Peer Review and Quality Control

Silvio O. Funtowicz* & Jerome R. Ravetz**

*Silvio O. Funtowicz
silvio.funtowicz@svt.uib.no
Professor
Centre for the Study of the Sciences and the Humanities (SVT)
University of Bergen
Allegt. 34
Phone: +47 55 58 27 05
Postboks 7805, NO-5020 BERGEN
5020 Bergen
NORWAY

**Jerome R. Ravetz
jerome.ravetz@gmail.com
Associate Fellow
Abstract

Peer review is the recognised standard for quality control in science. There is a growing awareness about the need to introduce modifications to the traditional peer review process in order to meet the challenges of a continuously expanding scientific enterprise, the emergence of new forms of science and the impact of new technologies. The growing application of science and science-based technologies to society requires innovative quality control mechanisms and the inclusion of new stakeholders. Collegial peer review is being rapidly transformed to review by an 'extended peer community, raising important issues to the governance of science.

Keywords:

peer review, quality, post normal science, extended peer community, trust, new forms of science, blogosphere, fraud, misconduct, retractions, reproducibility

Cross-references

Citizen Science; Ethical practices, institutional oversight and enforcement; Ethics and Values; Intellectual Property, Concepts of: Intellectual Property: Legal Aspects; New Forms of Science;
Norms in Science; Public Engagement with Science; Science and Industry; Science and Law;
Science and Technology Studies, Experts and Expertise; Science Communication;
Scientometrics; Universities, Science and Technology: United States; Trust, Political; Trust,
Sociology of;
Background

Peer review is a mechanism for quality control in science, including the assessment of proposed projects and of completed work. Its roots go back to the emerging science of seventeenth century, when novelty in natural knowledge became distinguished from technical invention, and the ideals of reproducibility and publicity became established. Peer review was acknowledged by the mid-twentieth century as the unquestioned norm in academic science.

Recent experience has shown the need for important modifications in the process of quality assurance, such as changes to the traditional, largely informal, procedures of collegial peer review, in order to take into account the emergence of new forms of science, the increased competitiveness of the research enterprise, the impact of new technologies and the inclusion of new stakeholders. Collegial peer review is being rapidly transformed to review by an "extended peer community".

Among contemporary forms of scientific practice, curiosity-driven research with no regard for applications has been increasingly marginalized. A diversification has occurred so that quality assurance must also be considered in such areas as mission-oriented and issue-driven research, techno-science and innovation research, and the provision of scientific advice for policy.

As the unquestioned norm in academic science, it was assumed that quality was unproblematic, owing to some special ethical disposition of scientists and their methods. With subsequent transformations in the social practice, embeddedness, and pervasiveness of science in society - which now extends into the economy, technology, policy and life forms and styles - corresponding changes in quality assurance, are urgently needed. The crucial requirement is for a more aware understanding of the encompassing community of those producing, evaluating,
sharing, profiting from, and even “suffering” the products and processes of science (see, e.g., Stampa, 1997).

In most other spheres of institutional activity, the formalization of quality assurance has become the norm, as for example through the wide-ranging standard-setting procedures of the International Standards Organization (ISO). In academic science, however, along with cultural pursuits like the arts, the methods have been and are, still, largely informal. Science has been almost unique in having self-assessment performed by accredited practitioners rather than by users or external critics. To what extent and in what ways this must change to keep pace with science’s expanding role in public and private life has become an urgent question in the conduct of science, in the governance of research and in the co-evolution of science and society.

The assurance of quality is not a straightforward task. This has been known since the time of the Romans, as indicated by the Latin motto *quis custodiet ipsos custodes?* (Who will guard the guardians themselves?). This motto implies an indefinite iteration. It is a reminder that, however, routine may be the tasks of quality control, full quality assurance demands yet higher levels of supervision at which informality, extended participation and explicit value judgments are necessary (Funtowicz & Ravetz, 1990).

These insights were introduced in industrial quality control by the pioneering work by William Edwards Deming at the end of the 1940s. His concepts of profound knowledge and total quality assurance were successfully applied by Japanese car manufacturers before being imported back, much later, to the US. Edwards Deming’s ideas were quite easily absorbed in the Japanese cultural landscape. These ideas included the establishment of quality circles, in which the assembly line was transformed into a kind of participatory process. Two aspects of this process
were quite revolutionary: the first was the extension of the community of quality controllers and evaluators to the work force, relying on practical knowledge, experience and commitment; the second was the encouragement of whistle blowing, whereby any member of the community, including the work force, was entitled to stop the process if they believed that quality was violated. On Edwards Deming's philosophy of quality, see *Out of the Crisis* (Deming, 1986).

