Student Paper for LIBR 559L – Morrison, Summer 2011

School of Library, Archival and Information Studies (University of British Columbia)

Peer review and authority in scientific research

Keshav Mukunda

Abstract

This paper studies the historical and philosophical roots of the current system of anonymous peer review in the field of scientific research. Using the intellectual disputes between the experimentalist Robert Boyle and the philosopher Thomas Hobbes in the late 17th century to frame opposing views of the structure and role of such review in the natural sciences, the paper argues that present attempts to refashion the prevailing environment of peer review – specifically through the use of open review models – can result in greater transparency and credibility in research, which in turn lends greater authority to the results of such work.

The historical emergence of peer review

The current arrangement of peer review, specifically of articles submitted to a scholarly journal, is so widespread that it is easy to forget just how recently this particular form of evaluation took shape. The relatively anonymous review of academic work by a small group of specialists in the field – typically just one or two experts – only truly emerged in its modern form during the middle and latter half of the 20th century. However, the ideas behind peer review, and the belief that work of a scholarly nature should be evaluated by one's contemporaries, whether experts or authorities in a position to judge, is no doubt much older (Kronick, 1990).

The initial human responses to the question of how we come to accept knowledge and understand the ways in which our world works were typically based on the authority given to certain individuals or bodies as a consequence of their position in a hierarchical social structure; religious leaders, monarchs, and so on. Such responses did not require much accord, either between individuals and those with authority in society, or between people and their unmediated observations of the world. But all of this began to change dramatically in Europe when, after the invention of the printing press in the 15th century, traditional structures that had dominated the field of knowledge began to be shaken by new opinions distributed widely and without constraint via the medium of print. Slowly the idea emerged that knowledge, of the sort that could be designated as 'matters of fact' (a term used explicitly by the 17th century natural scientist Robert Boyle in describing the outcomes of his experiments) could be advanced through social consensus or, in the case of the physical sciences, through the consensus of a group of

qualified specialists. Consensus of this kind was reached, if at all, only after the most thorough scrutiny of new pieces of knowledge and new observations of the world. However, in the initial stages of the spread of the new paradigm, this form of scrutiny or review was not particularly motivated by a desire to advance our collective understanding of ourselves and of our universe. For example, when Galileo ran into trouble with the Catholic Church in the early 17th century over heliocentrism, this preliminary form of review by the papacy was clearly an attempt to stop the spread of new ideas about the structure of the universe (Spier, 2002). But over time the review of work of a scholarly nature began to be undertaken by an informal community of academics and thinkers, whose principal aim – at least in spirit – was to generate reliable knowledge and not simply restrict the growth of scholarship.

One of the earliest such informal communities in England was the *Invisible College*, a group that included many notable scientists, mathematicians and natural philosophers. The members of the college were inspired by earlier ideas of Francis Bacon on how the newly evolving scientific method should create knowledge, and how such knowledge could be validated. To this end, they met regularly to observe experiments together and thereby, in their view, generate matters of fact about the world through the collective experience of natural phenomena. More generally, the college and other such networks were a means by which scholars could communicate their ideas to each other, and they served as precursors – and in later years, as alternatives – to academic journals. The Invisible College led to the formation of the *Royal Society of London for Improving Natural Knowledge* in 1663, and the members of this society resolved to meet periodically in order to communicate their ideas to each other, as well as to perform scientific experiments. It was this resolution that led to the publication, two years later, of what is generally regarded as the first scholarly journal, the *Philosophical Transactions* of the Royal Society of London. While another scientific periodical, the *Journal des Sçavans*, was also first published in Paris at almost the same time, today we would not recognize it as being a scholarly journal; much of the content of the Parisian publication consisted of "mail gossip", including theories and conjectures, observations, discoveries, as well as legal matters and obituaries of famous people (Guédon & Siemens, 2002).

