Moving science into the *Open*: Scalable post-publication review with commissions, rich tools and quality over quantity

Author consortium (names released upon hitting signatory-selected thresholds)[*]

[*]Corresp: fiete@mit.edu, bwyble@gmail.com, nick.halper@neuromatch, d.goodman@imperial.ac.uk

Abstract

There is a broad awakening underway in the scientific reseearch community that the current journal publication system is broken. The problems range from publishing and reviewing delays, to misaligned incentives, to not enough feedback, to expense, to inaccurate and gameable metrics for scientific merit. We provide a coherent and scalable proposal, which we call *Open*, for how the modern scientific publishing landscape for the sciences could, or in our view should, look. This paper is a call-to-action to move science into the *Open*. We believe that the time is ripe to implement this change, and feasible from both technological and sociological perspectives. We envision a scientific publishing and review system that is not only open, efficient, and richer in context, content, and functionality, but is *delightful*.

The many burdens of our current system

Some of the most exciting discoveries in the basic sciences, particularly in biology, sit cloaked for a year or two after being written up before they are made public in a journal publication. Authors work their way sequentially down a list of journals to publish the paper, devoting time and resources on rewrites, length and formatting changes, resubmissions, responses, and on and on, down the chain, rather than working on the next important question. Thus society, after funding the labors, must wait years to view and partake of its fruits. The full body of experts in the field, best-placed to evaluate the work and use it rapidly to propel their next breakthroughs, must wait until the paper is out.

Referees are imperfectly selected (not least because editors tend to pick them based on (recent) past rather than current interests and expertise, but also for many other reasons [1, 2]) and thus they are not always best-suited to evaluate cutting-edge work. On the topic of free and largely unrecognized scientific contributions: Referees must themselves generate multiple rounds of comment and response. The number of referees asked to contribute their time, uncompensated, in this way grows with the number of journals the paper goes through in the process. Scientists seldom receive credit for their contributions in reviewing, and in the rare instances they do, this credit is not fungible. The comments contributed by reviewers are usually not made available to readers of the article, a loss of a significant scientific product.

Priority or precedence, the first touchstone of scientific credit assignment, is muddied when the "submission" date on a paper reflects the date of submission to the last journal in the chain, and then too, the date of the last revision for that journal, rather than the date of first submission. The citation record, the second touchstone of scientific credit assignment, is also strongly distorted by journals that artificially restrict the number of references in papers. This restriction serves well high-profile journals that wish to concentrate citations to a few high-profile papers, thus boosting their own impact factors. Restricting allowed citations also means that fewer original articles and relatively more review articles are cited, concentrating credit to the writers of the review articles, who tend to be senior scientists. It also means that credit to the primary literature is now more heavily dictated by the review article, thus one or a few individuals have shaped the future citations of the entire field. Given the snowballing dynamics of citation, many relevant, prescient, and excellent articles can be ignored and forgotten, again a loss to

the field. There is no reasonable mechanism for improving credit assignment in the present publication process, unless one expects that the few referees looking at a paper know and note all missing relevant citations.

The current publishing process is also highly inequitable: large fees for publishing and/or reading scientific articles put them out of reach for many scientists in less affluent institutions and countries.

These are all serious problems for young scientists: A system that is highly stochastic and slow can impede and derail careers. Simple (and thus gameable, by Goodhart's law) citation metrics like publication in glossy journals for faculty positions and promotions distort what science is pursued, encouaring the chase for fast glamor over durability. Errors and biases in the evaluation processes can further frustrate junior scientists. And, lack of access to publications holds back young researchers in less-privileged institutions. They are therefore, rather idealistically and surprisingly given the high risk sensitivity of their career stage, the most likely to support more open and efficient processes of publication and review [3].

From the larger perspective of scientific communication and progress, the aggregate costs to the scientific community and to society in general, accrued from literal costs and the delays and distortions over all scientific publications annually, are unconscionable. What is redeeming today about a system that was long-ago established for the simple reason that paper, printing, and distribution were slow and expensive, that more recently has been distorted by metrics and profit motives that are not aligned with the goals of science [4, 5], and that is slow and staggeringly expensive [6, 4, 7]?

Preprint servers solve some key problems

Started in 1991, the electronic preprint server arXiv [8, 9], now widely used across fields of Physics, Astronomy, Mathematics, and Computer Science, and Statistics and expanded to Biology (bioRxiv), is familiar across science, even to those who have not used it.

In several fields that have embraced the arXiv, the norm is to post a manuscript to the electronic preprint server as soon as it is in a form to be submitted for the first time to a journal. Typically, journal submission is delayed for a few days after posting to allow time for (typically emailed) feedback, including suggestions for citations from the scientific community. In turn, members of the community scan daily the postings as a form of good preprint server hygiene, rather like regular flossing or reading the newspaper. The benefit of this system is that feedback can be broad, arriving in principle from all those currently working on precisely the most relevant topics, who chose to read the preprint. Few targeting mechanisms for expertise and matching interests are likely to be more accurate than self-identification. If I, as a reader, notice that a highly relevant work (from me or others) has not been cited, it is common practice to directly email the author and mention it. The author may choose to include it, but now it is a conscious decision, and she does so with shared knowledge between herself and me about her decision. This model is a strong corrective force on the citation record: It is not so easy to neglect to cite similar and competing works.