As long as science remained mainly academic, problems of quality were assumed to be resolved by self-correction, anticipating or responding to peer criticism; a process conceived as part of the very nature of the scientific endeavour. The informal systems of checking by peers seemed a rational and efficient response to the problem, rather than a culturally contingent mechanism characteristic of a particular epoch. Scientific facts were believed to be discovered by a reliable method, and scientists themselves were viewed as being endowed with certain superior moral qualities that protected them and their work from ordinary human failure or error. This self-correcting property of science could be explained in sociological terms, as in the four norms of scientific institutional practice expounded by Robert K. Merton in 1942 (Merton, 1973 [1942]): Communalism, Universalism, Disinterestedness and Organized Scepticism (CUDOS) defined the ethos of Modern science. Philosophically, it could be justified in the committed, determined attempts at self-refutation supposed by Karl Popper to be normal scientific practice (Popper, 1959), where the refutation potential becomes the hallmark of a scientific hypothesis.

**Science as the endless frontier**

With the onset of the industrialization of science after World War II, the self-conscious study of science as a social activity, including the methods of quality assurance, became inevitable. With the adoption of Vannevar Bush's report, *Science the Endless Frontier* (Bush, 1945), the stakes of
the scientific enterprise became higher, inviting scrutiny of the quality of processes and products. Growth in size, capital investment, scale, and social differentiation within science created divisions between managers and research workers, as well as between researchers and teachers in universities. A Gemeinschaft (community) of scholars could no longer realistically be assumed.

The earliest disciplined analyses of the quality of production in science were quantitative. Derek J. de Solla Price, who devised some measures of quality and provided analyses of its distribution, did the pioneering work. He noticed that at the leading British science reference library only a minority of journals was ever requested. The contents of the others could be inferred to have no interest, and hence to be of very low scientific quality (Price, 1963). This phenomenon is a reminder that “quality” is a relational attribute, and its definition “fitness for purpose or function” depends on whose purposes are dominant; not always perhaps those of a community devoted to the advancement of learning and achievement of reliable knowledge, but possibly only of those scientists working under constraints of “publish or perish.”

Price’s studies were continued in two directions. At the Institute for Scientific Information, Eugene Garfield produced more searching and sophisticated measures of quality, using citations rather than mere number of publications. Such attempts at quantification were bound to become controversial (Brooks, 1982; Garfield, 1970, 1972). It was impossible to avoid bias in the selection of the relatively small set of journals used for citations; those in conventional mainstream English-language research science were inevitably privileged at the expense of all others. Further, when quantitative measures of citations came to be used as indicators of academic merit, manipulative practices, including reciprocal citations, inevitably developed. The deep problems of a quantitative proxy for quality suddenly became acute.
In a more reflective vein, Jerome R. Ravetz applied the *quis custodiet* principle to analyse the vulnerability of the quality assurance system in science. He observed that the processes of peer review are conducted largely informally and (unlike research) are not themselves normally submitted to open scrutiny and review. They require different competences, which are not part of the formal training of scientists; and they also are more open to a variety of abuses, ranging from bias to plagiarism. One can understand in these terms the phenomena of low quality, both in scientific research and in technological development. Thus, while denying that the practice of science automatically produces a higher morality, Ravetz agrees that moral standards are necessary for the successful practice of science. On this basis, he stresses the importance of morale and morality (and ultimately idealism and leadership) in science (Ravetz, 1996 [1971]).

This analysis offers a background for the increasing interest in trust as an essential element of practice in science, in society, and in their interactions. The broader society has provided resources to the esoteric activities of science because it trusts the scientific community to make good use of them. There has always been an undercurrent of distrust, based on evidence either of pointless research or of malign applications. Once science became deeply involved in technology and related policy problems that crucially affect public health and welfare, the traditional relations of trust could no longer be assumed. It appeared to be necessary for the principles and practices of accountability to be extended from the institutions of political and economic governance (as, e.g., representative democracy) to those institutions that govern science and its applications.