Initially, what was published in the *Philosophical Transactions* was decided by the editor of the journal, or anyone else that the editor deemed suitable to comment on the work; there was no requirement that they had to seek assistance in this process. Although this was review of a certain sort, it did not really resemble the kind of peer review we see today. The first peer review in the form that presently exists today could perhaps be traced back to the year 1752, when, after almost a hundred years since its inception, the *Philosophical Transactions* established a committee that would review every article that was submitted for publication (Kronick, 1990). This committee could also solicit comments by other members of the Royal Society who were "knowing and well skilled in that particular branch of Science that shall happen to be the subject matter of any paper" (Royal Society of London, 1940).

For roughly two centuries after this, the review of submissions to a journal was primarily handled by editors alone. Occasionally, an editor would consult an individual or group deemed suitable to comment on the work, but one of the main difficulties was the creation of duplicate copies of a submitted

manuscript (Spier, 2002). This situation was greatly improved in the late 1890s when commercially produced typewriters began to be used to create copies of prospective articles to be distributed among small committees for peer review. Still, even at this point peer review was not so widespread, and in fact it was more common for articles to be judged solely by editors. A classic case, frequently cited to argue that the current system of peer review is not a necessary condition for scholarly quality, involves the four landmark papers published by Einstein in 1905 that changed the shape and direction of modern physics. None of these papers were reviewed by other physicists, and they were published purely at the discretion of the editors of the respective journals. In fact, it has been documented that of the 300 or more papers that Einstein published, only one actually went through anonymous peer review; in 1936, he submitted a paper to the American journal *Physical Review*, and was returned some critical comments on the validity of certain calculations that he had made. As it turned out, the reviewer was correct in their comments, but Einstein was completely taken aback at the process of anonymous peer review and simply reacted angrily to the report and withdrew his paper, preferring to publish the paper elsewhere (Kennefick, 2005). The kind of peer review that we are familiar with today was far more common in North America than in Europe at the time, but, as this example shows, it was still not as widespread. Thus, in the sciences for example, the prestigious journal *Nature* did not use a formal anonymous peer review system till 1967, and similarly most other journals only began to do so over the course of the twentieth century.

The basis for authority in scientific research

In his book *Novum Organum*, or the 'New Instrument', published in 1620, Francis Bacon outlined a system of logic by which reliable certainties about the world could be produced via a process of inductive reasoning, experimentation, and observation. This process is familiar today as the 'scientific method', but at the time it stood in contrast to the method of deductive reasoning which was the dominant logical method of arriving at knowledge. The Invisible College, which seems to have loosely formed in the 1640s, promptly dedicated itself to this Baconian program of the 'new science'. A principal member of the Invisible College, and later of the Royal Society of London, was Robert Boyle, perhaps one of the most brilliant experimentalists of all time. Based on their reading of Bacon's ideas, Boyle and his contemporaries set about creating a structural framework by which "matters of fact" could be created. In brief, this framework rested on the notion that the foundational experiments of the new science must be collectively observed by a multiplicity of witnesses. A single person might not provide reasonable proof, but as greater numbers of subjective experiences were brought to witness the same event – and *agree* on what had been witnessed – then the reliability of inferences could be greatly improved (Shapin & Schaffer, 1985). The assurance of the 'truth' of the deductions that could be made from common observations would naturally emerge, as the variation of individual testimonies counteracted and balanced each other out. So authority was granted to research by a process of collective witnessing and review; while in this case the reviewers were spatially co-located with the experiments and the researchers,

Boyle envisioned a systematic methodology of writing scholarly articles by which one could carefully describe experiments and observations, thereby allowing review by people in different places and at different times (Shapin & Schaffer, 1985). A key component was that the observations were made in a public space that was to be accessible by all; in Boyle's view, this would guarantee that no divisive arguments could take place about the nature of what precisely had been observed by all, although debate was encouraged within the clear parameters that had already been laid out by the act of mutual witnessing. Thus peer review could be seen to be a form of collective agreement on what constituted "matters of fact", and what did not, within the guidelines imposed by the common observation.