Manuscripts posted to the arXiv are immediately citable, with a unique identifier and stable system for future access of the manuscript. In many fields, arXiv articles are freely cited, until a published version of the article is available. When there is a question of scientific priority post-publication, it is decided based on the arXiv date rather than the journal "submission" date. The existence of multiple versions of a paper on the arXiv, and another in a journal, need not be a problem for the citation record because it is easy to build a system that links a journal paper with its corresponding arXiv record. Revisions to an arXiv submission may be made ad nauseum, but because all submitted versions are accessible to readers, such revisions do not in in principle iterfere with the fair assignment of priority. (Certainly, any such system for priority is better than the present one, where all rounds of revision are private between journals and authors, and nobody knows the review and revision history or who really submitted a work first.) The accessibility of all revisions acts as a strong control on quality, at least for members of the scientific community, who tend to care what their colleagues think of their work.

The direct-to-market nature of arXiv posting and feedback means the removal of most of the friction associated with publication for science and society. My colleagues can read what I did, build upon it, criticize it, or actively ignore it, all in a timely way. Indeed, some fast-paced fields of physics (e.g., high-energy theory) have dispensed with publication in journals almost entirely, trusting rather in the community of peers, their direct feedback, and citations of the arXiv preprints. Junior scientists in high-energy theory, who may require the imprimatur of journals

on their CVs for tenure-track faculty position applications (not all departments require this of high-energy theorists anymore), go on to submit their arXiv manuscripts for publication. Nevertheless, this publication step could be viewed as a post-facto seal of approval at the back-end, rather than a gatekeeping step at the front.

This model of publication in high-energy theory might be unique because of the relatively small size of this community, as larger fields have not yet followed suit. However, it is possible that with additional mechanisms to augment the arXiv model, it will be possible to dispense with gatekeeping and review coordinated by traditional journals in larger fields as well, as others have noted and as we argue next. November 2013 marked the launch of the bioRxiv electronic preprint server for Biology manuscripts, with the same functionality and philosophy as the arXiv. The bioRxiv represents fine-grained subfields of biology, from animal behavior, neuroscience and synthetic biology to clinical trials, with room for expansion in subject areas as well as the deposition of certain datasets. Use of the bioRxiv is burgeoning, with exponentially growing submissions and a rapid shift, in some subfields, to reading and discussing bioRxiv papers rather than the slowly arriving journal articles in journal clubs.

Critical unsolved problems: What we're missing

The main missing components in the current preprint archive system are 1) high-quality, high-speed feedback and review, and 2) information-rich and customizable multi-layered systems to help readers sort through high volumes of papers and decide what to read.

Preprint archives provide a forum for posting comments about manuscripts, but the feature is sparsely used. Directly emailing authors permits fast feedback, but lacks the option of anonymity and public viewability. Download, page view, and alternative metrics based on social media coverage provide some information on "what's hot", but these quantities provide little detailed information about the (de)merits of the work, and the social-media-based metrics are also prone to the echo-chamber amplification dynamics common to those forums.

Presently, journals perform the review function, but do so at the cost of introducing massive friction and time-delays into the process, as gatekeepers at the front-end of publication. Moreover, reviews are usually private, and not available to readers. Additionally, there is no good method for giving credit to referees for their expertise and the hard work they currently provide pro gratis, on the basis of which the (mostly for-profit) journal system is currently run [4].

We call for switching to a bottom-up, democratic, and open mechanism for refereeing and referee credit assignment, augmented by a potentially still-useful role for journals in scientific publishing. The mechanism simultaneously provides broader feedback to authors, much richer context and content to readers, and a way to build highly customizable and versatile search and recommendations for all.

Open: Post-publication review with commissions and rich tools

The philosophy of *Open* builds on the advocacy of many scientists and on a growing consensus [10, 11, 12, 13, 14, 15, 16, 7, 3]: that one's entire field is the unique "right" adjudicator for one's work, that bottom-up solutions are more efficient than top-down ones in science as in the marketplace for goods, and that reductions of friction and increases in transparency and accessibility are the goals we should strive for in scientific publication.

In brief, the proposed *Open* system incorporates a preprint archive with full regular functionality, augmented with commissioned and non-commissioned fully-visible anonymous peer review by scientists with unique but anonymous identifiers who gain reputation for their refereeing and commentary, with enhanced mechanisms for search and discovery of papers aided by the existence of publicly visible and searchable reviews. The role for traditional journals is as gallerists or curators of papers, who would add value by compiling papers for focused themes bundled with commissioned and paid-for commentary articles and pedagogical inserts, and thus gain subscribers and advertiser revenue based more closely on the work they do. We believe this model solves the majority of problems raised earlier.