Quality control in research science has become more difficult as the relatively inflexible technical requirements of the traditional printing process have first been relaxed, and then made obsolete by novel information and communication technologies and economic restructuring.
There no longer is a well-defined gateway to publication through the institutions that control reproduction of, and hence access to, documents. First through inexpensive photocopying and now through the Internet, it has become easy for anyone to distribute knowledge wares to an unrestricted audience. It is like a re-enactment of the period that followed the invention of the printing press and the publication of Gutenberg’s Bible around 1450. Now, as then, the monopoly of knowledge dissemination is collapsing and new and expanded interested audiences are being created.

In addition, the presence of the global media tends to bypass the traditional processes of evaluation that were conducted personally among colleagues. The demystification of scientific practice enables such events to become news, and, at the same time, is fostered by their being exposed. Top scientists become like media celebrities - needing the media for advertising themselves yet simultaneously hating them for their unwanted intrusions.

Isolated scientific results can and have become media events (Close, 1991). Seralini and colleagues rat study, relating cancer to genetically modified maize published in a peer-reviewed journal (Seralini et al., 2012), created not only great controversy among scientists (Butler, 2012) and stakeholders (see, for example, http://www.euractiv.com/innovation-enterprise/commission-science-supremo-endor-news-514072), but a new form of science-related reality TV show as Russian scientists planned to reproduce and broadcast the GMO experiment. The scientist in charge of the project told the news agency: “This is a unique experiment… There hasn’t been anything like this before – open, public research by opponents and supporters of GMO” (RT.com, 2012).
All those with an interest in the outcome, such as researchers, consumers, politicians, regulators, and the stock market, become potential stakeholders in the evaluation and promotion of the results. Thus, science arguably becomes accountable to a drastically extended peer community in the quality assurance process. The criteria of quality applied by these heterogeneous actors need not be identical to those of public knowledge generated within the tightly knit scientific networks of academic science.

**Pathologies of science in society**

These developments may be judged in different ways. While they may seriously disrupt the procedures of quality assurance in normal science, they can also bring needed public scrutiny to bear on innovation, controversies and, even, scandals. In 1999, the Netherlands Environmental Assessment Agency (PBL) was accused by an employee, senior statistician Hans de Kwaadsteniet, of lying and deceiving; his claim was that figures released by the Agency, based on virtual reality computer models were essentially unreliable. The whole episode, which lasted several months, ended with the transformation from a technocratic model of quality assurance to a more accountable, transparent and participatory paradigm (Petersen et al., 2011).

The “Baltimore affair,” centring on the US Nobel laureate David Baltimore’s laboratory at MIT, was perhaps in its time the most notorious case in which a dispute about scientific misconduct was blown up into a lengthy, visible, political saga that unquestionably damaged all the individuals and institutions involved (Kevles, 1998). The episode was symptomatic of an increasingly recognized problem of “deviance” in science, which carries the unspoken danger that, without timely correctives, misconduct might become the norm.
There has been a growing awareness and concern, expressed in studies and articles not only in the scientific press (Jha, 2012; Zimmer, 2012), about the quantity and quality of fraud and misconduct cases (Fanelli, 2009; Corbin, 2012). Several initiatives have been proposed to tackle misconduct and related issues, such as retractions (see, for example http://retractionwatch.wordpress.com ) and the reproducibility (see, for example, http://www.reproducibilityinitiative.org) of research results (Nature Biotechnology, 2012). The problem does not affect a single discipline but it is felt across a wide range of research areas; thus, for example, the Nobel Laureate Daniel Kahneman, has challenged psychologists to “clean up their act”, according to Nature (Yong, 2012).

All these developments affect the maintenance of trust, which is necessary for ordinary scientific practice and even more so for quality assurance. As in other professional domains, the normal tendency in science has been for misconduct to be covered up by the responsible institution (not necessarily by the community of scientists). In such situations, ultimate exposure does even more damage and further erodes the basis for mutual trust. Grigory Perelman, perhaps the most gifted mathematician of his generation, is quoted to have said, after refusing the European Mathematical Society top prize in 2007, that the award committee was unqualified to assess his work (Paulos, 2010). Perelman's mistrust towards his colleagues and the organized mathematical community is also illustrated by his decision to publish his proofs of the geometrisation conjecture in the non-peer reviewed archive for electronic preprints, arXiv. Nasar and Gruber report that Perelman had not mentioned the proof or shown it to anyone. “I
didn’t have any friends with whom I could discuss this,” he said in St. Petersburg. “I didn’t want to discuss my work with someone I didn’t trust” (Nasar & Gruber, 2006).