At the time, a number of other philosophers and scholars in England as well as in the rest of Europe raised numerous objections to Boyle's program, and one of the most interesting arguments came from the philosopher and political thinker Thomas Hobbes. While his arguments were far more nuanced and elaborate than we have the space for here, three of the chief criticisms were as follows: first, he claimed that while Boyle and his contemporaries insisted on the *public* nature of their activity, they were in fact carefully managing closed spaces, admissible only to those who were permitted by a "master who decided who could come in and who could not" (Shapin & Schaffer, 1985, p. 113). Thus the findings of the experimentalists were simply not witnessed by all. Secondly, although debate seemed to be encouraged, to Hobbes there were already those "who are most believed" (Hobbes, 1661), in the sense that there were intrinsic systems of authority already in place which tended to favour the statements and ideas of certain persons over others. The third point was more subtle; even if it were possible for *all* people to observe phenomena in common, Hobbes asked how such collective witnessing would be any different from a large number of individual isolated experiences. In effect, on the one hand he questioned the assumption that such experiments were indeed witnessed simultaneously and together, and on the other, the assumption that the only credible experience is that attained by collections of observers together; as Hobbes put it, "are not those phenomena which can be seen daily by each of you suspect, unless all of you see them simultaneously?"

These criticisms have a clear bearing on both the process of peer review as well as the claim that such review acts as a method by which new ideas are substantiated. Hobbes' second point above is worth noting in this context, as peer review is routinely influenced by such forms of authority, in which some people are believed and others are not. After all, the argument that verification comes through of a group of experts cannot be simply accepted any more, since the space of such review is sharply circumscribed and does not at present extend to all stakeholders; not just all scholars across the world, but also all those who will be affected by the structuring of knowledge in possibly discriminatory ways. Academic peer review is invariably cited as a process by which new knowledge is validated through careful scrutiny and verification, and a conventional argument posits that such review serves as a way to ensure quality control for scholarly work. While this might broadly be the case, we yet see clear evidence of some of Hobbes' criticisms here: in it's current form, peer review is largely a closed activity, carefully managed, with often implicit forms of hierarchies at play. Even at its best, when peer review is used with an honest

attempt at fairness, there can still be a great degree of arbitrariness and bias in it's results (Fitzpatrick, 2009, Guédon & Siemens, 2002).

There have been some significant cases where the peer review process has led to unscholarly – and occasionally fraudulent – work being published in journals, and, conversely, there are many examples of articles being rejected after negative peer review, only to have the authors later win Nobel prizes for their work (Nielsen, 2009). However, these generally seem to be exceptions and as such, cannot be controlled by any system of evaluation, however robust. The point to be made here is that there are, on the one hand, numerous cases of the discretionary and biased excessive scrutiny and frequent rejection of a submitted article due, for example, to the work being created by researchers in developing countries (Willinsky, 2006, p.106); and on the other hand, cases where works have been assumed to have greater reliability due to the reputations of the authors involved, leading to these authors being unreasonably given automatic credence. Thus, the kinds of consequences associated with the present system of anonymous peer review that are most crucially in need of attention are those pertaining to reducing the influence of the 'master' who decides what forms of knowledge are deemed legitimate, and those that aim to reduce the kinds of hierarchies that lead some to be 'more believed' than others.

The case for openness in evaluating knowledge

While it is clear that scholarly research – specifically in the physical sciences – has been immensely successful over the last few centuries, and a great many technologies have been derived from this research, it is also evident that most of society does not interact with the processes by which such investigations are undertaken. The routine, daily work of scientific research has been far removed from the public both spatially and intellectually. Of course, there is some justifiable reason for this; the scientific method requires that laboratories be controlled spaces, and in addition, the degree of intellectual work required to carry out advanced research requires years of specialization. Yet this purposeful removal of the scientific from communal space has, among many other things, led to strong feelings of wariness and cynicism among the general population. Many scholarly assertions have a certain quality of authority, and although some are displaced from their original contexts, they often come in the somewhat patronizing tone of wanting to help people dismiss their own creative knowledge gained from experience and common sense. Although the scientific process could initially be viewed as a careful refinement of common sense, today this is no longer the case, as the structure and composition of science has grown far beyond the reach of individual experience. As a classic example of this we could consider the scientific reasoning that supports the view of anthropogenic climate change: an individual simply could not definitively conclude whether such a theory is true or not just by careful observation from their window. Indeed, many people tend to doubt that climate change is happening based on the relatively benevolent view from their window, and yet it is certainly occurring, inasmuch as any statement can be made with certainty.