The process

The steps from arXiv to *Open* begin with posting a submission to arXiv or *Open* directly. Authors designate area(s), based on field and desired audience. Authors also add tags to their work, designating it as original research, review, or opinion, providing new methods, new results, etc. All submitted papers are immediately issued a unique citable identifier as already done on arXiv, and are considered published.

Each reviewer and author on *Open* will use an anonymous but unique account (e.g. linked to their Orcid handle) that verifies their institutional affiliation; this handle will be used for accruing reviewer reputation scores and for sending author/reviewer statistics to the author's/reviewer's landing page. Submitted papers are immediately open for reading and reviews by any user with sufficiently high "reputation" (reputation $\geq \theta_1$, as described below) and by solicited reviewers. After two weeks of such commentary or after sufficiently many reviews have been received as chosen by the authors, the authors have a month to post a response to the comments and if they wish a revised manuscript. Upon receipt of the author response, the revised manuscript receives a peer-reviewed stamp. All members of the community are now able to read the paper and (with a lowered threshold of θ_2 in reputation) to comment on it. They are also able to contribute to an "expanded citation list" for the article by suggesting related papers not cited by the authors, which can be appended to the article if the review itself receives strong enough reviews and is endorsed by the authors.

Commentary and ratings by our peers

Contributing thoughtful, constructive commentary and criticism takes considerable work; without a mechanism for apportioning credit that accrues across all comments posted by the person, participation rates stay low, as seen in existing comment forums for journal articles and preprint servers. A key addition in *Open*, which we believe will mean the difference between participation and non-participation by the community in reviewing or commenting on papers, is the ability of commentators to gain reputation and credit for their comments and referee reports, which would be used to assess their scientific contributions during hiring and tenure decisions.

Online scientific forums, such as Math Overflow and Math Stack Exchange – sites where any user can ask simple-to-research-level mathematics questions, and any user can spend some of her precious time helping to answer a question – have such mechanisms in place and manage to attract robust commentary, collaborative problem-solving, and expert advice-giving by practicing mathematicians ranging from graduate students to Fields medalists and everybody in-between. Some of us have posted non-trivial mathematics questions to the site and received extremely helpful answers within a day from people who don't know us.

How do they do it? The responses are rated, and responders accumulate reputation points based on how highly their past and present responses have been rated. Poor responses receive no points, crackpots get bad reputations over time, and the best responses and responders float to the top. The responders, with unique but potentially anonymous online IDs, now have an incentive to respond with thoughtful answers. Such a system, in which any registered member could post comments that are then scored by any other registered member, and in which reputation scores can keep increasing with well-rated comments, is the key for generating active, high-quality online feedback for each paper. Reputation and contributions are fungible: this information will be sent to members as certifications (details in FAQ below). These certifications can be used on CVs and tenure and promotion documents.

Value accretion

There are multiple ways in which outputs posted on *Open* gain value over time. First, each article gains visible context and commentary from reviewer comments. It can also be rated numerically along several dimensions by anyone posting comments, and the average of the article score on each of those dimensions, weighted by the reputation score of the reviewer, become multiple metrics associated with the article. For instance, papers could receive scores for originality, rigor, synthesis, code availability, dataset availability, topicality, conceptual contribution, etc.

Second, commenters can attach multiple tags to articles, identifying them as suggesting new methods, perspectives, conceptual insights, and so on. Readers can also suggest related works and references.

Third, authors can link articles to each other, creating a “meta” article. In these ways, articles continue to accrue content and value over time.

Better information for hiring and promotion committees than journal name proxy

A major argument for the importance of proxies of quality like journal name is that hiring and promotion committees have members who are not always experts in the area of the candidate’s research. Even when they take the effort to read the articles written by the candidate, they are in no position to properly contextualize and evaluate the methods, results, and impact of the work. Journal impact factor is then used as a crude estimate of the importance of the work. However, a better alternative is to provide evaluators with tools and rich data to evaluate the work of candidates than the name of the journal where it was published. With *Open* this takes the form of the open reviews, which will contextualize the key contributions (and limitations) of the work from the perspective of expert readers. In addition to free-form text in the reveiws, reader-added tags like "though-provoking conceptual insights", "durable", "confirmatory", "new technique", etc., will together paint a much more informative, detailed, and expert-guided picture of the work of candidates than journal name for non-expert members of hiring/promotion committees.

As noted above, reviewer activity by an individual is also quantified and certified, and can be used in hiring/promotion.

Quality over quantity in publication and review

Scientists voice two major concerns about post-publication open review. The first is the spectre of social media-like dynamics in paper review: that the number of reviews will snowball for a few papers, especially by well-known authors, to gain high reviewer repuation scores. The second is the perceived pressure toward publishing more papers, which they feel is growing and fear will be exacerbated without gatekeeping by journals. These are indirectly or directly concerns about volume over quality. A major goal of *Open* is to bend the curve toward quality over quantity. How can it do so?