Attempts to circumvent the need for trust by increasing bureaucratic surveillance are likely to be counterproductive in their own way, by erecting impediments to free inquiry and communication among colleagues.

The relations between social science and natural science have also been transformed during the last decades, with implications for quality control. Starting with the acceptance of natural science as the ideal of knowledge, essentially independent of social forces, there has been a gradual but accelerating shift toward recognizing all sciences as incorporating social constraints and biases. An early critical interaction was in connection with the astronomical community’s management of the eccentric Velikovsky (de Grazia, 1966). Later, the social science community embraced Thomas Kuhn’s disenchanted picture of “normal science” (Kuhn, 1970). Finally, post-Feyerabend studies of science re-examined the whole institution of scientific inquiry without presupposing any privileged status in relation to either virtue or natural knowledge (Bloor, 1991; Barnes et al., 1996; Collins and Pinch, 1993; Fuller, 1993).

When natural scientists, led by physicists, eventually confronted the emerging socialized picture of their discipline, the reaction was so strident that “science wars” became an appropriate label (Gross et al., 1997; Nelkin, 1996; Ross, 1996). Sociologists of science and post modernists were indiscriminately blamed for all the ills of science, including decline of public trust, budget cuts, resurgent Creationism, and even poor teaching of science. A physicist whose hoax article (Sokal, 1996) was accepted by a leading cultural studies journal, Social Text, helped to crystallize the
attack (Bricmont & Sokal, 1998). The implication was that the critics of science had no real quality control of their productions. The science warriors’ assumption was that within real science, such problems are prevented from occurring because of the verifiable empirical content of scientific research. In the ensuing debate, there was little mention of the ease of publication of erroneous or vacuous research in the standard scientific literature. Historical episodes, like Millikan’s manipulation of his oil-drop results in the course of a controversy on the charge of the electron, were discounted as mere embarrassments (Segerstale, 1995).

It has been presupposed thus far that “science” refers primarily to traditional basic research. However, among contemporary forms of scientific practice, curiosity-driven research with no regard for applications is giving way to new forms of science. Quality assurance must also be considered and performed in such areas as mission-oriented, issue-driven and innovation research, forensic science (Foster and Huber, 1997; Jasanoff, 1995), and the provision of scientific advice for policy (Jasanoff, 1990; Salter, 1988). In addition, the products themselves and the media through which they are diffused increasingly are diversified. For example, patents are now a common outcome of a research process, and this form of intellectual property is radically different from traditional published papers (Myers, 1995). Also, results are reported in non-peer reviewed electronic websites, unpublished consultancy reports and grey literature, or kept confidential within institutions or even totally sealed under lawyer-client confidentiality and legal settlement agreements.

Peer review extended and democratized
With traditional peer review as the norm, the challenges of quality assurance for these new products and processes are nearly unrecognizable. A genre of critical literature has developed, with some authors directing anger at the new contexts of scientific production (Huber, 1991), and others more clearly appreciating the problems they present (Crossen, 1994, Jasanoff, 1990, 1995).

A parallel diversification has occurred in the types of knowledge production that are accepted as legitimate. The democratization of knowledge now extends beyond the juries who assess the quality of technical evidence in courts (Jasanoff 1998) to include those who master previously esoteric aspects of their predicament (e.g., illness, contamination, pollution, oppression, discrimination, exploitation) through special-interest groups or Internet-based social media. In addition, claims of specialized or local knowledge are present in even more diverse contexts, as among indigenous peoples, and in systems of complementary or “traditional” medicine. These claims are commanding increasing commercial and political support among various publics, as well as gaining explicit recognition in numerous international treaty regimes. As a result, a new philosophy of knowledge appears to be emerging, based on a new disciplined awareness of complexity, in which a plurality of legitimate perspectives is taken for granted (Funtowicz & Ravetz, 1992). Modern science, with its characteristic methodology and social location, is part of this enriched whole, but not coextensive with it. The criteria and tasks of quality assurance must explicitly involve additional values and interests, incorporating even the ontological commitments of groups other than scientists. This new configuration has been termed post normal science.

