How to Write a Review for a Scientific or Mathematical Paper

Matthew Roughan

ARC Centre of Excellence for Mathematical & Statistical Frontiers (ACEMS), School of Mathematical Sciences, University of Adelaide, Adelaide, 5005, SA, Australia

Abstract

Reviewing a paper or grant proposal is a difficult and thankless task. Yet tens of billions of dollars in grants are dispersed on the basis of peer review. And careers are made and broken through publications that are judged through peer review. The goal of this report is to provide an education resource to improve the quality and efficiency of reviews

The task of the referee is to evaluate in a timely manner a paper for publication in a specific journal or conference proceedings. This involves determining if the work presented is correct, if the problem studied and the results obtained are new and significant, if the quality of the presentation is satisfactory or can be made so, and what revisions and changes to the paper are necessary and/or desirable. The evaluation must be with regard to the coverage and degree of selectivity of the specific publication.

Alan Jay Smith [1].

10

12

15

16

18

The most important job of a program chair, program committee, and external review committee is to deliver expert, fair, and thoughtful reviews

Kathryn S McKinley [2].

Although a person is not born with the knowledge or ability of how to be a good reviewer, the characteristics (*e.g.*, fairness, thoroughness, integrity) of that person certainly contribute to the activity. Unfortunately, it is rare to find a scientist whose formal training has incorporated instruction in the art of reviewing.

Benos, Kirk and Hall [3].

Build a man a fire, and he'll be warm for a day. Set a man on fire, and he'll be warm for the rest of his life.

Terry Pratchett, Jingo, 1997.

Email address: matthew.roughan@adelaide.edu.au (Matthew Roughan)

Preprint October 2, 2020

1. Introduction

27

28

30

32

34

35

36

38

42

43

47

50

51

52

53

54

57

61

62

63

64

65

Writing a review of a paper submitted to a journal or conference is a standard academic task. Likewise, reviewing a grant proposal. The task has much in common with reviewing work in other contexts as well.

The results of the review process can have a profound effect on people's careers. There is also evidence that bad peer review processes contribute to systemic injustice and discrimination. Even if the review itself is not biased by considerations of race or gender, unprofessional reviews disproportionately harm underrepresented groups [4] by reinforcing (through a supposedly legitimate channel) stereotypes that such groups are constantly bombarded with.

However, I have observed as an author, editor and program committee member and chair that too often the standard of reviews is low. We all have many experiences that bring this home – my most memorable one was being told my grant application was "too poetical."

I feel that if our system of reviewing had been around back then, Homer might hear from a reviewer that the Iliad was a "good first attempt but the story is really unbelievable and it can't possibly stand the test of time." They might also complain that it was "too poetical."

The experience of poorly reviewed papers is so universal there are now Facebook and Reddit groups devoted to "Reviewer #2 Must be Stopped". There are great memes floating around² about what happens to papers to satisfy such reviewers, and collections of the at best unhelpful, and at worst downright spiteful comments that reviewers have made³. And there are several satirical articles such as "If the Script for Avengers: Infinity War Was Peer-Reviewed" [5] and "If the Script for Star Wars: The Last Jedi Was Peer-Reviewed" [6].

Apart from such anecdotal reports, scientific studies of the review process have been performed, *e.g.*, [7] resubmitted successful papers with fictitious authors. Few re-submissions were detected, and 16 of 18 reviewers then rejected the papers claiming serious methodological flaws!

Much of the problem arises from asking over-busy people to do a thankless task again and again. I can't fix that. But one problem that can be addressed is that training in such tasks is usually left to individual graduate project supervisors, and it's patchy at best [3, 8]. This report aims to be a practical resource to help new (and old) participants in the academic world to fill that gap.

Apart from improving the standard of reviews, I hope that this will enable junior staff and even students to do reviews [9]. Apart from the value their energy can bring to the task, writing a review can provide insights into writing that are invaluable to the reviewer as well as the reviewee.

This document aims to lay out the process of writing a review. There are already several good reports on this topic, but I wanted something that I could adapt and grow with time according to my own understanding of the task, which is not always in concord with current trends.

The document is based largely on my experience in mathematics and computer science, but I believe many of the comments are relevant to a much wider set of fields.

I started this with the ambition of writing a very short (2 page) document. It's grown a little out of control. Sorry!

2. General Advice

In general, we should approach writing a review like any other professional task. The key task is to make a clear recommendation, and do it on time, but we can add a great deal of detail to refine this goal.

Underlying this are a set of fundamental goals best stated by Allman [10]: clarity and consistency (of decision making), and fairness. However, these are somewhat abstract objectives. What follows is concrete advice on how to achieve these goals in the specific context of a review.

• **Know the guidelines:** [11] Each journal or conference or grant scheme (we will call these generically *venues*) will have specific processes for its reviews. A reviewer should know and follow these. But also, a reviewer should be aware of the guidelines for the authors, particularly the constraints under which they submit their papers (*e.g.*, page limits), and the advice that the journal gives the authors. There are few things more frustrating for an author than a reviewer criticising a paper for following the journal's guidelines.

¹Reviewer #2 is now shorthand for the reviewer who does badly thought through or simply nasty review.