First, each person will be limited to a maximum number of reviews per year – review writing will be a scarce (and fungible, see above) resource, gaining value from its scarcity, so that reviewers are thoughtful and judicious in selecting what they review. Second, every year each author will have a single star to award to their own output(s) per year ("If you read one paper from me this year, this should be the one."). Similarly, each reviewer will have one star each year (per year in which they reviewed >=5 papers) to award to the best paper they have reviewed that year. This single-star annotation can be a powerful signal of the quality of a paper. These steps strongly incentivize quality over quantity.

Third, reviewers of reviews will be urged to evaluate a review on the quality and newness of perspective relative to reviews already posted, discouraging redundancy or high multiplicity in reviews of the same paper. To do so, people scoring reviews of a paper will be asked about the quality of a review and whether it contributes a unique/novel perspective relative to what has already been written about the paper. Reviews that are solid in quality but do not contribute a new perspective relative to earlier reviews will receive a lower score.

Fourth, papers will be scored (along each dimension) by the weighted average of their review scores (on that dimension), with a separate standard error of measurement (SEM) score that is based on the volume of reviews and the scores of the reviewers. The volume score will provide the only "crowd" quantity metric on how much crowd attention a paper received, amongst a multitude of quality metrics.

Wading through the ocean: Layers of tools for sharing and discovery

The *Open* data (which would consist of the papers, reviews, tags, badges, dates, links to datafiles and videos, and all other metadata) would be highly searchable with fully user-customizable rich filters, together with additional collaborative filtering recommendation tools [17, 18]. Search criteria could be selected through various filters that can be applied to article content and article comments or reviews (e.g. “I would like to see all articles that have a 4 or higher rating from my favorite anonymous commentor with handle XYZ, on the tag of new hypothesis

and in a particular subfield, that moreover shares data from the paper") and other article metrics (weighed score on some dimension, highest mean ranking, number of citing articles, etc.).

The system would additionally allow and facilitate the layering on of recommendation systems by third-party systems that run on top of the same data, and we expect that an ecosystem of recommendation systems will organically arise through open-source platforms to serve this system.

A role for traditional publishers? Curation, overviews, and tool development

Finally, is there any role remaining for publishers under the post-publication review model? The answer can be a constrained "yes". Gallerists, curators and aggregators in museums and the commercial world cannot accept or reject the public release of an artwork, news item, or product, but they play a different vital role. Discerning curators and aggregators collect gems from the sea of possibilities, gallerists envision themes and organize shows with interpretive material, helping consumers discover prominent or new promising artists or artworks. Radio stations curate songs by genre and taste, and people choose to tune in. These gallerists, curators, and aggregators monetize their contributions even when they do not directly sell the products they highlight, by collecting advertising and subscription revenues from their followers. Scientific journals may occupy the same niche with respect to published *Open* papers.

In the new publishing model, we propose that journals be viewed as galleries and their editors as curators, showcasing what in their view are interesting (already published) works, and providing additional perspectives, commentary, reviews, overviews, and pedagogy. Although this is a part of what journals already do, at present they act primarily as gatekeepers for the release of papers. Science glossies could invite or commission authors of a recent article to provide a more condensed or accessible version of their work as a feature in their pages. Clearly, journals could not earn revenue from downloads of the original research articles because they are already freely available on *Open*; however, users and institutions would likely be willing to pay fees to journals if the journals add sufficient value through thoughtful curation and added commissioned content. Commissioned content and condensed versions of *Open* papers could generate revenue on a per-download basis, similar to the traditional fee structure that journals currently use to charge for content.

As importantly, journals, individuals, and any other third-party sites can design and provide customized search and recommendation formulae operating on top the *Open* data, by factoring the user's interests and various other metrics, tags, and words using collaborative filtering and related machine learning approaches [17, 18]. This product would be complementary to the highly customizable filters that users of *Open* could create for themselves.

In these ways, there would be a better alignment between the compensation of for-profit journals and the value that they would bring to the community, than in the current model where a substantial portion of their compensation derives from the original content provided free by scientists and the donated work of reviewing. Importantly, individual articles would be cited and referred to by their *Open* identifier even if they were featured in a journal, thus minimizing the use of journal publication metrics in promotion decisions. In other words, journals would exist as curators and their revenue model would hinge on their curatorial skill, but inclusion of an article in the journal would not mean that the article's citation is attached to the journal.

For featured articles, this is a flipped model of journal publication: instead of authors shopping serially down a list of journals to publish, the article is immediately published and instead journals queue up, soliciting or bidding to feature it. Authors can consider the bids in parallel until they select one (or none). Importantly, both selection steps take place post-publication, while scientists and the public have immediate access to the published work on *Open*. For the broader readership, journals would add value by putting works in perspective through overviews and collections, and this process can be slower without slowing down experts who need access to cutting-edge results.

