Peer Replication

Samuel J. Lord[1,✉], **Arthur Charles-Orszag**[1], **Kristen Skruber**[1], **R. Dyche Mullins**[1], **and Anders Rehfeld**[2,✉]

[1]UCSF & HHMI, 600 16th St, Genentech Hall N314, San Francisco CA 94143
[2]Department of Growth and Reproduction, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark

ABSTRACT: To help end the replication crisis and instill confidence in our scientific literature, we introduce a new process for evaluating scientific manuscripts, termed "peer replication," in which referees independently reproduce key experiments of a manuscript. Replicated findings would be reported in citable "Peer Replication Reports" published alongside the original paper. Peer replication could be used as an augmentation or alternative to peer review and become a higher tier of publication. We discuss some possible configurations and practical aspects of adding peer replication to the current publishing environment.

Correspondence: ***samuel.lord@ucsf.edu & rehfeld@sund.ku.dk***

Introduction

Readers of the scientific literature face at least three major challenges: 1) there are too many papers for a single human to read, even in highly specialized fields; 2) most published research is unlikely to be reproducible (Drude et al., 2021; Errington et al., 2021), either because the original experiments are flawed or because the methods are not sufficiently described; and 3) peer review has proven burdensome to both authors and reviewers and yet ineffective at screening out low-quality science or outright fraud (Baxt et al., 1998; Schroter et al., 2008). While peer review likely increases the quality of manuscripts to some degree, often there is little difference in the core elements of a preprint and its final published form (Carneiro et al., 2020). One reason for this may be that peer review evaluates a manuscript based on the opinions of very few people at the beginning of its lifetime rather than its ability to stand up against the key aspects of the scientific method—orthogonal testing and replication. Moreover, the current publishing system introduces perverse incentives: authors are motivated to present splashy but not necessarily robust findings and there is nothing that prevents reviewers demanding superfluous controls or entirely new experiments (Cleaver et al., 2023).

Publishers and preprint servers are experimenting with tweaks to the system that take aim at some of these problems (ASAPbio, 2019; Sever, 2023): eLife is piloting an approach where peer reviewers do not accept or reject papers, Review Commons and Peer Community In perform journal-blind peer review, PubPeer hosts post-publication reviews, and Lifecycle journals propose to evaluate a manuscript as it matures (Nosek, 2020). Pre-registering clinical trials has helped to reduce problems of p-hacking and post-hoc data slicing, but similar checks on robustness haven't become popular in the basic sciences, where exploratory studies are not as amenable to pre-registration. Here, we introduce an augmentation (or alternative) to peer review, termed "peer replication," in which fellow researchers attempt to reproduce the key findings of a manuscript, ensuring trustworthy science as well as more detailed reporting.

Schemes similar to peer replication have already been implemented to some extent. For example, the Institute for Replication arranges "replication games" workshops to replicate selected papers within economics and political sciences, and publish replication reports in meta-papers (Brodeur et al., 2023). Epic Research arranges clinical trials with multiple independent groups running the experiment in parallel (Epic Research, 2020), and some complex datasets have been independently analyzed by multiple teams (Aczel et al., 2021; Wagenmakers et al., 2022). For over a century, the journal Organic Syntheses has ensured that the instructions it publishes are easy to follow and actually work by repeating the synthesis in an editor's laboratory (Kamm et al., 1921). While synthetic chemistry is especially amenable to peer replication, many techniques in other fields—Western blots, chemical analysis, flow cytometry, transient transfections, reanalysis of raw data, etc.—could be subject to the same treatment.

A Proposed Process of Peer Replication

1. Journal editor selects a submitted manuscript or an already peer reviewed paper for peer replication. Authors would opt-in to the peer replication track during submission.

2. The editor invites researchers to serve as "peer replicators" on the manuscript, which means that they are invited to replicate selected key experiments or analyses in the manuscript. In consultation with the editor, peer replicators decide how to do so and can involve relevant co-authors. The replication plan would then be pre-registered.