²https://amlbrowndotcom.files.wordpress.com/2015/10/11219225_10153619513398446_2657606012680909527_n.jpg

³https://shitmyreviewerssay.tumblr.com/

• **Reviews are confidential:** [1, 3, 12] work submitted to a journal or conference or grant scheme MUST be treated in confidence. That is, reviewers must not speak of or write about the ideas outside of the review process.

- Moreover, you cannot (as a reviewer) use those ideas yourself until the work is actually published⁴, and not using ideas you know about can be harder than you think.
- **Disclose conflicts of interest:** [1, 3, 12, 13] Reviewers should disclose and (usually) avoid conflicts of interest. Journals and conferences sometimes have specific guidelines as to how to define a conflict of interest, but it is often obvious. If I have a reason outside of the article's quality to either wish it published, or wish it rejected, then I have a conflict of interest [14]. This can arise because of a personal or business relationship with the author, or because of a personal or business agenda (*e.g.*, the work supports an idea I am otherwise being paid to promote). Conflicts of interest should be disclosed (for instance to the editor of the journal), but do not always rule out writing a review [9]. In some cases the conflict is minor, and in other cases all of the parties experienced enough to write a review would have some conflict [13]. It is a question of managing these conflicts, and that is the job of the editor, not the reviewer. If you are not sure then talk to the editor.
- Write a defensible review: Most reviews are anonymous⁵. This is so that reviewers can write without fear or favour. For instance, so that a junior academic can write what they believe about the work of senior academics who might hold sway over their future career. However, anonymity is *not* a license to do things that you would not do in public [13]. I always write a review as I would if I knew it could be made public. That does not mean I shy from honest criticism, but it does impose a level of rigour on what I write, and I write comments politely, as I would if the author was in the room with me, and armed with a sub-machine gun.
- **Be constructive:** Criticism should be written in such a way as to help an author write a better paper [1, 15]. Constructive criticism doesn't just list problems, but also tells the author how to change the work to improve it. Definitely avoid things such as sarcastic or abusive comments or anything else unprofessional.
- **Be concrete:** Criticism needs to be clear and concrete [1, 15]. Vague comments are not helpful. For instance, don't say "there is additional work that should be cited," say specifically what should be cited. I shall say more about this in Section 5.
- **Be objective and impartial:** A review process must be fair. To be fair, it should be as objective as possible [1], otherwise it brings into doubt its own validity. Reviewers, for instance, should not seek to manipulate the process to advance their own ideas [3]. I will say more on this in Section 6, but note that fairness and objectivity are the foundations of GOOD SCIENCE.
- Approach the review with a positive attitude: Too often we resent the time involved in reviewing, or otherwise find ourselves with a negative attitude. In reviewing for high-quality venues sometimes people feel they are the "defenders of the faith," and must preserve quality. Of course we do aim to preserve quality but the attitude you should adopt in reviewing is that someone has worked hard on their manuscript, and your job is to help them get their good ideas published. Try to see the big picture, and not get caught in the minutiae of little corrections. No plan survives contact with the enemy, but if you start with this thought in mind, then you look for the good in a paper, which results in a more constructive review, and you may find you enjoy reviewing more as well.

There are always exceptions to rules, but they should be discussed with the person or organisation managing the process.

All of the above rules require work – more work than writing a sloppy and careless review. However, remember that in these days of "publish or perish," an author is engaged in a high-stakes activity in submitting their paper, and if you cannot be bothered being professional about the review, you don't deserve to be involved.

Some commentators on reviewing [8] suggest you do three "readings" of a paper: (i) to get overall impression, (ii) to concentrate on nuts and bolts, *i.e.*, the detailed technical read, and (iii) (after a break from the paper) to look at presentation, but also to check your own view of the paper and revise your review. I agree with this approach, in principle, but you also need to be efficient and put the work in where it is needed. Sometimes one read is all it takes to see the critical flaw in a paper. Sometimes a paper may take much more effort to assess properly.

⁴Pre-print archives alleviate this problem a little.

⁵We could argue about the merits of open review, but that's not the current process in most cases.

3. What is a Review?

The answer to "what is a review?" seems obvious, but it is not. Superficially it is a report describing the value (or lack thereof) of a submitted manuscript or project proposal, but we can be more specific about the types of values that are important in an academic or scientific context. Specifically, when we are reviewing we aim to assess several qualities. Each involves subtleties that are sometimes overlooked.

The qualities roughly in order of importance are listed below.

- 1. **Correctness:** There is, of course, a desire that published papers are "correct" [12, 16]. A key task of a reviewer is to assess this. However, a reviewer cannot truly assess correctness. They don't often have access to the raw data or code of the authors, and even if they did they would not usually have time to repeat all results in detail. So correctness is about determining correctness of process. We attack it through questions such as
 - Is the problem being solved clearly stated?
 - Is the stated problem actually solved?
 - Did the authors use the correct tools and techniques?
 - Are they aware of the limitations and assumptions implicit in these tools?
 - Do they analyse errors in their data?
 - Is their logic sound?
 - Are statements backed up by data or other sources?
 - Does it contradict other established work?⁶

And there are other examples of such questions in [1, 8, 11, 12]. However, the meta-concept of these is to look for flaws in execution, not results.