In science, we have long functioned within a model where journals decide which manuscripts get published, because that is where the costs and bottlenecks have resided for hundreds of years. Now publication is cheap; the new bottleneck is human cognitive bandwidth. Journals might not merely survive but thrive if they serve by improving the problems with the new bottleneck, rather than trying to maintain a stranglehold on the production step.

To summarize, the proposed *Open* is a fully peer-reviewed article repository, in which each paper can accumulate value and content over time, in the form of the reviewer commentaries attached to it, the author responses,

and the revisions the paper undergoes. There are zero publication delays, and no delays attached to peer review, which is immediate and ongoing. And the model strongly incentivizes the contribution of reviews by allocating fungible credit to reviewers.

Ongoing efforts

An entire body of papers espouses related ideas for open review, open access, post-publication review, meta-reviewing of reviews, etc. (for a few examples, see [10, 11, 12, 13, 14, 15, 7, 3]), presenting multiple alternatives for publishing to the community. A few journals and conference paper review portals (Elife, OpenReview) have made large and successful strides along some of these directions. OpenReview [19], used at scale for reviewing papers submitted to premier international Computer Science conferences like NeurIPS and ICLR among others, includes an affinity matching algorithm to find appropriate reviewers, obtains 4-6 double-blind reviews per submitted paper within a constrained time-frame, and builds in an author response period with back-and-forth discussion period between authors, reviewers and "area chairs" (analogous to editors). Partially related efforts toward preprint review in the biosciences include ReviewCommons [20] and ASAPbio [21].

A recent groundbreaking announcement from the White House OSTP [22] and science funding agencies in Europe [23] adds tremendous momentum to the movement toward real open-access publication and for an overhaul of the scientific publishing system. In some countries, national bodies that officially weighed journal impact factors as a metric in evaluating scientists are moving decisively away from using such metrics in recognition of the distorted incentives such metrics create [24, 25]. These are all signs of a turning international tide.

In the Biological Sciences, the journal ELife has announced that it will select papers for review and commission reviewers, but following that step, will not make explicit accept and reject decisions and will make all the reviews public at a time selected by the authors. This is a large and important step in a direction that partially aligns with the vision of *Open*. The *Open* model further reduces gatekeeping and adds multiple layers of functionality.

Our goal is to make scientific publishing and review not only open, efficient, better at apportioning credit, and richer in context, content, and functionality, but as critically, to make it *delightful*.

The efforts bubbling up in the ecosystem of open academic publishing are an encouraging sign of the growing appetite for change and dramatic improvement in how we communicate our science. These models represent necessary and exciting progress. At the same time, current efforts tackle subsets of the problems, needs, and incentive misalignments, but to become fully mainstream requires in our view a system that tackles them all.

First, we think of the "iphone" as a metaphor in our vision of a comprehensive solution to the current publishing problems: there were other smartphones or personal assistant-phone hybrids on the market, but the alchemy of the iphone was to create a system that had some very useful tools, felt fast, effortless and intuitive to use, and was fun and beautiful to boot. We want to do this with a publishing system that tackles multiple problems and incentives at the same time, in our version of a publishing "iphone".

Second, we seek to innovate and immediately open-source the tools and platforms that we develop and make them very easy to use by anybody (other individual scientists, organizations, etc.) so anybody can easily create their own vision of what the "iphone" of publishing might look like.

Common concerns and questions

Q: The pre-publication review process helps improve my papers, how will I get the feedback I need to achieve this?

Peer review by a small number of reviewers selected based on the best but necessarily imperfect knowledge of a few editors has, in aggregate, at most a marginal impact on paper reliability and quality [**?**] relative to wholly unreviewed papers deposited on preprint servers. However, it is undeniable that a reader who has relevant expertise and is committed to carefully reading a paper and providing feedback, can improve it.

Under *Open*, we expect more rather than less useful feedback because all people invested in the results and implications have the opportunity to comment (and everybody knows that it's in their best interest to comment early and productively if they wish to have an impact on the revision). *Open* (like arXiv) allows the posting of

revisions, and the primary posted version is the latest (earlier versions can be accessed, but are not part of the primary view). Thus, this feedback can be used to revise and improve papers, if authors wish to do so.

The volume of review for a given paper will likely better match the volume of total attention the paper will receive (a matching strategy [26]), which is arguably also a much more efficient way to allocate review resources than the current system.

Q: I don't have the time to monitor and continually respond to comments on my 3-years-ago published paper. But if review is continuous and ongoing, there will also be an (unwritten) obligation to respond.

Authors may post a time-window from the date of publication within which they will plan to respond (publicly) to reasonable comments on the paper and make any corresponding edits they choose on the paper (if they wish to do so beyond the mandatory initial 1-month period). We expect this time-window will evolve to a consensus duration in a field-dependent manner. A couple of weeks to a couple of months may be a fair interval for this process, depending on field.

Q: The review and response process will devolve into long dialogues between commenters with opposing views.