3. After consultation with the editor, the authors of the original manuscript can share reagents (especially those that may not be commercially available) with peer replicators, assist with questions on the experimental design, and allow access to specialized equipment or custom software if necessary for performing peer replication experiments. The editor's role here

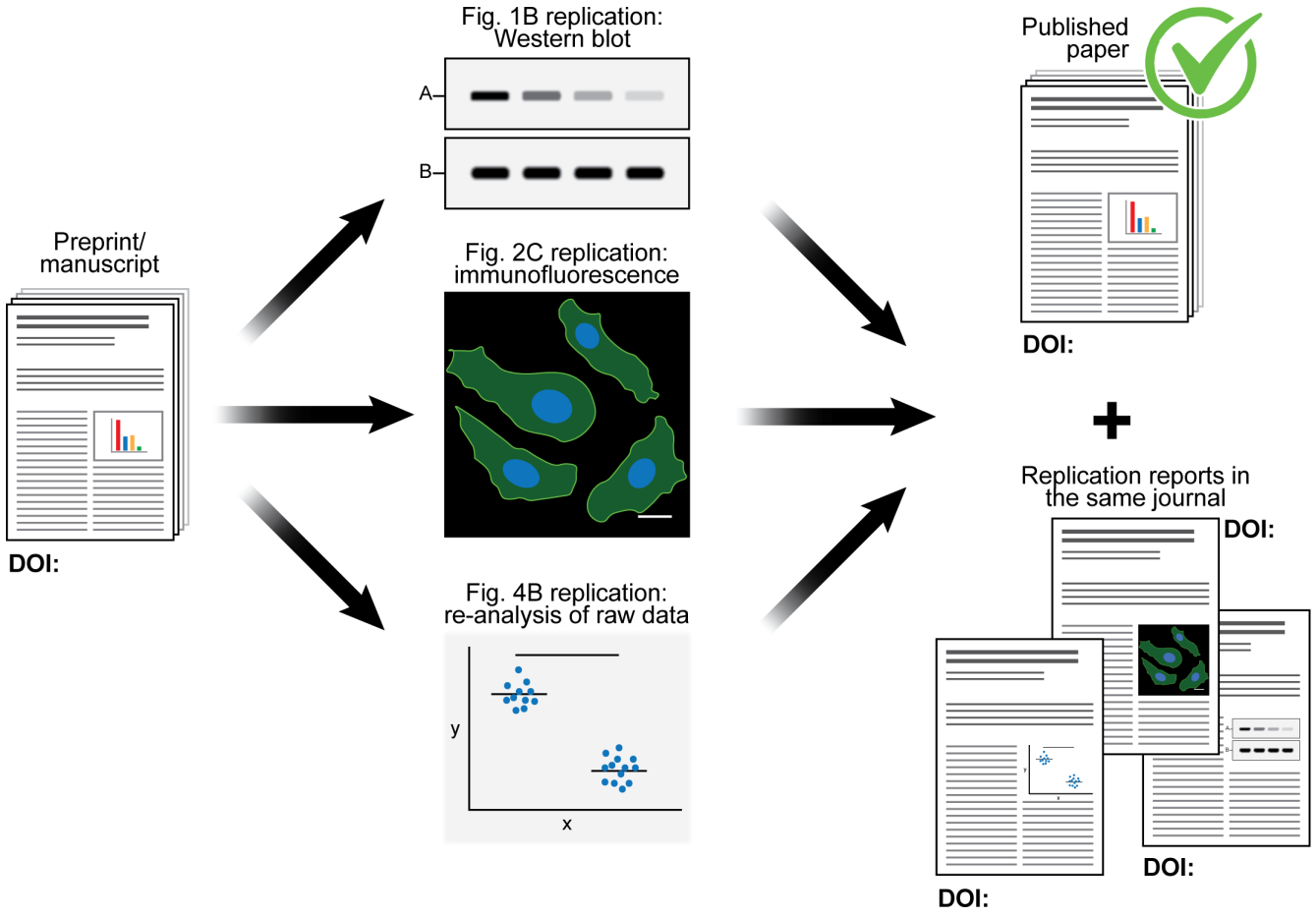

Fig. 1. A scheme summarizing an example Peer Replication process. Left: Manuscript is selected for peer replication by editor and peer replicators are invited to replicate key experiments or analyses. Center: Key experiments and analyses are replicated by peers. Right: Peer Replication Reports are published in the journal together with the published original paper. The journal would then indicate that the published paper has been replicated. This process could take place within a journal or be organized independently by a 3rd party. This could also occur post-publication and with or without traditional peer review.

would be to maintain independence between the labs while also promoting expediency.

4. Once a replication attempt is completed, the results are published online alongside the original manuscript in the form of a short and citable "Peer Replication Report" with its own DOI.

5. Once all Peer Replication Reports are published, the manuscript would receive a "Peer Replicated" label. Journals could develop their own policies for dealing with failed replication attempts. Some may simply label the manuscript "Not Peer Replicated," while others may reject the manuscript and only publish the Peer Replication Reports.

The process proposed above is only one potential model. Peer replication could be implemented outside of journals altogether and instead be arranged by an organization like ReviewCommons (ASAPbio, 2019) prior to submitting for publication. Or peer replication could take place post-publication, with the Peer Replication Reports either being published at the same journal, at a dedicated sibling journal, or at a 3rd party database that certifies rigorous research, similar to the "publish, review, curate" model (Krummel et al., 2019; Science Colab, 2023; Sever, 2023; Stern and O'Shea, 2019). In fact, the simplest implementation would be for journals to be much more willing to publish brief replication reports when they are submitted and link them to the original paper.

Peer replication could supplant or augment traditional peer review and become a third tier of publication: i) preprints, ii) peer-reviewed papers, and iii) peer-replicated papers (with or without prior peer review). Journals could thus choose to implement peer replication in multiple different ways to achieve the goal of elevating reproducibility as a key factor that is communicated to the readers.