An incorrect paper (one that doesn't meet these standards) should be rejected, or only accepted with major revisions (that are checked by a subsequent review). However, there are often minor discrepancies that are less serious, and we need to accept that most submissions will not be perfect in every detail we wish, and not reject every paper [1]. Space and time are limited, and no work is ever perfect but a reviewer can help improve the correctness by providing helpful feedback.

- 2. Novelty: Novelty is not originality. A paper should have novel content, but this is sometimes confused with the idea that all papers should be composed of only original ideas. Good science is built incrementally by building on the work of others, and from repetition studies [17]. Synthesis of existing ideas into new tools is also very useful, as is transference of ideas from one field to another. However, authors should be aware of, and describe the context in which their work exists, and how their work relates to this context. And there should be something about the paper that adds to our knowledge, if only by confirming our existing knowledge.
- 3. Reproducibility: Good science is reproducible [11]. A paper should describe its experiments, tools and techniques at a level of detail suitable for a reader to be able to reproduce the results. This is not always possible within the constraints of sometimes short page limits, or the constraints of privacy ethics. However, there are ways to make a paper more reproducible: release of data and computer code are good examples [18]. A longer version of a paper may be placed in an archive or other store to provide missing details. All papers should aim for high standards in this regard, and reviews should provide advice about how to improve the reproducibility work described by a manuscript.
- 4. **Presentation quality:** The quality of presentation of a work affects the ability for readers to understand it and gain knowledge from it. It also affects your ability to review it properly if the paper is too poorly written [1, 16]. This is often treated as a hygiene issue, *i.e.*, the paper is either clean or not⁷. However, I believe that we should be providing more weight to well-written papers, and trying to help authors in all cases to improve the presentation of their work. In this regard concern yourself primarily with the quality of communication, not the adherence to strict (and often outmoded) English rules, though these play a distinct part in good communication. However, a

⁶One has to be quite careful about this, because contradictions can be correct, and such works (when correct) are often the most valuable [1].

⁷A person is either clean or not. Washing your hair ten times a day does not make it significantly cleaner.

reviewer in not a copy editor. It is not your job to find typos or spelling mistakes (see Section 6). A paper that does not pass a basic English hygiene test should not even be reviewed in more detail.

5. **Importance or interest or significance:** A final filter for papers is whether the work is interesting [12, 16]. There is an infinite pool of boring papers that satisfy the above criteria so some extra discrimination is needed. That is why the importance of the work should be considered. But note that "importance" is near last in my list, because it is too often today treated by editors and reviewers as the prime criteria. This last filter is entirely subjective, and reviews should be objective as far as is possible. The true test of a paper's significance is how it impacts future research and applications, so let's allow that future to happen.

Of course we don't want to publish boring and trite results, but it is intrinsically hard to predict what work will be truly important in the long run. Said G.H. Hardy about his own work:

"No discovery of mine has made, or is likely to make, directly or indirectly, for good or ill, the least difference to the amenity of the world."

His discoveries are now hugely important and practical: some feed into cryptography, and he is also joint father to the Hardy-Weinberg principle in population genetics, a crucial and much used result. If even the author of an idea can predict its importance with poor facility, what hope does a reviewer have? There is a large amount of actual evidence that peer reviewers [19] are not good at predicting the success of future work.

There is a secondary problem created by pushing the idea of importance to its extreme: Nosek *et al.* [17] note that "The real problem is that the incentives for publishable results can be at odds with the incentives for accurate results." The need to make results exciting leads to overzealous selling of the latest fashion versus building solid foundations of long-term value. Creating false incentives creates a disconnect between what is good for authors, and what is good for science.

One of the most damning phrases a reviewer can use is "this research is incremental," implying that it adds nothing significant to our knowledge. But science is incremental. And the value of work is hard to predict. Do not be lured down this rabbit hole.

A subset of "significance" that can be more usefully and objectively assessed [16] is "does the manuscript align with the topics of interest to the journal or conference?" However, I find this is not too often an issue – authors are reasonably sensible about choosing venues, except in odd cases, which are usually easy to reject on other grounds as well. It can be an issue, however, for multidisciplinary papers that have no obvious home, so I prefer to be inclusive.

In summary, I think of significance as another hygiene test. Rather than looking for "the next big thing" we should exclude only obviously boring papers. I realise this is controversial in these days where journals and institutions are all craving "impact factor." But I doubt their ability to obtain real impact through chasing the latest trends. True impact comes from new and consequently untrendy ideas.

- 6. **Ethical concerns:** Ethics is not last on this list because it is least important, but because it is relatively rare (in my experience) that in reviewing a paper you may become concerned about some ethical issue. There are three broad classes of potential concern⁸:
 - Scientific misconduct: misconduct formally [22] is considered to be falsification (distortion of data) or fabrication (making up data)⁹ and plagiarism [24] (see the next comment as well).
 - Authorship ethics (*e.g.*, see [25]): a submitted paper should be the authors' own original work, not previously published and giving appropriate credit to all involved [24]. Authors on a paper should have made a substantial intellectual contribution to the paper [24]. Multiple submissions¹⁰ are generally considered unacceptable. Conflicts of interest between authors, publishers and commercial interests should also be disclosed. There are other issues to consider and norms vary between fields, so be aware of potential differences.

