

approaches. They should report a comprehensive range of emissions sources and sinks. The Paris agreement asks nations to provide this information. Ideally, carbon balance sheets should address consumption-based emissions and account for the leakage or export of emissions (when entities transfer production of emissions to other locations through trade, for example), estimated at the national level to be three times the physical quantity of traded goods⁷. Some local governments, such as that of King County in Washington state, which includes Seattle, are starting to incorporate these considerations into their commitments.

Adding a function to NAZCA that evaluates the degree of implementation for each climate action would help to compare and identify actions that are being carried out and those that exist only on paper. (This could use the ranking system of the Climate Action Tracker, which rates national climate pledges as 'inadequate', 'medium', 'sufficient' or 'role model'.) Data sets on countries' populations, gross domestic products and land areas, for instance from the World Bank, could be linked with NAZCA data to provide visualizations and metrics to understand the breadth and scope of global climate actions.

By promoting meaningful and accurate record keeping, NAZCA could become the gold standard for climate-action reporting networks. ■

Angel Hsu is assistant professor at Yale-NUS College and Yale School of Forestry and Environmental Studies, Singapore. Yaping Cheng and Amy Weinfurter are research associates and Kaiyang Xu and Cameron Yick are research assistants at the Yale Data-Driven Environmental Solutions Group, New Haven, Connecticut, USA. e-mail: angel.hsu@yale.edu

1. Heede, R. *Clim. Change* **122**, 229–241 (2014).
2. Ecofys & University of Cambridge Institute for Sustainability Leadership. *Better Partnerships: Understanding and Increasing the Impact of Private Sector Cooperative Initiatives* (2015).
3. Schücking, H., Kroll, L., Louvel, Y. & Richter, R. *Bankrolling Climate Change* (urgewald, groundWork, Earthlife Africa Johannesburg & BankTrack, 2011); available at <http://go.nature.com/wpqvqz>
4. US Department of State. *2016 Second Biennial Report of the United States of America* (US Department of State, 2015); available at <http://go.nature.com/y7dom8>
5. Hoffmann, M. J. *Climate Governance at the Crossroads* (Oxford Univ. Press, 2011).
6. Espey, J. et al. *Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development* (Sustainable Development Solutions Network, 2015); available at <http://go.nature.com/oqk4yd>
7. Wiedmann, T. O. et al. *Proc. Natl Acad. Sci. USA* **112**, 6271–6276 (2015).



William Whewell, peer-review pioneer.

PAUL D. STEWART/SPL

Troubled from the start

Pivotal moments in the history of academic refereeing have occurred at times when the public status of science was being renegotiated, explains **Alex Csiszar**.

Referees are overworked. The problem of bias is intractable. The referee system has broken down and become an obstacle to scientific progress. Traditional refereeing is an antiquated form that might have been good for science in the past but it's high time to put it out of its misery.

What is this familiar litany? It is a list of grievances aired by scientists a century ago. If complaining about the faults of referee systems is nothing new, such systems are not as old as historical accounts often claim. Investigators of nature communicated their findings without scientific referees for centuries. Deciding whom and what to trust usually depended on personal knowledge among close-knit groups of researchers. (Many might argue it still does.)

The first referee systems that we would recognize as such were set in place by English scientific societies in the early nineteenth century. But these referees were never intended to play the part of supreme scientific gatekeepers. That notion emerged in around 1900 (see 'Past notes'). It was exactly then that some began to wonder whether referee systems might be fundamentally flawed. In this sense, peer review has always been broken.

Today, with the debate about the future of peer review more fraught than ever, it is crucial to understand the youth of this institution. What's more, its workings and its imagined goals have evolved continually, and its current tensions bear the marks of this. The referee system has become a mishmash of practices, functions and values. But one thing

stands out: pivotal moments in the history of peer review have occurred when the public status of science was being renegotiated.

SCIENTIFIC PUBLICISTS

In 1831, William Whewell, a Cambridge professor and philosopher of science, proposed a scheme to the Royal Society of London. He suggested that it commission reports on all papers sent for publication in the semi-annual *Philosophical Transactions*. Written by teams of eminent scholars, these reports might, he argued, be “often more interesting than the memoirs themselves” and thus a great source of publicity for science¹. Besides, authors would be grateful to know that their papers would be read carefully by at least two or three people. The society was just then launching a new journal to be called the *Proceedings of the Royal Society*, a cheaper monthly periodical to include abstracts of papers presented at the society. It had pages to fill and seemed the ideal place for these new reports.

At the time, editors of scientific journals made publishing decisions by personal fiat, perhaps in consultation with some trusted helpers. For publications that belonged to a scientific academy or society — such as the *Philosophical Transactions* — the vote of some committee of eminent persons would determine a manuscript's fate. (The temptation to conflate these practices with modern referee systems has led to the stubborn myth that the origins of the scientific referee can be traced back as far as the seventeenth century.)