When seen from the point of view of the academic researcher, this kind of dichotomy is seen as a gap or deficit in the civic understanding of science. Ironically, this idea of a gap that needs to be reduced reveals the general reluctance on the part of scholars to allow the public to be partners in the evolution of scientific research, and itself contributes to the magnification of the gap. However, from the perspective of the public, the polarization between these two modes of thought could be seen as a purposeful withholding of certain ways of thinking that are pertinent to society as a whole. Frequently the scholarly community attempts to reconcile this separation by presenting knowledge, in it's final polished form, to the wider public. But simply stating matters of fact is not sufficient; how such conclusions are arrived at is just as important, if not more so (Wynne, 1992). The example of anthropogenic climate change illustrates why it is vitally important for more open interaction between the public and the process of scientific research, not just between the passive public and technoscientific terminology. What is required is a way to tell people "what science is like in the making" (Shapin, 1992).

It is here that the open access movement has the potential to play a significant role. Specifically, the process of *open peer review* that we describe in the next section could allow nonspecialists to see exactly how it is that scholarly knowledge is created, and why research that is tested through a system of careful scrutiny and assiduous review leads to more reliable matters of fact. After all, given the purportedly crucial role that peer review plays in creating and establishing an authoritative framework for new scholarship, it is relevant to ask why the entire process of review is currently kept more or less hidden from external scrutiny. Reviewers' comments are typically kept anonymous, are not published along with the final paper, and are commonly restricted to only one or two peers. Also, once a paper has passed this refereeing process, the tendency is to accept its results, a priori, as being validated and therefore accurate. In this sense, the current system of closed peer review has begun to look not unlike the earliest judgements of knowledge by small groups of authorities placed high up in the social hierarchy. But it is not necessary for new ideas and theories to be protected and hidden away in order to flourish. In fact, their utility and authority comes from their ability to persevere, especially under examination from numerous points of view.

Open peer review

Among the future directions that peer review might take, as attempts are made to mitigate some of the shortcomings of the current system, one important alternative is *open peer review* (Fitzpatrick, 2009, Pöschl, 2010). In addition to making the process of review more transparent, and possibly more conscientious, open peer review could also play a crucial role in allowing nonspecialists to see exactly how it is that scholarly knowledge is created, and why such an approach leads to more reliable knowledge about our world.

There are numerous forms that open peer review takes, and only the general outline of such processes will be described here, based on the method used by the open access journal *Atmospheric Chemistry and*

Physics, as well as by a number of other publications by the European Geosciences Union (Pöschl, 2010). It is important to stress that open access together with 'community peer review' does not run counter to the obvious strengths of the typical peer review system, but instead augments and expands it to include honest discussion and review by all interested people, whether scholars or members of the public. While preserving the usual voluntary anonymity for designated referees, this system also documents every discussion point, interactive comment, innovation, flaw, and correction in the submitted paper, and these are publicly accessible as well. As a remarkable consequence of this access to each step in the process, the public can directly see just what the peer review process looks like, and it allows one way to observe how scientific knowledge is made and verified. In addition, articles are easier to review and evaluate by specialists in the field if more of them are allowed to read the submitted paper and be directly consulted. Designated reviewers are typically given a certain time frame in which to respond to papers. They also have the opportunity to let their diligent work in the careful reading of manuscripts be publicly acknowledged, and if they wish to let their name be used, then they can be included as a part of the creation of the article itself. This process has provided a better way to manage "quality assurance" of the research itself, since the work is scrutinized by more people and is not "locked away behind . . . barriers" (Pöschl, 2010, p.1). Such interactive open peer review approaches have compared favourably with other conventional methods on numerous quantitative indicators, and using open peer review has been shown to have no statistically significant effect on the quality of the reviews themselves (van Rooyen, Delamothe, & Evans, 2010). Similar open frameworks could be put in place to improve current online journals, using whatever variations are required for each area of specialization or type of publication. Besides making sense from a structural standpoint, interactive open peer review is also in line with the fundamental philosophy of science and rationalism, which values collaboration, free speech, the rigorous inspection of results, and efficiency in the communication of ideas. This model further presents a means by which the scientific and public worlds could be reconciled through collaboration.