⁸Definitions of scientific misconduct may vary between jurisdictions. Definitions included here are an approximation of US standards. However, specific locales and venues may have their own standards or additional requirements, e.g., see [14, 20, 21].

⁹A surprising number of scientists admit to some form of poor practice in at least some of their work [23], *i.e.*, around 2% admit to falsification or fabrication at least once, which is much higher than the rate of documented cases.

¹⁰Multiple submission is submitting the same work to more than one venue at the same time, even if the intention is only to publish it in once.

Scientific ethics: in research with human participants or animal subjects the ethical issues in such research
are very complex. Such research needs to be considered and approved by an appropriate Institutional Review
Board (IRB).

Journals and conferences may have their own specific ethics guidelines, e.g., [25], that should also be adhered to.

A reviewer is not always in the position to assess all of these, and is not expected to be able to do so comprehensively, but should bring any problems to the attention of the editor.

However, ethical questions are rarely simple, and differences of opinion are common. And honest errors are *not* misconduct (though they should be corrected). Ethical questions must be therefore handled with care as consequences for bad decisions can be serious (*e.g.*, career ending).

4. What Goes in a Review?

Many journals and conferences have their own form for a review, and this must be followed. However, I find it useful to keep in mind the key components of a review when composing it. These are:

- **Metadata:** I always start a review by reiterating the author, title and any reference number, just so that the review can't be misplaced.
- A clear recommendation: [1, 13] the recommendation is the critical component of a review. *Make a decision*. Sometimes reviewers forget it, but more commonly, they are vague. Please make this precise, and map it to the categories for the specific journal or conference. Commonly it will be a points based rating (e.g., 1-5, but make sure you calibrate your scores to the correct scale) or provide a recommendation of the form:
 - accept:
 - accept with minor changes;
 - major changes (implicitly requiring another round of reviews); and
 - reject

Some journals will have sub-categories, or just different categories. Adapt!

Note that being too permissive or too accepting are both bad. Likewise rating all papers 3 out of 5 (or otherwise on the fence) is not useful. You are there to help create a dividing line [1]. *Make a decision*.

Some academics believe the recommendation should be only reported to the editor, not the authors *e.g.*, [3, 13]. I disagree! It is easier (for the reviewers and editor) to hold a secret court, but I believe that transparency of decision making is a key aspect of fairness. Nevertheless, you must follow the policy of the particular venue in question.

- A short summary of the paper: [1, 13, 15] A short summary (1-5 sentences) to show what you believe the paper to be about is often helpful for an editor (who may not read the paper in detail). It's like an extra version of an abstract, as seen from another perspective, and helps put context around the following comments.
- A short list of the paper's key strengths (if any): It's worth putting a short bullet list of 2 or 3 key strengths of a paper. It helps you approach the review with a positive attitude. It also makes sure that the author receives some positive feedback in what is likely to be a sea of negativity (from other reviewers).
- A short list of the paper's key weaknesses (if any): One difficulty for an editor or author when they receive a long list of criticisms is to determine which are most important, and which are least. Providing a short list of the main weaknesses (again a bulleted list of 2 or 3 is enough) aids greatly.
- A critique: [1] The main body of the review is usually a list of detailed criticisms of the manuscript. This is discussed in detail in Section 5.

It is hard to say how long a review should be. It is common to say 1-2 pages [13]. However, the length of the critique depends a lot on the quality of the paper. I find that it is longest for borderline papers (I have received and written reviews that were closer to 5 pages). Borderline papers that could be accepted, but need lots of work have the longest critiques. Really good papers often need little feedback, and poor papers often have a few important issues to deal with before they reach the detailed critique stage so a review could be as little as half a page.

It's hard to say much more because the details of this section depend so much on the particular journal, conference, topic, style and particular manuscript. However there are two useful points I think that can help:

1. Divide your critique in some way so as to give advice about which points are most important, and which are less. *e.g.*, major flaws vs minor flaws vs typos [8, 10, 15].

247

248

249

251

252

254

256

257

259

260

261

262

263

265

266

267

269

271

273

275

276

278

280

282

283

284

285

287

288

290

291

292

2. Make clear which are necessary changes, and which are stylistic suggestions [1].

Without these guides, authors and editors cannot easily disentangle the important pieces of a potentially long list of comments, or resolve contradictory comments from multiple reviewers (which happens more often than you might hope).

Finally, please ensure that your recommendation and your critique are well aligned [1, 15]. It is difficult for an editor to use a review where either

- there is a recommendation to reject, but blandly positive or non-committal criticisms; or
- a recommendation to accept, but strong negative feedback.

Avoiding these cases seems obvious, and yet not all reviewers do.

5. Concrete comments

The critique of a paper should be (as noted above) concrete and constructive wherever possible [15]. In order to make my own comment more concrete I provide the following examples:

- don't just say "it's wrong" explain why;
- don't say "this work has been done before" without a citation;
- don't say "some claims are questionable." Any claim can be questioned, even if the answer is always "It's correct!" [26].

Also be specific: say exactly where and what should be improved (page, column and line number if possible).