Because ongoing back-and-forth discussions are not the purpose of this review forum, commenters will be limited to one review and one follow-up per paper. Each commenter will also be limited to a maximum number of papers that they can comment on each year; this will disincentivize high-volume low-quality commentary and guard against bot-like use of the system.

Q: The process will be biased: commenters and junior scientists might be strategically deferential toward famous and influential authors.

The plan is for commenters to use anonymous but unique IDs that are privately (and cryptographically) but not publicly linked to their actual identities, thus reducing the incentive for deferential review. Further, the fact that comments or reviews are themselves rated and contribute to one's reputation points, which can be used for important tenure and promotion decisions, means that reviewers are disincentivized from writing flattering or critical comments unless they are fair and contain actual high-quality content. (Responses to all the questions above assume well-intentioned participants in the review process. For trolls and bots, please see the next question.)

Q: The process will be overwhelmed by trolls and bots or others with political or commercial agendas who post self-serving comments.

First, reviewers must register to obtain a unique (anonymous) ID; the registration process will involve validation of the scientific (including institutional) credentials of the reviewer. This is similar to the Orcid system for identification. Second, there will be a threshold on reviewer quality: paper ratings from those with reputation scores below a threshold will not be included in computing weighted averages (this nonlinearity is to prevent a large number of ratings from bots from dominating the ratings provided by legitimate reviewers). Third, the limit on the number of reviews each user can post will highly restrict such activity. Fourth, the social sciences as well as translational medical fields and those with direct commercial and political implications may pose unique challenges to a democratized post-poublication review process. The proposed model may be best suited, at least at first, to the non-applied and non-medical sciences.

Q: How could my reputation points be used for tenure review if my commenter ID is not outwardly linked to my name?

Your commenter ID would be a unique ID linked privately to you through an encrypted database. *Open* would generate and provide you, through the encrypted linkage database, certificates for your contributions (including reputation points) in your actual name without outwardly linking your reviewer ID to your name. Other reviewing entities could utilize the same encrypted database if they adopt the credit allocation system.

Q: I'm already drowning in potential papers to read. I depend on editors to screen out less-good papers. I'm also worried that this system will perpetuate a herd mentality so that only the most famous authors and a few breakthrough papers are read.

It is debatable whether editors of for-profit journals use metrics best-aligned to the scientific interest in selecting papers to accept and profile. Indeed, the history of scientific publishing as a for-profit industry highlights the incentive distortions the industry as introduced into science [4]. Such issues of profit or other biases aside, it is still questionable whether a small body of gatekeepers could produce a selection of papers that better serve the broader interest than a broader, bottom-up process.

With papers, reviews, and various rich tags available in one common system, and made highly searchable as descibed above, the search for good or appropriate papers should become easier rather than harder. Explicit randomization or exploration settings should further help to find papers outside one's narrow interests and outside the "popular" stream. Highly customizable search should provide a powerful tool to counter rather than perpetuate a herd mentality. *Open* could add features that allow users to create their own shareable curated paper lists to the community ("Maya's recommended readings"; this could then effectively be Maya's own journal). Finally, those journals, journal editors, and other members of the scientific community who fight against the herd mentality dynamic will still have the opportunity to continue doing so through curation, highlighting important research, and writing or commissioning review articles or collections.

Q: Many faculty search committees will not consider a candidate without a publication in a glossy (high impact factor) journal. Will the idealistic move to Open *hurt postdocs looking for a faculty position?*

First a note on history and statistics then pragmatics. The journal impact factor (JIF) was proposed as a way to rate journals for librarians making subscription decisions; rating scientists with the JIF is increasingly recognized as the tail wagging the dog [27, 28]. Using publication venue as a proxy for individual paper quality is statistically deeply flawed, not only because of false negatives (excellent papers not getting in) but because the citation distribution is highly skewed: most articles published in high JIF venues are cited very little, with the high mean coming from a very small number of papers in the tail [29, 27, 30]. Independent of these issues are various other shortcomings of JIF metrics, which have been catalogued elsewhere [30], including the myopia of weighing papers over only two years and therefore substituting flash over durability, the encouragement to publish studies of lower statistical power, and the gaming of JIFs by journals that artificially limit the number of citations a paper can include.

Now to the pragmatics. There is no substitute for reading the work of individual candidates. However, even when a committee member invests the time to read a candidate's articles, assessing the scientific context, advance, and durability of the results requires very specific expertise that almost no committee members possess for any given search. Thus, the common reliance on high impact-factor journal names. As regular hiring committee members, we believe that the ability to read not only the paper, but also the review comments from designated well-regarded reviewers will fill in the context around a paper, providing more specific information about advance and durability than the proxy of journal name.

Thus, a richer, more rigorous, broad, and robust set of indicators of quality than we have at present and which we hope will be provided by *Open* will better serve hiring committees and candidates, than a single misleading number attached to the venue in which a candidate's work appears.

Q: Who would pay for the running and upkeep of the system?

