were conducted and by whom, and anyone or anything else that contributed to the research process.

## Figures

- Are any figures legible, relevant, and integrated into the text?
- Do all figures have alt-text to help those who might be visually impaired?
- Are figure files provided in their raw formats?
- Is the association between figures and the underlying data/methods needed to recreate them visible?
- Are any three-dimensional figures presented in a three-dimensional format?

## Tables

- Are any tables legible, relevant, and integrated into the text?
- Are the tables provided in both machine-readable and human-readable formats?
- Is it clear where the information contained within the tables have been sourced from?

## **Supplementary files**

- Are all supplementary files provided in non-proprietary formats and in a sustainable manner?
- Are all supplementary files referenced within the main text?
- Is there information within the supplementary files that is critical to the research, and should thus be in the main manuscript?
- Are all supplementary files also archived within a suitable repository outside of the journal?
- Do all supplementary files have Digital Object Identifiers (DOIs) to help with their citation and re-use?
- Is all information within the supplementary files also of a scientific standard, including with appropriate references?

## **Abstract**

- Is the abstract concise, conveying the main research findings?
- Are any key conclusions missing?
- Does the abstract over-embellish the key findings?

## **Introduction**

- What type of study is this (e.g., original research replication, exploratory, meta-review)?
- What is the primary purpose for conducting this study?
- Is this part of a larger study, or a one-off?
- What are the specific research questions addressed?
- Are all primary hypotheses that were tested sufficiently described?
- Is there any evidence of HARK-ing (Hypothesising After the Results are Known)?
- What is the scope of the study?
- Is the published literature covered sufficiently in depth and breadth?
- Are any relevant previously published articles missing?
- Are any appropriate angles or context missing?
- Is the history of the research sufficiently described?
- Are there any unnecessary or irrelevant references included?

## **Methods**

- Are the methods reported in sufficient detail and clear enough to allow another researcher to gather the same data and run the identical analyses?
- When was this study started and for how long was its duration?

## **Data collection**

- Are the supporting data included in the manuscript or in a relevant repository?
- Are the data presented in a way that is consistent with the FAIR principles? (Findable, Accessible, Interoperable, Reusable), Wilkinson et al., 2016.
- When were the data collected?

- Did the authors develop their analysis plan, including choices of variables, without looking at the data, for instance prior to gathering data or with a dummy data set?
- What were the conditions under which the data were collected or extracted?
- Are all information sources described, including accession dates?
- Were data collected and vetted by all, some, or just one of the authors?
- Were data obtained from previously published sources, or unpublished sources?
- Is justification given for data that were collected but not including in the final analyses?
- Are potential biases assessed and, if so, mitigated (e.g., publication bias, missing data, or selective reporting within studies)?
- Is it clear how representative the data are of reality?
- Is it clear about how the data might change in the future based on new acquisitions or publications? As in, how stable are these data to future research?
- Are there any geographical, political, or societal influences or constraints on the data?
- Are complete and explicit database search terms including all logical connectors and operators sufficient to exactly duplicate search, provided? Also, have all other sources of data (e.g. literature cited sections of published reviews, personal unpublished data, collaborative group project data available from cited sources or online) been indicated?

### **Involving fossil specimens**

- If relevant, possible and allowable, are all voucher specimens of the studied taxon/taxa deposited in an appropriate and publicly curated collection, with accessing numbers?
- Were all relevant permits and permissions obtained to study specimens, and to publicise them?
- Are all relevant contextual information provided? This includes:
  - Taxon name
  - Taxonomic level
  - Age of specimen, both stratigraphically and in Ma, with justification given
  - Horizon of discovery, including bed, member, formation, and group
  - Lithology specimen discovered in, and environmental interpretations
  - Geographic location of specimen, including GPS co-ordinates and latitude and longitude
  - Discoverer of specimen, and date of discovery
  - Previous taxonomic diagnoses and information, including synonyms
  - Relevant type and referred specimens
  - Original, revised, or emended diagnoses
  - Taphonomy of specimen and mode of preservation
  - Methods of collection
  - Methods of preparation
  - Completeness of specimen and articulation
  - Any associated fossils

### **Involving quantitative analysis**

- Why was this precise study design selected?
- Are any limitations or published criticisms of the methods appropriately mentioned?
- Were any assumptions, simplifications, or transformations to the data made?
- Are any methods of weighting applied, and justifications for this given?