Incentives Aligned

The incentive for a researcher to volunteer their lab's time and resources to try to reproduce someone else's experiment

would be simple: credit in the form of a citable published Peer Replication Report in the same journal as the original manuscript. Unlike peer review, the referees will receive compensation for their work in the form of citations and another publication to include on their CV. To minimize the burden, peer replication would need to be initially limited to simple experiments using assays the replicating lab already uses.

Authors would be willing to subject their work to peer replication for multiple reasons. Above all, most scientists want to publish real findings and witnessing a colleague repeat your results is rewarding in itself. Additionally, the process would in some ways be easier than traditional peer review, which puts a huge burden on the authors to perform additional experiments and defend their work against reviewers. Peer replication turns the process on its head, with the referees doing the work of validating the manuscript's findings. Finally, if funding agencies and promotion committees favored replicated publications, authors would be further encouraged to participate in peer replication.

Ancillary Benefits

A successful replication is clearly superior to opinions of peer reviewers based on reading alone, but the peer replication process would introduce other benefits. For example, it would require that adequate protocols are provided to the scientific community if the findings have any chance of a successful replication. Additionally, peer replication would address the issue that replication studies and negative results have generally been difficult to publish (National Academies of Sciences, 2019).

Furthermore, peer replication transforms the adversarial process of peer review into a cooperation among colleagues to build scientific rigor. Another set of eyes and brains on an experiment could introduce additional controls or alternative experimental approaches that would bolster and even expand the original finding.

This approach also encourages sharing experimental procedures among labs in a manner that can foster future collaborations, inspire novel approaches, and train students and postdocs in a wider range of techniques. Too often, valuable hands-on knowledge is sequestered in individual labs; peer replication would offer an avenue to disseminate those skills.

Finally, peer replication would reduce fraud. It would be nearly impossible for a researcher to pass off fabricated data or manipulated images as real if other researchers actually attempt to reproduce the experimental results.