Many other example of concrete comments could be added. One could write an entire report just on this topic. However, I would rather introduce a thought process that would be helpful in general. I aim to think "what will or should the authors do with this comment?" The answer is that authors will "triage" the comments in some variant of the following manner:

- the comment is easy to fix, so just do it;
- the comment is hard to fix, but the reviewer raises an important point, and so this also should be assiduously addressed;
- the comment is confusing and hard to understand, so do something and pretend you have fixed the possible problem;
- the point raised is essentially unbounded (e.g., discuss the paper in the context of Western philosophy in general and particular), or at the very least is about a different paper than the one we wrote, so we will argue for why the paper should not be changed in that way; or
- the comment is incorrect, but we have to find a polite way to say that without starting an argument with the reviewer.

See, for instance, [27] for more detail, but the point I am trying to make is not how to revise your paper but that in each case the authors will do something to "fix" the paper, but it might not be what you as a reviewer wanted done. Vague and abstract comments result in confusion and poor responses. *Clear, constructive comments result in the right behaviour from the authors.*

You may hope that a paper you don't like is going to be rejected, and hence it doesn't matter what the author does, but you need to provide helpful criticism because other reviewers (and the editor) may decide otherwise. More generally many papers (even ones I don't like) get accepted somewhere.

The last point – how authors address incorrect comments – is interesting. A bad response is a knee jerk complaint about the reviewer. A mediocre (but common) response is to (as stated) attempt to politely circumvent the comment. A really good response, however, is to think deeper about why the misunderstanding occurred, and attempt to clarify the manuscript. This is still a good outcome. As a reviewer, you don't have to be 100% correct, as long as you listen to reasonable return arguments, and your editor interprets the ensemble of reviews appropriately. However, this is not an excuse to be careless!

6. What Not to Do

6.1. Avoid bias

The authors' reputations, and your feelings about the authors, whether they be positive or negative, are not relevant to the review¹¹ [1, 2, 28–30]. I have heard it said by senior editors that they look more favourably on a paper by X (where X was a very famous statistician). But this type of prejudice and discrimination – *because that's what it is!* – leads to negative outcomes like disadvantaging early-career researchers and countries without large funding for research. Bias in reviewing is, however, alive and well, and much more complex than you might think [31, 32].

Many conferences now have a strict "double-blind" or masked review policy whereby the authors are anonymised in submissions to avoid overt as well as unconscious bias¹². They recognise that even the appearance of bias can have negative effects [28]! On the other side, the only good reason to preserve anonymity of *reviewers* is to preserve their ability to write a frank review without fear of reprisal.

Reiterating this point because it is *very* important, the authors of the paper should not matter. Neither should the institution and research group to which the authors belong. If it is not entirely obvious issues such as race, gender and religion of authors are completely out of bounds. The content is the thing being assessed. Bad content by famous people should be rejected. Good content by unknowns should be accepted. Thus reducing bias increases quality [2].

Avoiding bias is hard. Anonymising authors is one approach to enforcing a higher degree of *fairness* on a system that is implemented by (mostly) well-meaning but flawed humans. The reality and appearance of fairness is a crucial issue for the reputation of science as a whole. Genuine fairness is so difficult to achieve that some organisations are even considering funding by lottery as a mechanism to avoid bias [33] and other problems.

However, blinding is not universal and you are likely to be asked to review papers that are not anonymised. Some advice is provided in [31]. Apart from advice that is generally consistent with the practice of good reviewing (put sufficient care and professionalism into it) they suggest you "Challenge yourself with thought experiments: would your evaluation change if the investigator were of a different gender (or race, or from a different institution)?"

6.2. Don't be a copy editor

A reviewer is not a copy editor [8]. Your job is not to correct typographical, grammatical and spelling mistakes. Inevitably we end up doing this to some extent, but I think of it as a favour to the author, not my job as a reviewer. And if a manuscript is replete with errors, then I will not review it and it gets sent back with only the comment that it is not in a suitable state for review.

6.3. Don't forget Hanlon's razor

Hanlon's razor says "Never attribute to malice that which is adequately explained by stupidity." It is useful to keep in mind. More often than not a problem is the result of an oversight rather than deliberate misconduct. You should not overlook ethical considerations, but be aware that accusations of unethical behaviour are usually treated *extremely* seriously. A careless or ill-thought statement can have impacts on people's lives and careers and often mistakes are accidental, minor and easily correctable with some thoughtful advice.

6.4. Don't avoid reviewing because you don't know everything

A reviewer does not have to review every component of a paper [9]. Sometimes a review's expertise is needed for a particular component of the paper. At an editor's discretion, a reviewer may work on just this component, otherwise it is exceedingly hard to obtain reviews for multidisciplinary papers. Reviewers should be clear about their areas of competence, however [13].

¹¹ The authors may matter in some other contexts. For instance some grant applications specifically list the reputation of the authors as a criteria.

¹²There are objections to double-blind reviewing, for instance as detailed in [2, 28], so it is not yet used everywhere but its use appears to be increasing.

6.5. Don't avoid reviewing because you aren't senior enough

A reviewer does not have to be a senior academic. It is beneficial from all points of view to have junior staff [2, 9] and even students [34] participate in the review process. Junior staff often bring a fresh perspective and energy to the process [35]. They also avoid some of the problems with old fogeys like me having inflexible, fixed points of view. Junior staff can also gain valuable education in what it takes to publish papers by participating from the other side [2]. For this reason I don't like the commonly used term *peer* review. That can put people off participating.