Whewell was not much concerned about preventing shoddy papers from being printed; he was not proposing a new mechanism to inform publishing decisions. Instead, he was one of many people campaigning to increase the public visibility of science and give a unified identity to the scientific enterprise in England. (It was he who, a few years later, coined the word ‘scientist’ to this end.) This movement had begun in 1830 and is now most remembered for Charles Babbage's *Reflections on the Decline of Science in England*, a screed about the paucity of state funding for, and public recognition of, science. But its more consequential legacy is the referee system.

Whewell was cribbing from a century-old custom at the French Academy of Sciences in Paris of writing reports that evaluated inventions and discoveries in the service of the king. There, researchers who were elected to the academy were paid by the state as a reward for scientific eminence, and politicians seemed to value their opinions. Indeed, to be an expert (a French word not yet common in English) was almost by definition to be a writer of reports. Whewell reckoned that those French *académiciens* must be doing something right.

The proposal to turn the Royal Society into a corps of expert judges in the

style of the French academy was met with enthusiasm. But translating the report-writing practice across the Channel proved more complicated than Whewell expected.

NEWS OR VIEWS?

Whewell agreed to write the first report. His collaborator was a former student at Cambridge, John William Lubbock, a mathematically inclined astronomer who was also the Royal Society's treasurer. They jointly selected a manuscript submitted by George Airy, another up-and-coming astronomer. The paper, ‘On an inequality of Long Period in the Motions of the Earth and Venus’, used sophisticated mathematical methods to calculate how the orbits of these planets were influenced by the gravitational force each exerted on the other.

Whewell and Lubbock took turns reading the manuscript — copying technologies at the time left much to be desired. Both instantly knew what they thought of it. And they completely disagreed.

They argued about the paper for months. Both wrote draft reports, which could not have been more different. Whewell's focused on the significance of the problem and on Airy's remarkable conclusions. Lubbock's picked at the inelegant ways in which Airy had constructed his equations. Most fundamentally, they argued about what a reader's report ought to be. Whewell wanted to spread word of the discovery and to place it in the bigger picture (think *Nature's* News & Views and *Science's* Perspectives). “I do not think the office of reporters ought to be to criticize particular passages of a paper but to shew its place,” he told Lubbock. If they picked out flaws, he warned, authors would be put off. Lubbock had other priorities: “I do not see how we can pass over grievous errors,” he wrote.

Feeling that they had reached an impasse, Lubbock went to the author himself to deliver his suggestions for improvement. Airy was understandably irritated that his manuscript was being subjected to this strange new procedure. “There the paper is,” he wrote to Whewell, “and I am willing to let my credit rest on it.” He had no intention of changing his text. Lubbock threatened to pull out, but ultimately relented and swallowed his criticisms, acknowledging that this was “the first report which the Council have ever made” and trying to see the bigger picture. He thanked Whewell for putting his “shoulder to the wheel” and signed his name to the report².

With disaster averted, Whewell's version of the report was read publicly at the society on 29 March 1832, and was printed in the *Proceedings*, while Airy's full paper appeared in the *Transactions*. Lubbock's critiques never became public.

Not long before, the Astronomical Society of London (now the Royal Astronomical



PAST NOTES
How organized academic review has evolved over 300 years.

1665 Henry Oldenburg, secretary of the Royal Society in London, creates the *Philosophical Transactions* to simplify his correspondence. He uses no referee system.

1699 France's Royal Academy of Sciences is given power by Louis XIV (pictured centre, with academy members) to report on and approve books for publication and bypass the royal censors.

1752 After vicious satires of the *Philosophical Transactions*, the Royal Society establishes a committee to vote on what to publish.

1831 Cambridge professor William Whewell convinces the Royal Society to commission public reports on manuscripts. Might referees increase the visibility of science?

1833 By now the reports have become private and anonymous.

1892 A pamphlet ‘On the Organisation of Science’ published in London by ‘A Free Lance’ kick-starts a movement to standardize the selection and distribution of scientific papers. Might referees be guardians of the literature?

1892 A paper surfaces that was rejected by a Royal Society referee in 1845, outlining the kinetic theory of gases more than a decade before James Clerk Maxwell's famous paper. Might referee systems be fundamentally flawed?

1968 British physicist John Ziman describes the referee as “the lynchpin about which the whole business of Science is pivoted”. Outside the United Kingdom and North America, many editors and scientists remain largely unconvinced.

1973 External refereeing becomes a requirement for publication in *Nature*¹⁰.

1991 An e-mail/FTP server at xxx.lanl.gov for freely sharing unreviewed physics preprints goes live. Later relocated to the web at arXiv.org, it becomes a touchstone for discussions about the end of peer-reviewed journals.

2006 *PLoS ONE* launches as an open-access journal that eschews ‘importance’ as a factor in peer review.

2007–11 *EMBO Journal*, the *Frontiers* series and *BMJ Open*, among other journals, experiment with open peer review, publishing reviewers' names or notes alongside papers.

► 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.