FAQs

What about clinical trials, complex new methods, particle physics, animal research, long-term studies, or other complicated or expensive experiments? A lot of experiments will not be feasible to attempt a one-to-one replication. But most papers will have some core aspects that can be peer replicated. While it would not be possible to repeat a clinical trial, referees could run independent analyses on the raw data (Wagenmakers et al., 2022). Similarly, a peer may not be able to build a new microscope design from scratch, but they may be able to bring their own samples to the author's lab and perform their own imaging to test the equipment (Millett-Sikking and York, 2019). Editors—in consultation with the authors and referees—will determine the set of key experiments that will undergo replication, balancing rigor and feasibility. Those findings would then be highlighted in the main paper as well as the Peer Replication Reports. Ultimately, some fields will be more amenable to this process and may provide good testing grounds for peer replication. While this may leave some of the most complex experiments unreplicated, it will be up to the readers to decide how to judge the paper as a whole. Furthermore, peer replication will encourage authors to design experiments and describe the protocol adequately to ensure replicability. It should also be noted that modern clinical trials generally meet much higher standards for pre-registration, reporting, and sample sizes than the typical research article found in most journals. In fields where replication is especially challenging, rigor must be achieved in other ways.

Won't this make publishing even slower? Journals may choose to simply publish peer-reviewed manuscripts with an "Awaiting peer replication" label, which would be no slower than the current system. Regardless, editors will need to select only key experiments for replication, minimizing the time and effort burden on everyone involved. Many experiments could be replicated in a few weeks at most. Peer review is already a slow process (Royle, 2020), often taking months and multiple rounds of revision, and in many cases ending in rejection. With the rapid adoption of preprints across many domains of science, new results can be disseminated quickly, so the need to accelerate the publication of the final form is alleviated. If in practice it does take more time than traditional peer review, peer replication should increase the strength of the findings enough that it will be worth it.

Who will fund the cost of the replication experiments? Ideally, funding agencies and research foundations would create small on-demand grants to cover the costs of peer replication experiments. Institutes or funding agencies could even require that the work they fund is peer replicated and allocate earmarked funding for such peer replication studies. Initially, we hope to run a pilot where the replications are funded. Alternatively, the publication fees paid by the original authors could fund the replication, in which case the editor would need to carefully select only low-cost experiments. Some labs will be able to find the time and funds to run a few simple experiments and immediately get authorship on a paper, especially if existing assays and equipment can be reused. This may be particularly appealing to primarily undergraduate in-

stitutions and lab courses, which have many early-career researchers who would benefit from learning new techniques and building their list of publications.

What if a replication attempt fails? Pre-registering a replication plan could help insure rigor and fidelity of the peer replication process. A failure to replicate does not necessarily negate original findings. For instance, a replication attempt may fail due to important variables not accounted for in the methods. In that case, the editor could work together with both the authors and peer replicators to track down the problem, strengthening the methods section and technique for future experiments. Ultimately, editors would decide whether the failed replication is essential to the manuscript and/or if the attempt at replication was adequately undertaken. Importantly, the Peer Replication Report should be published regardless of the outcome of the replication experiments to ensure transparency and maintain an unbiased incentive for the peer replicators to accept performing the task.

What if the replicating lab doesn't try hard enough or isn't well-equipped to replicate the experiments? Peer replicators will be incentivized to perform solid work, because the Peer Replication Reports will be published with the names of the peer replicators on the report, meaning that the researchers undertaking the peer replication will have to stand by their results. Editors may choose to recruit multiple independent replicators for the same experiments, just like they typically solicit multiple reviews. Furthermore, any researcher disagreeing with the findings of a Peer Replication Report would be welcome to address this in future publications, fostering data-driven discussions of key research findings. Ultimately, it is still beneficial to the scientific community to be able to know what other attempts or approaches did not succeed.