It is important for junior staff to seek advice from mentors when they are learning, and this is one situation where I believe confidentiality may be bent, with the consent of the editor [3].

Some conferences now run a *shadow program committee* where PhD students and junior researchers are given the submitted papers for an important and high-quality conference (with permission from the authors) and asked to conduct their own review processes (with some mentorship). Engaging these younger people to read high-quality submissions is better than just handing them random papers, and learning the process together has many benefits. Doing so in a low-stakes system where the results are not critical to anyone's career means students can afford to make mistakes to learn more effectively. However, it is a great deal of work to organise a whole parallel process and so shadow PCs are relatively rare still.

6.6. Don't self promote

Your review should not be self-promoting. Think very carefully about suggesting to an author that they should cite one or more of your own papers [13] or otherwise push your own projects. You don't need to avoid telling an author about your work entirely, but think about how it would look if the review were made public.

6.7. Don't show off

A reviewer is not there to prove themselves. You don't have to seem smart or make comments to prove you did something. Just do a good review.

6.8. Others ...

There are more bad behaviours reviewers can indulge in than I can list here or in any document of reasonable length. But a good summary of many of the others is given in [26].

7. Other Issues

7.1. Become efficient

Peer review is not an old system. Although the idea stretches back to the 17th century and the foundations of learned societies it did not become so widely used until after WWII [33], for example for distribution of funding. It is now used to distribute some tens of billions of dollars in the US alone. The system has become increasingly demanding over time. Many organisations have increased the number of reviews they seek per manuscript or application over time, to improve the quality of their process in the face of increasing competition. And academics are expected to publish more and apply for more grants in the increasingly complex and competitive environment. The result is that everyone is expected to do more.

As a consequence you may end up writing very many reviews. I have quite literally written hundreds. This cannot be allowed to take over your life. You must become efficient.

The time taken to write a review is like the length of a review: it varies tremendously, from a couple of hours for a very clear cut and easy to review paper to some that have taken me days to properly review.

However, like any other task, you become faster with practice. You should also begin to think about processes that help you.

And while due care is needed, I do not try to craft the sentences of a review with the care I might in my own manuscripts. It is somewhat ironic and always annoying, to have a reviewer complain about *your* writing in *their* broken English. But remember, a published paper is a public document that may be out there essentially forever. The review is a transient document designed to help your paper get published. It does not need to meet the same standard of writing care (though still must be well enough written to be understood clearly).

7.2. Report your confidence

A critical issue for an editor attempting to resolve contradictory comments (this happens more than you would think) is to have information about the level of confidence you have in your review. It is desirable that all reviews are conducted by true experts in the particulars of a manuscript. However this is often impractical. Often in a niche or rapidly-evolving area, there are few such experts. And often they are busy with many other reviews or have serious conflicts of interest with the current authors. Less expert (but hopefully still knowledgeable) people are sometimes tapped. The key issue is for the editor to understand the limits of your expertise, particularly with reference to sections of a work that are out of your field.

7.3. Comments to the editor

Sometimes a review form includes a section called "comments to the editor," or something similar. This is a section that can be used to write confidential comments that go to the editor, but not the author. This section should *not* be used other than for comments such as your confidence in your review (such as described above), or to document potential conflicts of interest. An author has the right to understand as much as possible the information that the editor has used to form a decision about their paper. They have a *moral right* to transparency. They probably have the right (in some jurisdictions) under freedom of information laws to obtain such comments. Don't use this to hide information from authors. It's sneaky, underhand and not the intention of this part of a form.

7.4. Review the paper you are reading

A paper should be assessed on its merits. This seems obvious, but often reviewers seem to review some other paper that they wanted to read, and complain that the current paper is not that. Cormode calls this "moving the goal posts" [26], but its worse than this – you are essentially saying that they are playing the wrong game altogether. The paper should say what it is about, and be assessed on that basis, not what you as the reviewer would like it to be.

7.5. Be on time

A serious impediment to the rapid advancement of science is simply how long it takes to get work published [36]. Excessive delays can also damage peoples' careers [1]. A review should be timely [1, 3, 13, 36]. Different venues have different expectations, but reviewers should aim to adhere to the expected timeline, and agree only to reviews such that you believe this is possible [12, 13]. But don't be lazy. The system only works when enough people participate.

7.6. Differences by type of review

For the most part this report has treated all reviews (journal, conference, grant proposals, ...) as the same. There are many commonalities but they are not entirely the same. For instance:

- Many grants are not submitted anonymously, and indeed the quality of the applicant is often one of the criteria to
 be assessed. Race and gender bias must still be avoided, but other characteristics of the authors such as their past
 publications are likely to be relevant.
- Conferences have wildly varying expectations on reviewers: some expect reviewers to cover up to 50 papers in a short amount of time. This imposes limits on what a reviewer can do. Other conferences expect reviews of a quality that meets the standard of any journal. Other conferences review proposals for talks rather than actual papers. You have to know the venue to understand what it will require.
- Reviewing a PhD or Masters thesis (yet another task in a similar vein) is usually not anonymous, *i.e.*, the student will be (at least ultimately) aware of who the reviewer is. If you follow my advice (see above) lack of anonymity should not change your approach to reviewing, but a PhD is a much larger document than a typical journal paper, and so the review is likely to be correspondingly longer.