Most funding agencies currently permit payment of publication fees for "open-access" articles in both open-access and hybrid non-open/open journals. These fees per are high, amounting to thousands of dollars per article, with generous profit margins built in for journals [4]. Similarly, scientific society publications use their journals to support other society activities, again showing that scientific publication is a profitable industry and that the current subscription and open-access fees are much higher than the costs to publish.

By contrast, the arXiv budget is approximately $2M/year for about 180,000 submissions/year, amounting to an operating cost of $12/manuscript. The discrepancy between >$3000 per open access paper and $12 per arXiv preprint leaves a lot of room for innovation at low cost.

In *Open*, overhead costs are expected to be vastly smaller than journals and only modestly greater than the arXiv: it will be operated as a non-profit, in a federated ownership model by universities and libraries. Given the sophistication and ease of use of typesetting programs like overleaf, together with standard templates for paper formatting, the costs of production can be very small. Reviewed conference papers in fields of Computer Science

typically use such templates to produce conference proceedings. We envision that the expected small per-paper costs will be easily funded by grant agencies that supported the research and university libraries who elect to support the service. In addition, direct support from public and private bodies as currently provided for the arXiv and bioRxiv, could cover costs while allowing all services and articles to be freely available to authors and readers.

Q: How do we make this switch and achieve the change we want to see?

There are four important pieces. The first is construction of scalable, robust, rich and usable systems for post-publication review. This work is partially underway, and continues to be developed as outlined above and by some of us.

The second is voluntary adoption by senior and junior members of each scientific field. We call on scientists at all levels to lead by engaging with and demanding comprehensive post-publication review systems like *Open*. On our part, we aim to make the value added by *Open* so great that scientists will gradually be drawn to it, not as much from revolutionary fervor but for its sheer utility (and delightful experience).

If, during the present transition time, post-publication review of a manuscript is not viewed as a publication, there is no cost whatsoever to first, before journal submission, putting all our manuscripts on a post-publication review site like *Open* for high-quality feedback and earlier accrual of citations.

The third is recognition and acknowledgement from academic institutions: We call for University Presidents, Provosts, and Department heads to lead by stating that i) papers submitted through post-publication review channels are regarded as bonafide and full-fledged peer-reviewed publication in the tenure and promotion processes, and that ii) they will explicitly consider and give credit for scientific contributions to constructive peer-review, according to certificates issued.

The fourth is support and mandates from libraries and funding bodies so the system is sustainable. By our calculations, *Open* will be vastly cheaper for libraries, funding bodies, authors and readers to operate than the costs paid by these same entities in the current system. Governmental institutions like NSF and NIH and European agencies like the European Research Council have recently and relatively precipitously moved toward open science in the form of widespread open access publishing, data-sharing and code-sharing [23, 31, 32]. It is important to recognize the friction of scientific publishing as another major barrier to the efficient sharing of results of taxpayer-funded science. Funding agencies can mandate that research funded by them be published through *Open*, or that funds for publishing be distributed to university libraries and *Open*. Private scientific funding bodies like the HHMI, the Gates Foundation, the Simons Foundation, Google, Facebook, CZI, the Moore Foundation, the Arnold Foundation and the Wellcome Trust, which have already taken steps toward open science and open access publishing (for instance the HHMI through Elife, which recently announced a move toward post-publication review, and various funders through OpenReview, ASAPbio and ReviewCommons) can play an important role by also mandating that research funded by them be published in *Open*, and by financially supporting the infrastructure. The Medical Research Council (UK) and the Canadian Institutes of Health Research, and multiple governments through the European Molecular Biology Conference have funded alternative non-journal review structures through ASAPbio. The NIH, though not yet funding post-publication review, has made the move to require that data obtained as part of a grant from 2023 onwards be publicly released (sharing.nih.gov); support for *Open* and other post-publication review venues is a logical partner step, if not a logical precursor for NSF and NIH. It will allow data to be more easily accessed and discovered through links from the publication.

We call for private and governmental science funding agences in the United States and elsewhere to lead by mandating that all papers be published through a post-publication review system like *Open*. And we call for these agencies and universities to fund efforts to build such an ecosystem. We believe that the time is ripe to implement this change, and feasible from both technological and sociological perspectives. As scientists we must seize this moment.

Acknowledgements

We are grateful for thoughtful and helpful conversations with many individuals who have been invested in building a new publishing landscape: Michael Eisen, Prachee Avasthi, Bodo Stern, Leslie Voshall, Erin O'Shea. We are also

grateful to many other colleagues, including Daniel S. Fisher, Roger Levy, Tony Zador, Mark Harnett, Ben W. Pearre, Gerry Rubin, and our trainees, who have provided ideas, suggestions and feedback based on their experience and perspectives. We thank Andrew Saxe for the single-star suggestion. We thank Ned Talley, Jim Gnadt, and Ken Whang for sharing their perspectives from the NIH and NSF. We are grateful to univeristy libraries for their engagement, advice, and support, particularly Roger Levy, Erin Stalberg, Chris Bourg, Laura Hanscom, and the rest of the leadership at MIT libraries and to Peter Suber of Harvard libraries.