Will the peer replicators be incentivized to report a positive replication even if the experiments fail? Few would be willing to commit fraud in order to bolster someone else's paper. Publishing fake results in the Peer Replication Report would require two or more independent labs colluding to fabricate data with none of the co-authors blowing the whistle, which seems unlikely.

Doesn't this mean that my competitor might learn some of my techniques? Any manuscript evaluated by the peer replication process can first be uploaded as a preprint, so the primary authors can claim priority. While primary authors may feel as if they are "giving away" their techniques to replicators, it will be under the banner of collaboration and strengthening the findings of the primary authors. Ultimately, widespread adoption of one lab's technique benefits the originating lab, and peer replication bolsters this process. Alternatively, the replicating lab (in consultation with the editor) may choose an alternative technique to answer the scientific question, adding orthogonal robustness.

Will all manuscripts be able to undergo peer replication? In addition to the fact that some studies are too large or complex to be amenable to peer replication, there would never be enough time or resources to repeat even all the simple experiments that are published in the thousands of journals that exist. Editors, and the availability of peer replicators, would ultimately dictate which papers get selected. Manuscripts submitted to prestigious journals would be more likely to get peer replicated, since the peer replicators get their Peer Replication Reports published in the same high-profile journal. However, in less prestigious journals, a replicator-initiated approach could be used, where researchers may apply to become a peer replicator on a particular part of an already published manuscript. Editors would then decide if these applicants are accepted as peer replicators. This would enable researchers who are normally not invited as peer reviewers to contribute to the process while at the same time providing them with an opportunity to get a first-authored publication.

Moving Forward

The next steps for this idea are to gather additional feedback from stakeholders, including researchers, institutions, funding agencies, and journals. And then we need a journal, preprint server, or 3rd party to experiment with implementing a version of peer replication either in addition to their current peer review model or as an alternative.

In the meantime, authors can add replication robustness on their own. First and foremost, researchers should repeat their experiments multiple times, perform their data collection and analysis blinded, and perform statistics on biological replicates (Eisner, 2021; Lazic, 2010; Lord et al., 2020). Researchers may also seek their own form of peer replication by recruiting collaborators and colleagues to perform replications of some key experiments and adding them as coauthors on the subsequent manuscript. Ultimately, authors, reviewers, editors, funders, and institutions celebrating work that has been replicated will help shift scientific norms towards robustness. While peer replication has the promise to help align incentives and cure the replication crisis, we can each contribute to robust science and reproducibility today.

Acknowledgements

This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (R01-GM061010 to RDM), by the Howard Hughes Medical Institute Investigator Program (RDM), by a BRIDGE - Translational Excellence Programme grant (to AR) funded by the Novo Nordisk Foundation (NNF18SA0034956) and by a Lundbeck Foundation Clinical Postdoc grant (R381-2021-1478 to AR). We thank Richard Sever, Prachee Avasthi, Brian Nosek, Tim Errington, Stuart Buck, Lil Fritz-Laylin, Wolfgang Draxinger, and John P. A. Ioannidis for helpful discussions and reviewing early drafts of this work.

Bibliography

Aczel B, Szaszi B, Nilsonne G, van den Akker OR, Albers CJ, van Assen MA, Bastiaansen JA, Benjamin D, Boehm U, Botvinik-Nezer R, Bringmann LF, Busch NA, Caruyer E, Cataldo AM, Cowan N, Delios A, van Dongen NN, Donkin C, van Doorn JB, Dreber A, Wagenmakers E-J. 2021. Consensus-based guidance for conducting and reporting multi-analyst studies. eLife 10. doi:10.7554/eLife.72185

ASAPbio. 2019. ReimagineReview – A registry of platforms and experiments innovating around peer review. ASAPbio.org. https://reimaginereview.asapbio.org

Baxt WG, Waeckerle JF, Berlin JA, Callaham ML. 1998. Who reviews the reviewers? Feasibility of using a fictitious manuscript to evaluate peer reviewer performance. Ann Emerg Med 32:310–317. doi:10.1016/s0196-0644(98)70006-x