There are other differences [1], but if you understand the motivations behind this report they should be deducible.

7.7. Be objective

Finally, but not unimportantly let us reiterate the ultimate goal of objectivity. A review should be objective as far as possible. A reviewer must be careful how much of their own "knowledge" (actually their own opinion) they bring to the review. You must bring something, or you fail to bring your expertise, but do not allow your beliefs and expectations to overrule the findings of the paper. Don't make a decision about the paper, and then justify it. Build the decision from the paper and then explain it.

8. Conclusion 424

427

428

429

430

431

432

434

435

436

439 440

441

443

444

446

447

448

449

451

452

454

455

456

457

458

459

460

462

463

464

465

467

468

470

471

473

475

476 477

Many reviews are written badly and this is rarely treated with the seriousness it deserves. However, an author puts their life and soul into their work. To treat it carelessly is an abuse of power, and should be considered almost criminal¹³.

I myself am not perfect (yet), but I am trying to be better. Many people contributed to my understanding of the review process over the years, but this report contains the resulting opinions of only one academic. It was written at least in part to help me firm up my own ideas about writing reviews, and thus improve my own standard of conduct.

This report is backed up in places [1, 3, 10, 13, 26, 37, 38], and it is interesting how much commonality there is in recommendations from different fields of research.

However, some of my opinions are clearly not shared by all academics or research bodies. I have seen much evidence of that fact. You are welcome to disagree with my opinions; all I ask is that you have a good reason; a reason that you can articulate, and are willing to state in public.

Above all, please follow the "Golden Rule" of reviewing, McPeek et al. [36] "review for others as you would have others review for you."

References 438

- [1] Alan Jay Smith. The task of the referee. Computer, 23(4):65-71, 1990. http://www.eecs.berkeley.edu/Pubs/TechRpts/1989/6154 html
- [2] Kathryn S McKinley. More on improving reviewing quality with double-blind reviewing, external review committees, author response, and in person program committee meetings, 2015. http://www.cs.utexas.edu/users/mckinley/notes/blind.html, accessed Oct 9, 2019
- [3] Dale J. Benos, Kevin L. Kirk, and John E. Hall. How to review a paper. Advances in Physiology Education, 27(2):47-52, 2003.
- [4] Nyssa J. Silbiger and Amber D. Stubler. Unprofessional peer reviews disproportionately harm underrepresented groups in STEM. *PeerJ*, 2019. DOI: 10.7717/peerj.8247.
- [5] Bill Sullivan. If the script for 'Avengers:Infinity War' was peer-reviewed. *PLOS Blogs*, May 2018. https://blogs.plos.org/scicomm/2018/05/10/if-the-script-for-avengers-infinity-war-was-peer-reviewed/.
- [6] Bill Sullivan. If the script for 'Star Wars: The Last Jedi' was peer-reviewed. PLOS Blogs, March 2018. https://blogs.plos.org/scicomm/2018/03/13/if-the-script-for-star-wars-the-last-jedi-was-peer-reviewed/.
- [7] Douglas P. Peters and Stephen J. Ceci. Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences*, 5(2):187195, 1982. doi: 10.1017/S0140525X00011183.
- [8] Mathew Stiller-Reeve. How to write a thorough peer review. Career column, Nature, October 2018. https://www.nature.com/articles/d41586-018-06991-0.
- [9] Thomas M. Annesley. Seven reasons not to be a peer reviewer and why these reasons are wrong. Clinical Chemistry, 58(4):677–679, 2012. ISSN 0009-9147. doi: 10.1373/clinchem.2012.182618. URL http://clinchem.aaccjnls.org/content/58/4/677.
- [10] Mark Allman. Thoughts on reviewing. ACM Computer Communication Review, 38(2), 2008. Editorial Contribution.
- [11] Thomas M. Annesley. Now you be the judge. Clinical Chemistry, 58(11):1520-1526, 2012. ISSN 0009-9147. doi: 10.1373/clinchem.2012. 195529. URL http://clinchem.aaccjnls.org/content/58/11/1520.
- [12] Reviewing a manuscript: An overview of the OSA peer review process. Technical report, The Optical Society, 2016. osapublishing.org.
- [13] Kevin D. Haggerty. How to write an anonymous peer review. The Chronicle of Higher Education https://www.chronicle.com/article/How-to-Write-an-Anonymous-Peer/131475?cid=rclink, Haggerty April 2012.
- [14] Australian code for the responsible conduct of research. National Health and Medical Research Council/Australian Research Council, 2018. https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018#block-views-block-file-attachments-content-block-1.
- [15] Thomas M. Annesley. Writing an effective manuscript review: The 6 "Be's" to success. Clinical Chemistry, 59(7):1028-1035, 2013. ISSN 0009-9147. doi: 10.1373/clinchem.2013.208280. URL http://clinchem.aaccjnls.org/content/59/7/1028.
- [16] Peter Thrower. Eight reasons I rejected your article. *Elsevier Connect*, September 2012.
- [17] Brian A. Nosek, Jeffrey R. Spies, and Matt Motyl. Scientific utopia: Ii. restructuring incentives and practices to promote truth over publishability. Perspectives on Psychological Science, 7(6):615–631, 2012. doi: 10.1177/1745691612459058. URL https://doi.org/10.1177/1745691612459058. PMID: 26168121.
- [18] Jonathan B. Buckheit and David L. Donoho. *Wavelets and Statistics*, volume 103, chapter WaveLab and Reproducible Research, pages 55–81. Springer: Lecture Notes in Statistics, 1995.
- [19] Ferric C. Fang, Anthony Bowen, and Arturo Casadevall. Research: NIH peer review percentile scores are poorly predictive of grant productivity. *eLife*, 2016. https://doi.org/10.7554/eLife.13323.001.
- [20] Call for papers, acm internet measurement conference, ethical considerations. https://conferences.sigcomm.org/imc/2019/cfp/, 2019.