- If methods have been previously published, are relevant citations given?
- What statistical tests were formed, and why were they chosen?
- What is the total sample size?

#### **Software environment**

- Are all software packages and scripts identified, including their versions, availability, and annotated full code or set of commands, provided?
- Is the full model specification provided, including specific parameter choices and their justification?
- Are all software and scripts available in a public repository and under an appropriate FOSS license?

#### **Results**

- Are the results presented in a clear and coherent manner?
- Are all results reported, or are some omitted?
- Are the results reported in a way that is supported by the data?
- Are the results sufficiently illustrated by relevant figures?

#### **Involving quantitative analysis**

- Are the full outcomes of statistical tests reported, including basic parameter estimates of central tendency (e.g., means), other basic estimates (e.g., regression and correlation coefficients), and variability (e.g., standard deviation) and associated estimates of uncertainty (e.g., confidence or credibility intervals)?
- Are appropriate test statistics reported, including the degrees of freedom and p-values?
- For Bayesian analyses, have all results been reported, at a minimum including information on choice of priors and MCMC (Markov Chain Monte Carlo) settings (e.g. burn-in, the number of iterations, and thinning intervals)?

#### **Involving fossil specimens**

- Is the morphology described in a systematic manner?
- Are sufficient measurements provided?
- Is sufficient justification given for the erection of any new taxa?
- Are significant anatomical features appropriately described and figured?
- Are all relevant synapomorphies and apomorphies identified and described?
- Are all relevant morphological features compared to other related taxa?

#### **Discussion**

- Is the strength of the evidence evaluated and interpreted?
- Are the sample sizes large enough to justify the authors' conclusions? If presenting significance tests, how much power would this study have to detect statistically significant weak, moderate and strong effects?
- Are potential sources of bias, and their consequences, appropriately evaluated and interpreted?
- Do the results contribute to the research field, irrespective of whether they are 'negative' or a replication study?

- Are the new results placed into sufficient context of the relevant literature?
- Are all arguments well-balanced?
- Are the full implications of the new results discussed in sufficient detail?

## Conclusions

- Are the conclusions supported by the results?
- Are the conclusions concise and not over-embellished?

## Discussion

Above, we have provided guidance on a specific, itemised check-list for peer review of palaeontology research. This is an important step in standardising the peer review process, and improving the quality of published palaeontology research. Its effectiveness will be realised when it is up-taken by palaeontology journals, who can integrate these guidelines into their reviewer policies as part of ensuring that they are facilitating peer review to a sufficient quality standard. These guidelines can also be used as an objective checklist for authors. If these items are all appropriately checked prior to submission, then substantial time, money, and effort can be saved during the review process.

This checklist is not designed as a replacement for peer review, or for critical thinking and the application of wider forms of expertise. Nor do we make recommendations about how researchers should subjectively assess the ‘impact’ or ‘quality’ of a study. The primary design is to provide a baseline standard for research such that minimum standards of reporting quality are met, which is part of the palaeontology community’s collective responsibility to assure. They are a compliment to traditional peer review, and when combined can improve the creativity, intellectual diversity, and operational functionality of the process. With this checklist in mind, it also becomes contingent on reviewers to make sure that they themselves have the scholarly expertise to check such things. This will include a relevant knowledge of any methods, as well as a mastery of the relevant scholarly literature.

Upon implementation of these guidelines, there are some important implications to further consider. Just because research passes these guidelines, it does not automatically mean that it should be accepted for publication; however, the onus will then be on editorial standards to justify rejection of research, should it be scientifically sound. Furthermore, if we have appropriate, community-driven and objective standards for defining quality in peer review, then it might be the case that we no longer need to have secondary and redundant forms of post hoc evaluation based on things like journal brands or impact factors. The validity of a piece of research is qualified through an objective quality assurance process, and not just arbitrarily assumed based on a binary process of whether it was peer reviewed or not. For something to be standardised, it has to attain a specific level of technical quality for comparative purposes. For example, examining the utility and functionality of both peer review, and the quality of published scholarly research. These guidelines provide this.

## References