Brodeur A, Dreber A, Hoces de la Guardia F, Miguel E. 2023. Replication games: how to make reproducibility research more systematic. Nature 621:684–686. doi:10.1038/d41586-023-02997-5

Carneiro CFD, Queiroz VGS, Moulin TC, Carvalho CAM, Haas CB, Rayêe D, Henshall DE, De-Souza EA, Amorim FE, Boos FZ, Guercio GD, Costa IR, Hajdu KL, van Egmond L, Modrák M, Tan PB, Abdill RJ, Burgess SJ, Guerra SFS, Bortoluzzi VT, Amaral OB. 2020. Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. Res Integr Peer Rev 5:16. doi:10.1186/s41073-020-00101-3

Cleaver O, Prince VE, Wallingford JB. 2023. We have seen the gatekeepers, and they are us. Dev Biol 501:A15–A17. doi:10.1016/j.ydbio.2023.06.016

Drude NI, Martinez Gamboa L, Danziger M, Dirnagl U, Toelch U. 2021. Improving preclinical studies through replications. eLife 10. doi:10.7554/eLife.62101

Eisner DA. 2021. Pseudoreplication in physiology: More means less. J Gen Physiol 153. doi:10.1085/jgp.202012826

Epic Research. 2020. About Us. EpicResearch.org. https://epicresearch.org/about-us

Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, Nosek BA. 2021. Investigating the replicability of preclinical cancer biology. eLife 10. doi:10.7554/eLife.71601

Kamm O, Marvel CS, Clarke HT, Davis AW. 1921. Alkyl and alkylene bromides. Org Synth 1:3. doi:10.15227/orgsyn.001.0003

Krummel M, Blish C, Kuhns M, Cadwell K, Oberst A, Goldrath A, Ansel KM, Chi H, O'Connell R, Wherry EJ, Pepper M, Future Immunology Consortium. 2019. Universal Principled Review: A Community-Driven Method to Improve Peer Review. Cell 179:1441–1445. doi:10.1016/j.cell.2019.11.029

Lazic SE. 2010. The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis? BMC Neurosci 11:5. doi:10.1186/1471-2202-11-5

Lord SJ, Velle KB, Mullins RD, Fritz-Laylin LK. 2020. SuperPlots: Communicating reproducibility and variability in cell biology. J Cell Biol 219. doi:10.1083/jcb.202001064

Millett-Sikking A, York A. 2019. High NA single-objective light-sheet. Zenodo. doi:10.5281/zenodo.3244420

National Academies of Sciences. 2019. Reproducibility and replicability in science. Washington (DC): National Academies Press (US). doi:10.17226/25303

Nosek BA. 2020. Bridging Divides and Promoting Diversity of Ideas for Knowledge Production in the Academy Proposal: Center for Open Science to the Charles Koch Foundation 2023, https://osf.io/x945r. Center for Open Science.

Royle S. 2020. Same Time Next Year: crunching PubMed data. Quantixed.org. https://quantixed.org/2020/05/08/same-time-next-year-crunching-pubmed-data/

Schroter S, Black N, Evans S, Godlee F, Osorio L, Smith R. 2008. What errors do peer reviewers detect, and does training improve their ability to detect them? J R Soc Med 101:507–514. doi:10.1258/jrsm.2008.080062

Science Colab. 2023. Biophysics Colab. https://www.sciencecolab.org/biophysics-colab

Sever R. 2023. Biomedical publishing: Past historic, present continuous, future conditional. PLoS Biol 21:e3002234. doi:10.1371/journal.pbio.3002234

Stern BM, O'Shea EK. 2019. A proposal for the future of scientific publishing in the life sciences. PLoS Biol 17:e3000116. doi:10.1371/journal.pbio.3000116

Wagenmakers E-J, Sarafoglou A, Aczel B. 2022. One statistical analysis must not rule them all. Nature 605:423–425. doi:10.1038/d41586-022-01332-8