¹³ Some organisation, such as the Australian Research Council Research explicitly describe failure to conduct peer review responsibly as misconduct [24].

478 [21] Cope ethical oversight. https://publicationethics.org/oversight, accessed August 31, 2020.

484

499

500

511

- 479 [22] David Goodstein. On Fact and Fraud: Cautionary tales from the front lines of science. Princetone University Press, 2010.
- [23] Daniele Fanelli. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLOS ONE*, May
 2009. https://doi.org/10.1371/journal.pone.0005738.
- 482 [24] Australian research council research integrity policy, June 2019. https://www.arc.gov.au/policies-strategies/strategy/ 483 arc-research-integrity-and-research-misconduct-policy.
 - [25] Ethics in publishing: Instructions to authors. Elsevier, https://ees.elsevier.com/journal/ELS_Plagiarism.html, July 2015.
- ⁴⁸⁵ [26] Graham Cormode. How not to review a paper: The tools and techniques of the adversarial reviewer. *SIGMOD Rec.*, 37(4):100–104, March ⁴⁸⁶ 2009. ISSN 0163-5808. doi: 10.1145/1519103.1519122. URL http://doi.acm.org/10.1145/1519103.1519122.
- 487 [27] Bill Sullivan. Revising your first scientific manuscript. PLOS Blogs, April 2018. https://blogs.plos.org/scicomm/2018/04/24/ revising-your-first-scientific-manuscript/.
- 489 [28] Richard Snodgrass. Single- versus double-blind reviewing: An analysis of the literature. ACM SIGMOD Record, 35(3), 2006.
- 490 [29] SecDev: FAQ on double-blind reviewing, 2019. https://secdev.ieee.org/2019/double-blind-faq/#qg1, accessed Oct 9, 2019.
- 491 [30] Mark Allman, Olivier Bonaventure, Anja Feldmann, Dina Papagiannaki, and Darryl Veitch. ACM Sigcomm IMC statement on double blind 492 reviewing, 2019. https://www.sigcomm.org/content/imc-statement-double-blind-reviewing, accessed Oct 9, 2019.
- 493 [31] Anna Kaatz, Belinda Gutierrez, and Molly Carnes. Threats to objectivity in peer review: the case of gender. *Trends in Pharmacological*494 *Sciences*, 35(8):371 373, 2014. ISSN 0165-6147. doi: https://doi.org/10.1016/j.tips.2014.06.005. URL http://www.sciencedirect.
 495 com/science/article/pii/S0165614714001084.
- 496 [32] Anna Kaatz, You-Geon Lee, Aaron Potvien, Wairimu Magua, Amarette Filut, Anupama Bhattacharya, Renee Leatherberry, Xiaojin Zhu, and
 497 Molly Carnes. Analysis of National Institutes of Health R01 application critiques, impact, and criteria scores: Does the sex of the principal
 498 investigator make a difference? Acad Med., 91(8), 2016. doi:10.1097/ACM.000000000001272.
 - [33] Sandra Bendiscioli. The troubles with peer review for allocating research funding: Funders need to experiment with versions of peer review and decisionmaking. EMBO Reports, 20(12), 2019. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6893288/.
- 501 [34] J.W. Navalta and T.S. Lyons. Student peer review decisions on submitted manuscripts are as stringent as faculty peer reviewers. *Adv Physiol Educ.*, 34(4), 2010.
- 503 [35] Virginia Gewin. Rookie review. *Nature Jobs*, 478:275-277, 2011. https://www.nature.com/naturejobs/science/articles/10.
 504 1038/nj7368-275a.
- [36] Mark A McPeek, Donald L DeAngelis, Ruth G Shaw, Allen J Moore, Mark D Rausher, Donald R Strong, Aaron M Ellison, Louise Barrett,
 Loren Rieseberg, Michael D Breed, Jack Sullivan, Craig W Osenberg, Marcel Holyoak, and Mark A Elgar. The golden rule of reviewing.
 The American Naturalist, 173(5):E155–E158, 2009.
- 508 [37] Mark Allman. A referee's plea. http://www.icir.org/mallman/plea.txt, accessed Sept.6th, 2018, 2001.
- [38] Mark Allman. What ought a program committee to do? In USENIX Workshop on Organizing Workshops, Conferences, and Symposia for
 Computer Systems (WOWCS), 2008.