

JUDGEMENT DAY

Many governments are assessing the quality of university research, much to the dismay of some researchers.

BY BRIAN OWENS

Two years ago, academics at Lancaster University, UK, found themselves in the uncomfortable position of being graded. They each had to submit the four best pieces of research that they had published in the previous few years, and then wait for months as small panels of colleagues — each containing at least one person from outside the university — judged the quality of the work. Those who failed their evaluations were offered various forms of help, including mentoring from a more experienced colleague, an early start on an upcoming sabbatical or a temporary break from teaching duties.

The university did not undertake this huge exercise just to make sure that the researchers

were pulling their weight. The assessment was a drill to prepare for the Research Excellence Framework (REF), a massive evaluation of the quality of research at every university and public research institute in the United Kingdom, which is set to take place in 2014.

The idea of the drill “was to identify areas where we could help people develop their profiles”, says Trevor McMillan, Lancaster University’s pro-vice-chancellor for research. Happily, he says, the results suggested that the university would score more highly than it did on the most recent national evaluation, in 2008.

But other mock evaluations have proceeded less smoothly. In a survey of more than 7,000 UK academics published on 3 October by the University and College Union (UCU) in

London, almost 12% reported having been told that failure to meet their university's REF benchmarks in a drill could lead them to be transferred to a teaching-only contract before the real REF (see go.nature.com/eqiirr). Almost 10% said that they faced denial of promotion. At Cardiff University, around ten academics were pressured to switch to teaching-focused contracts after they scored poorly on a practice exercise, so as not to drag down their department, says Peter Guest, an archaeologist at Cardiff and the university's UCU liaison on the REF. This form of game-playing is discouraged, but not expressly forbidden, by the REF — however, making career decisions solely on the basis of the evaluation is against the university's own policies, as well as those of many other institutions, says Guest.

All of the Cardiff cases were resolved in a day or two, with managers being “forcefully reminded” of the rules by the UCU, says Guest. But the experience shows how tempting it is for institutions to make career decisions on the basis of predicted REF scores, which are highly subjective. This is neither reliable nor fair, says Guest. (In response to questions about the incident, a spokesman for the university said in an e-mail: “We have been running a long-term programme for over four years to ensure our academic staff are on contracts that reflect what they actually do.”)

Even many academics who did score well in the mock evaluations resent them. Around the United Kingdom, researchers view these national assessments as a bureaucratic imposition that can stifle creativity.

UNDER PRESSURE

Most academics at Lancaster saw the mock REF as little more than a “mildly annoying” bit of bureaucracy, but the real thing is a different matter. “We have our department's top research professor working on preparing our REF submission, and it's taking up about a third of his time,” says one member of the mathematics and statistics department. “It seems like a waste of talent.” Too many researchers are focused on winning grants and trying to predict what kind of work will be rewarded in the next assessment, rather than doing the best science they can, says Dorothy Bishop, an experimental psychologist at the University of Oxford, UK. “I think a lot of science is just not very well done these days because people are trying to do too many things.”

But university administrators and the government have come to rely on these evaluations to help them decide how to disburse funding. And the idea has been so popular with educational leaders that other countries are following the United Kingdom's example, with similar exercises cropping up in Australia, Italy, Germany and elsewhere.

In the late 1980s, the United Kingdom became the first country to systematically evaluate the quality of its university research.

The REF is the latest incarnation of these check-ups. Previously known as the Research Assessment Exercise (RAE), the evaluations are widely credited with helping to improve the country's research system. Between 2006 and 2010, citations of UK articles grew by 7.2%, faster than the world average of 6.3%; and the country's share of citations grew by 0.9% per year, according to a 2011 analysis conducted by publishing company Elsevier for the government.

The assessment is used by the UK government to distribute more than £1.6 billion (US\$2.6 billion) a year in block grants to universities. More than 70% of the pot goes to the top-scoring 20 or so universities — last year, the University of Oxford got more than £130 million in quality-related funding — whereas the smallest, least research-intensive institutions make do with just a few tens of thousands of pounds. Assessment results are eagerly assembled into league tables, showing which universities are performing best in which subjects (see “Top 5”).

