

► Society) and the Geological Society of London had also begun to experiment with similar reports. It was a geologist, George Greenough, who introduced the term ‘referee’ in 1817, importing into science a term he knew from his days as a law student³. But it was the Royal Society’s system of reports that caused the British scientific world to take notice. The practice gradually spread to other societies, including the Royal Society of Edinburgh and the Linnean Society of London. But it was not really until the twentieth century that journals unaffiliated with any society slowly followed suit.

ANONYMOUS JUDGES

The struggle between Whewell and Lubbock represented two distinct visions of what a referee might be. Whewell was the authoritative generalist, glancing down on the landscape of knowledge. He was unconcerned with — and probably not in a position to critique — the details. Such referees were, according to the Royal Society’s president, “Elevated by their character and reputation above the influence of personal feelings of rivalry or petty jealousy”⁴. Lubbock was a younger specialist, Airy’s equal. This allowed him to take a fine-tooth comb to Airy’s arguments; it also put him in the position of reviewing a direct competitor.

Initially, Whewell’s vision won out. But the system began to transform even as it lurched into existence. After a couple of years, the reports became shrouded in secrecy. The last *Proceedings* issue to include one was in mid-1833, and no negative reports were ever published. A letter Whewell wrote in 1836 shows that he himself had changed his view: he describes the referee as a defender of a society’s reputation, working behind the scenes to exclude publications that do not belong. Neither the Royal Society’s archives — nor the personal papers of those involved — are clear on how this happened, but we should not be surprised that it did. In England, unlike France, there was little precedent for public authorities judging from on high what constituted good or bad science. Signing one’s name to explicit criticism of a colleague would have been ungentlemanly.

More familiar was the anonymous critic who purported to speak for the public, epitomized by the anonymous book reviews that dominated English periodicals throughout the period, from the *Quarterly Review* to the lowly *Mechanics’ Magazine* (the practice survives today in *The Economist*). Through anonymity, as one uncredited editor argued in 1833, “the individual is merged in the court which he represents, and he speaks not in his own name, but *ex cathedra* (with full authority)”⁵. Justifications of the anonymity of the scientific referee took a similar view.

It took just a decade for the referee to become an established scientific persona,

and not a noble one. An 1845 exposé in a London magazine painted a picture of referees as scheming judges quite possibly “full of envy, hatred, malice, and all uncharitableness”. Hidden away in some secret chamber, this scientific judiciary, the article implied, used the cover of anonymity to advance their personal interests — perhaps through undetectable acts of piracy — at the expense of helpless authors⁶.

It was only near the turn of the twentieth century that the idea began to take hold that editors and referees, taken as one large machinery of judgement, ought to ensure the integrity of the scientific literature as a whole. Amid calls to curtail the “veritable sewage thrown into the pure stream of science” (a suggestion⁷ by the physiologist Michael Foster in 1894), English scientific societies debated combining their publishing apparatuses, with a standardized referee system overseeing all of scientific publishing. (The plan was abandoned, in part because it would have meant convincing publishers of independent journals, such as the *Philosophical Magazine*, to go out of business.)

“The referee was reimagined as a universal gatekeeper with a duty to science.”

Nonetheless, the referee was gradually reimagined as a sort of universal gatekeeper with a duty to science. As this idea gained ground, many began to worry that the system itself might be intrinsically flawed, a force that impeded creative science and which ought to be abolished. Such worries culminated in what was surely the first formal inquiry into the workings of referee systems — in 1903, by the Geological Society of London. The inquiry found that opinion was sharply divided on the subject, receiving several vitriolic statements about the injustices and inefficiencies of the systems in use. The ‘referee’ was in such disrepute that they nearly banned the use of the term in all society business.

But referee systems survived, and were slowly set up by independent journals as well. Outside the Anglophone scientific world, referee systems remained rare. Albert Einstein, for example, was shocked when an American journal sent a paper of his to a referee in 1932. The idea that any legitimate scientific journal ought to implement a formal referee system began to take hold in the decades following the Second World War.

APOTHEOSIS AND FALL

In the 1960s, refereeing emerged as a symbol of objective judgement and consensus in science. The referee was, in the words of the physicist and science writer John Ziman, “the lynchpin about which the whole business of Science is pivoted”⁸. Just as in 1830s England, the relationship of science to the public was at the foreground of these

changes. The scientific community was once again working hard to solidify perceptions of its role in society. The very phrase ‘scientific community’ dates from this time. Researchers wanted to preserve autonomy while holding on to the massive government funding that had come their way since the Second World War. Allocations for basic research in the United States, for instance, swelled by a factor of 25 in less than a decade⁹.

‘Peer review’ was a term borrowed from the procedures that government agencies used to decide who would receive financial support for scientific and medical research. When ‘referee systems’ turned into ‘peer review’, the process became a mighty public symbol of the claim that these powerful and expensive investigators of the natural world had procedures for regulating themselves and for producing consensus, even though some observers quietly wondered whether scientific referees were up to this grand calling.

Current attempts to reimagine peer review rightly debate the psychology of bias, the problem of objectivity, and the ability to gauge reliability and importance, but they rarely consider the multilayered history of this institution. Peer review did not develop simply out of scientists’ need to trust one another’s research. It was also a response to political demands for public accountability. To understand that other practices of scientific judgement were once in place ought to be a part of any responsible attempt to chart a future path. The imagined functions of this institution are in flux, but they were never as fixed as many believe. ■

*Alex Csizsar is associate professor of the history of science at Harvard University, Cambridge, Massachusetts, USA.
e-mail: acsizar@fas.harvard.edu
Twitter @alexcsizar*

1. W. Whewell to P. M. Roget, 22 March 1831; Royal Society of London Library [DM/1].
2. J. W. Lubbock to W. Whewell, 27 January 1832; Trinity College Library, Cambridge [a/216/61].
3. George Greenough Papers; University College London [Add. 7918/1621].
4. *Proc. R. Soc. Lond.* **3**, 140–155 (1832).
5. *New Monthly Magazine* **39**, 2–6 (1833).
6. *Wade’s London Rev.* **1**, 351–369 (1845).
7. *Nature* **49**, 563 (1894).
8. Ziman, J. *Public Knowledge: An Essay Concerning the Social Dimension of Science* (Cambridge Univ. Press, 1968).
9. Kaiser, D. *Nature* **505**, 153–155 (2014).
10. Baldwin, M. *Making Nature: The History of a Scientific Journal* (Univ. Chicago Press, 2015).

CORRECTION

In the Comment ‘Cracking the Indus script’ (A. Robinson *Nature* **526**, 499–501; 2015), the sentence about Bryan Wells’s estimate of the number of Indus script signs has been revised to more accurately reflect Wells’s contributions. He estimated the number of signs at 676, not 958.