Conclusion

It is important to remember that the current system of anonymous closed peer review is "a little like democracy . . . [in that it is] the least objectionable form of evaluation" that exists currently (Guédon & Siemens, 2002). However, just like democracy, the existing peer review process can also be improved; we can build on the long-standing tradition of scholarly review and attempt to reduce some of it's inconsistencies and errors by using more open methods.

Open peer review and other similar community review processes have two important implications. First, they lead to more transparency and greater evenhandedness in the treatment of new research, especially in the context of the relations between scholars in developed and developing nations. Second, they have the potential to reduce the widening gap between the professional researcher and the public, by engendering a better civic understanding of how scholarly research is assessed and verified, and showing how this actually does lead to more dependable opinions about the ways in which our physical

world works. In this respect, some of the criticisms raised by Hobbes – regarding Boyle's interpretation of the scientific method of generating knowledge – can be partially addressed by using open peer review.

More generally, the tendency toward being closed, by restricting access to information and modes of thought, is damaging in the long run and puts a greater weight on our economic and human resources, than a general commitment toward being open. The benefits of reverting to an open system, even if there are initially some doubts and mistakes made, have greater and more far-reaching implications for the improvement of the human good.

References

Fitzpatrick, K. (2009). *Planned Obsolescence*: Chapter One: Peer Review. Retrieved from http://mediacommons.futureofthebook.org/mcpress/plannedobsolescence/one/

Guédon, J.-C., & Siemens, R. (2002). The credibility of electronic publishing: peer review and imprint. *TEXT Technology*, *11*(1), 17–35. Retrieved from http://web.viu.ca/hssfc/Final/PeerReview.htm

Hobbes, T. (1661). *Dialogus physicus De natura aeris, translated by Simon Schaffer*. London: Londini.

Kennefick, D. (2005). Einstein versus the physical review. *Physics Today*, *58*(9), 43. doi:10.1063/1.2117822

Kronick, D. A. (1990). Peer review in 18th-century scientific journalism. *Journal of the American Medical Association*, *263*(10), 1321–1322. doi:10.1001/1990.03440100021002

Nielsen, M. (2009). Three myths about scientific peer review. Retrieved from http://michaelnielsen.org/blog/three-myths-about-scientific-peer-review/

Pöschl, U. (2010, March). Interactive open access publishing and public peer review. *IFLA Journal*, *36*(1), 40 –46. doi:10.1177/0340035209359573

Royal Society of London. (1940). *The Records of the Royal Society of London*. London: Morrison & Gibb.

Shapin, S. (1992). Why the public ought to understand science-in-the-making. *Public Understanding of Science*, *1*(1), 27 –30. doi:10.1088/0963-6625/1/1/006

Shapin, S., & Schaffer, S. (1985). *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life: including a translation of Thomas Hobbes, Dialogus physicus de natura aeris by Simon Schaffer*. Princeton, N.J.: Princeton University Press.

Spier, R. (2002, August). The history of the peer-review process. *Trends in Biotechnology*, *20*(8), 357–358. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167779902019856

van Rooyen, S., Delamothe, T., & Evans, S. J. W. (2010). Effect on peer review of telling reviewers that their signed reviews might be posted on the web: randomised controlled trial. *341*. doi:10.1136/bmj.c5729

Willinsky, J. (2006). *The Access Principle : the case for open access to research and scholarship*. New York: MIT Press. Retrieved from http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=10611

Wynne, B. (1992). Public understanding of science research: new horizons or hall of mirrors? *Public Understanding of Science*, *1*(1), 37 –43. doi:10.1088/0963-6625/1/1/008