There are now many initiatives, increasing in number and significance all the time, for involving wider circles of people in science, decision-making and implementation on policy issues. The
contribution of all the stakeholders in cases of post normal science is not merely a matter of broader democratic participation. Novel and emerging techno-scientific issues, promoted and produced as innovation research, are in many ways different from those of traditional research science, professional practice, or industrial development. Each of those has its established means for quality assurance of the products of the work, whether through peer review, professional associations, or the market. However, for these new forms of knowledge production, the maintenance of quality depends on open dialogue between all those affected. In post normal science, it is called an “extended peer community,” consisting not merely of persons with some form or other of institutional accreditation, but rather of all those with a desire to participate in the resolution of the issue. Since this context of science involves politics and policy, we might see this extension of peer communities as analogous to earlier extensions of the franchise in other fields, such as women's suffrage and patients' rights.

Extended peer communities are already being created, in increasing numbers, either when the authorities cannot see a way forward, or when they know that without a broad base of consensus, no policy can succeed. They are called citizens' juries, focus groups, consensus conferences, or any one of a great variety of other names; and their forms and powers are correspondingly varied. But they all have one important element in common: they assess the quality of policy proposals, including a scientific element, on the basis of the science they master combined with their commitment and the knowledge of the ways of the world. In addition, their verdicts all have some degree of moral force and hence political influence.

These extended peer communities will not necessarily be passive recipients of the materials provided by experts. They will also possess, or create, their own “extended facts”. These may
include craft wisdom and community knowledge of places and their histories, as well as anecdotal evidence, neighbourhood surveys, investigative journalism and leaked documents. Such extended peer communities have achieved enormous new scope and power through the Internet. Activists scattered among large cities or rainforests can engage in mutual learning and coordinated activity, providing themselves with the means of engagement with global vested interests on less unequal terms than previously. Along with the regulatory, evaluative function of extended peer communities, another, even more intimately involved in the policy process, is springing up. Particularly at the local level, the discovery is being made, again and again, that people not only care about their own environment and health but can also become quite ingenious and creative in finding practical, mixed social and technological means for its improvement; and local people can imagine solutions and reformulate problems in ways that the accredited experts, with the best will in the world, do not find “normal”. This is most important in the phases of policy-formation, and also in the implementation and monitoring of policies. Thus in addition to extending the traditional processes of quality assessment, participants can enhance the quality of the problem solving processes themselves.

Thanks in large part to the new powers created by information and communication technologies and the new social media, the extended peer community can on occasion inject itself into quality debates of a traditional sort. A notable instance of this tendency is the activity of the so-called deniers in the debate on global warming science. Some have focused on the evidence adduced for the phenomenon of global warming, independently of whether it is actually happening. The most famous of these is Steve McIntyre, a retired mining engineer, who has criticized the tree-ring analysis conducted by Michael Mann and others. Part of his criticism was about the persistent refusal of Mann to supply him with his data, which were obtained on public funds and
so would be presumed to be in the public domain. His more serious criticism related to the statistical techniques underlying the “hockey stick” picture of global temperatures in recent centuries. Although his arguments were never formally recognized by the academic scientists he criticized, he did gain important respect from key science journalists (Harrabin, 2011).

McIntyre has his own site in the “blogosphere”, but he and others became a powerful presence through popular websites like “Watts Up With That?” (http://www.wattsupwiththat.com accessed 10.9.13), which have attracted millions of visits and also won prizes for quality. In this way, the new social media have given strength to the extended peer community in science in a way reminiscent of the contribution of printing to the Reformation.

Quality assurance, thus, can be seen as a core commitment of post normal science, replacing “truth” as science’s ultimate regulative principle (Funtowicz & Ravetz, 1993). Defined in terms of uncertainties and decision-stakes, quality assurance encompasses public interest, citizen, and vernacular sciences. In a period of domination by globalized corporate science (Gibbons et al., 1994), this effort to make scientists accountable to interested groups presents a coherent conceptual alternative for the survival of the public knowledge tradition of science. Collegial peer review is thereby transformed into review by an extended peer community. This new form of quality assurance will be given its institutional structure and routines by those heterogeneous actors who put it into practice.
Bibliography


