Written testimony of

Brandon Stell PhD

CNRS (French National Center for Scientific Research)

President and Co-Founder of The PubPeer Foundation

Before the House Committee on Science, Space, and Technology

For the hearing titled

PAPER MILLS AND RESEARCH MISCONDUCT: FACING THE CHALLENGES OF SCIENTIFIC PUBLISHING

July 20th, 2022

## Introduction: History and motivation of PubPeer

Chairman Foster and all distinguished members of the committee, thank you for this opportunity to join the hearing today. I am a US citizen currently living in France and a neuroscientist with the French National Center for Scientific Research (the CNRS), I am also the President and cofounder of The PubPeer Foundation, a nonprofit organization that maintains the website PubPeer.com. My testimony today is my own and does not necessarily reflect the views of the CNRS.

In the Fall of 2012, I launched PubPeer with the help of two colleagues aiming to provide a forum for scientists to discuss the scientific literature, and today I run it with the help of Boris Barbour, another CNRS researcher.

My motivation for the website is to capture discussions of the scientific literature that we scientists typically have in and around the lab, and share them publicly to help other scientists evaluate the literature. I believe that a body of collective public evaluation like this one could also help refocus the assessment of science onto results and data in individual published articles and away from the use of names of journals and secondary metrics that are rampantly used today as proxies for quality of science and scientists. The essential characteristics of PubPeer are: the ability to discuss any scientific article, speed, simplicity, a permanent right of reply for authors, the availability of anonymity and exhaustive content moderation. In the ten years since we launched, the site has grown immensely, thanks to the input

of our users, who have made the site what it is today. The site has witnessed the emergence of a community of dedicated and expert reviewers while exposing low-quality research.

A few statistics help give some idea of PubPeer's scale. We are a worldwide site in terms of the origin of the publications discussed. Our database contains over 130k comments about 45k publications in 4k journals. Although the majority are focused in biology and medicine, many other subjects are represented, such as chemistry, psychology and engineering. Currently we receive around 3500 comments and over 700k page views per month**.**

Today I will share with you some thoughts and observations from the first ten years of running this website.

## Fundamental problems in some of the published literature

When we launched the site I naively thought that comments would be similar to scientific discussions but the overwhelming majority were about more fundamental problems with data. An astonishing number of comments began flooding in that pointed to serious flaws in articles. It was as if there was a backlog of these fundamental problems and the website provided a new release valve to share them without passing through the usual mechanisms for correcting the record. Prior to the website, the process for sharing such problems had largely been to write to the authors, their institutions, and the journals in hope that one of them would recognize the problem and correct the record. However, even when this process was successful, it was extremely slow, and conflicts of interest often discouraged any action from ever being taken, because authors, journals and institutions all have reasons for not wanting to acknowledge or publicize such problems. The site allows this blockage to be circumvented: issues can be immediately displayed online so that anyone interested can be made aware and authors can respond. This process has the advantage of immediately informing those interested in building on a publication (e.g. scientists, policy makers, pharmaceutical companies, etc) about potential flaws, and shortening, by months or often years, the time it takes to find out about such issues saves researchers' time and ultimately tax dollars that would have been spent trying to build on flawed research. Over the last ten years we have received comments pointing out these fundamental flaws in articles ranging from apparently high-profile, ground-breaking research from the best institutions, to articles that were clearly only meant to be fillers for evaluation committees interested in publication metrics. It has become crystal clear to us that the reputation of a journal does not guarantee quality of the research. **It should be stressed that journal and institutional responses remain, on average, very ineffective and inefficient, even when evident problems have been made public, but the site always serves to alert researchers directly, bypassing official inaction.** Furthermore, numerous regular PubPeer users are now clearly more expert than journal and institution staff when it comes to the forensic examination of the literature.