References

[1] David Pascual-Ezama, Derek Dunfield, Beatriz Gil-Gómez de Liaño, and Drazen Prelec. Peer effects in unethical behavior: Standing or reputation? *PLOS ONE*, 10(4):e0122305, apr 2015.

[2] Nihar B. Shah, Behzad Tabibian, Krikamol Muandet, Isabelle Guyon, and Ulrike von Luxburg. Design and analysis of the nips 2016 review process, 2017.

[3] Michael B Eisen, Anna Akhmanova, Timothy E Behrens, Diane M Harper, Detlef Weigel, and Mone Zaidi. Implementing a "publish, then review" model of publishing. *eLife*, 9, dec 2020.

[4] *Is the staggeringly profitable business of scientific publishing bad for science?* Guardian, 2017.

[5] Neal S Young, John P. A Ioannidis, and Omar Al-Ubaydli. Why current publication practices may distort science. *PLoS Medicine*, 5(10):e201, oct 2008.

[6] Richard Van Noorden. Open access: The true cost of science publishing. *Nature*, 495(7442):426–429, mar 2013.

[7] *Coupling pre-prints and post-publication peer review for fast, cheap, fair, and effective science publishing.* http://www.michaeleisen.org/blog/?p=1820, 2016.

[8] Paul Ginsparg. Winners and losers in the global research village. *The Serials Librarian*, 30(3-4):83–95, may 1997.

[9] Arxiv.

[10] Dwight J. Kravitz and Chris I. Baker. Toward a new model of scientific publishing: Discussion and a proposal. *Frontiers in Computational Neuroscience*, 5, 2011.

[11] Nikolaus Kriegeskorte, Alexander Walther, and Diana Deca. An emerging consensus for open evaluation: 18 visions for the future of scientific publishing. *Frontiers in Computational Neuroscience*, 6, 2012.

[12] Brian A. Nosek and Yoav Bar-Anan. Scientific utopia: I. opening scientific communication. *Psychological Inquiry*, 23(3):217–243, jul 2012.

[13] Ulrich Pöschl. Multi-stage open peer review: Scientific evaluation integrating the strengths of traditional peer review with the virtues of transparency and self-regulation. *Frontiers in Computational Neuroscience*, 6, 2012.

[14] *A New Publishing Model in Computer Science*. http://yann.lecun.com/ex/pamphlets/publishing-models.html, 2013.

[15] Jaime A. Teixeira da Silva and Judit Dobránszki. Problems with traditional science publishing and finding a wider niche for post-publication peer review. *Accountability in Research*, 22(1):22–40, oct 2014.

[16] Bodo M. Stern and Erin K. O'Shea. A proposal for the future of scientific publishing in the life sciences. *PLOS Biology*, 17(2):e3000116, feb 2019.

[17] F.O. Isinkaye, Y.O. Folajimi, and B.A. Ojokoh. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3):261–273, 2015.

[18] Xiaomei Bai, Mengyang Wang, Ivan Lee, Zhuo Yang, Xiangjie Kong, and Feng Xia. Scientific paper recommendation: A survey, 2020.

[19] Adam Saunders Soergel, David and Andrew McCallum. Open scholarship and peer review: a time for experimentation, 2013.

[20] Reviewcommons, https://www.reviewcommons.org/.

[21] Asapbio, https://asapbio.org/.

[22] OSTP issues guidance to make federally funded research freely available without delay. https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/, August 2022. Accessed: 2022-10-12.

[23] Jean-Claude Burgelman, Corina Pascu, Katarzyna Szkuta, Rene Von Schomberg, Athanasios Karalopoulos, Konstantinos Repanas, and Michel Schouppe. Open science, open data, and open scholarship: European policies to make science fit for the twenty-first century. *Frontiers in Big Data*, 2, dec 2019.

[24] *China to end the reign of the impact factor?* International Science Editing, https://www.internationalscienceediting.com/journal-impact-factor/, 2020.

[25] Smriti Mallapaty. China bans cash rewards for publishing papers. *Nature*, 579(7797):18–18, feb 2020.

[26] R.J. Herrnstein. *The Matching Law*. Harvard University Press, 1961.

[27] May R. Berenbaum. Impact factor impacts on early-career scientist careers. *Proceedings of the National Academy of Sciences*, 116(34):16659–16662, jul 2019.

[28] Eugene Garfield. From citation indexes to informetrics: Is the tail now wagging the dog? *Libri*, 48(2), 1998.

[29] Bruce Alberts. Impact factor distortions. *Science*, 340(6134):787–787, may 2013.

[30] Frieder M. Paulus, Nicole Cruz, and Sören Krach. The impact factor fallacy. *Frontiers in Psychology*, 9, aug 2018.

[31] *About the Data Management Sharing Policies*. NIH, 2020.

[32] Max Kozlov. NIH issues a seismic mandate: share data publicly. *Nature*, 602(7898):558–559, feb 2022.