## The underlying cause

Perhaps the underlying reason for these issues is not too surprising. With the expansion of science we continue to create new and important advances at a faster rate than ever; recent examples being the ability to rapidly create new vaccines to treat a global pandemic and the James Webb telescope sending back mind-boggling data about the outer reaches of the universe and the beginning of time. However this expansion creates challenges for identifying and supporting the best science and scientists. Job postings for faculty researchers and funding opportunities can now receive hundreds or thousands of applications and shortcuts for screening those applications become more and more tempting. It's much faster to look up metrics about the journal where an article is published than it is to read the article, and applicants (perhaps falsely) believe that those metrics are the key to the advancement of their careers. In that atmosphere it's easy to see how many of the issues we see raised on PubPeer find their way into the literature. Instead of publishing a boring result in a journal with lower metrics, an author could choose to select only experiments that worked, choose less stringent statistical tests or in the rare extreme case rearrange the data in an attempt to give a falsely positive result that would land in a journal with higher metrics. A sensational example of the problem is paper mills, which produce articles for the sole purpose of artificially inflating publication and citation metrics, sometimes by a lot, while hoping that nobody ever reads them to see that they are fake. They do get published, sometimes by reputable publishers, but perhaps that isn't too surprising, since journals collect fees for every article they publish, regardless of its quality. Although I believe these paper mill articles cause little harm to the overall progress of science—in my field they would only rarely be confused for real scientific results—they do highlight the underlying problem with scientific publishing: **incentives need to shift to place a higher importance on the content of articles**.

## How can we fix the problem?

If the underlying problem is that the content of scientific articles is undervalued, what can we do to provoke change? The metrics surrounding articles are now ingrained in the community and are unlikely to disappear anytime soon, even if they should. But commentary on sites like PubPeer can provide parallel sources of information that can be much more informative. **If scientific commentary continues to grow and involve diverse sections of the community, evaluation committees could start relying on it more than the current metrics, so that incentives might shift back towards solid, reproducible results that stand up to public scrutiny.** If contributing to this public body of evaluation were rewarded when evaluating researchers for funding, promotion and prizes, it is likely that scientists would participate to a greater extent.

Currently we rely on the journals for curation of scientific literature but there have been many recent encouraging developments from various groups that provide platforms for community curation of the scientific literature in what are known as overlay journals. These platforms enable the scientific community to take control of curation and provide a more democratic system of deciding what are the most important discoveries. The development and growth of such platforms should be encouraged.

Finally, our experience with PubPeer suggests another key intervention. The great majority of disputes on our site would be instantly resolved by access to the original data, which is often unavailable because many publications only show examples and summaries that readers must accept on trust. If data sharing were mandatory, it would strongly discourage low-quality work and misconduct, because data falsification would be harder and riskier, while detection of all problems would be facilitated. But those benefits only accrue if data sharing is required of researchers who wouldn't wish to do it; in other words it should be mandatory.

For all the above, the federal government, through its funding of science, could play a huge role, but that potential influence is largely unrealized today. To our knowledge, funding agencies like the NIH and NSF have no procedures to exploit information available through community curation sites, including PubPeer. In addition to providing additional evaluation of publications referenced in grant applications, the information from these sites could be used to reward scientists who make exceptional contributions to public evaluation.

**Legal context**

Online forums for discussing science rely on commentary that is often critical and it is not uncommon for scientists to react to this criticism with legal threats to the platform. We regularly receive legal threats[1] from disgruntled authors who prefer not to respond with their data on the website, but because the United States provides us with protection under Section 230 of the Common Decency Act, they rarely become more than threats. Without the protection provided by Section 230 the PubPeer Foundation would be forced to defend these frivolous threats and the legal fees could force us to shut down the website. The SPEECH act also provides partial protection from foreign actions.

1. https://www.aclu.org/cases/sarkar-v-doe-pubpeer-subpoena-challenge

Biography

Brandon Stell is a US citizen who currently resides in Paris, France where he leads a research team with the CNRS (Centre Nationale de la Recherche Scientifique). His team studies a part of the brain called the cerebellum and how it processes sensory information encountered during movement to adapt behaviors to a changing environment. In 2012 he co-founded the website PubPeer.com to provide scientists with a forum to discuss the published research literature. He now runs the website with another CNRS researcher, Boris Barbour, and it has grown to be one of the leading sites for scientific discussion with a dedicated community of users who have helped strengthen the scientific record by exposing and correcting its weaknesses.