“The reputational aspects of it can be as important as the financial aspects,” says McMillan. Some smaller institutions that are strong in particular subjects — as Lancaster is in physics — have reported that they have an easier time attracting students in those areas as a result of the assessments. And it is not just students. “One of the consequences is that people really want to come to a department that did well in the RAE,” says McMillan. “We've found it easier to recruit high-quality staff in physics.”

For the REF, universities submit a selection of work from most of their active researchers to one of dozens of subject-specific panels known as Units of Assessment that correspond roughly, but not exactly, to university departments. The panels evaluate the quality of the research using peer review and metrics such as citation indexes. And they will also, for the first time, look at the economic and social impact of a university's research.

Even critics of the assessments agree that they have had some positive effects on the country's research system. Because the exercises judge academics on the quality of their

research, many departments have tried to cut back on other demands, such as administrative work, says Guest. Furthermore, the results make it clear which departments and academics are not pulling their weight, and allow universities to make strategic decisions about how to invest resources.

Royal Holloway, University of London, faced that very situation after the first research assessment in 1986, which ranked the university's psychology department in last place

nationwide, says Kathy Rastle, a cognitive psychologist and the department's director of research. Recognizing that it would not be able to boost its rating by hiring established stars, the department sought instead to attract and develop young talent. “We try to focus on people we feel have great potential,” says Rastle.

Early-career psychologists at Royal Holloway are now offered “substantial, but tailored” start-up packages, she says, with hardly any teaching commitments for the first two years. They also get help from more experienced colleagues in preparing funding proposals.

In the 2008 RAE, after two decades of nurturing junior staff, the department was ranked among the top ten in the country. It has ambitions to go even higher. “I look forward to the REF as an opportunity to show what we've done, and to move up the ranks,” says Rastle.

AN IDEA SPREADS

As other countries begin their own national research evaluations, they hope to achieve

the same kinds of benefits. This year, Italy published the results of an evaluation begun in 2011 (see *Nature* <http://doi.org/nrx>; 2013); its goal is to increase meritocracy in the country's universities, where academics of the same rank and seniority currently receive the same salary, regardless of output. “There are no incentives to improve your research performance,” says Giovanni Abramo, who studies bibliometrics and research evaluation at the National Research Council of Italy in Rome. “Now some of the money the government gives to universities will be based on this evaluation.”

The Italian effort evaluates only three journal articles from each researcher with teaching commitments, whereas Australia assesses all research output as part of its Excellence in Research for Australia (ERA) initiative, most recently in 2012. Only a relatively small pot of funding rides on the results: this year,

TOP 5

Analysts used information from the 2008 UK Research Assessment Exercise to rank academic departments by quality.

CHEMISTRY

1. University of Cambridge
2. University of Nottingham
3. University of Oxford
4. St Andrews/Edinburgh
5. University of Bristol

MATHEMATICS

1. Imperial College London
2. University of Warwick
3. University of Oxford
4. University of Cambridge
5. University of Bristol

PHYSICS

1. Lancaster University
2. University of Cambridge
3. University of Nottingham
4. University of St Andrews
5. University of Bath

PSYCHOLOGY

1. University of Cambridge
2. University of Oxford
3. University of Birmingham
4. University College London
5. Birkbeck

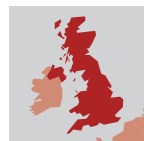


IMPACT

A Nature special issue
nature.com/impact

STAND AND BE COUNTED

Assessment of academics has spread throughout the world, but each country does it differently.



UNITED KINGDOM

Name: Research Excellence Framework

Next assessment: 2014

A sample of researchers submits four examples of work published since 2008;

department heads often decide who is included. Each department provides a description of the economic and social impact of its work.

Submissions are evaluated by expert panels that assign a quality profile to each university. Quality will account for 65% of the score, impact for 20% and research environment 15%.

Results are made public and used to distribute the government's 'quality-related' research funding, which in 2013 was worth more than £1.6 billion (US\$2.6 billion).



GERMANY

Name: Research Rating

Next assessment: Unknown

Four pilots of the Research Rating have been conducted so far, in chemistry, sociology, electrical engineering and

English and American studies.

Panels of 15–20 people evaluated the quality of a selection of publications from each research institute. Panels also sought to promote young researchers and technology transfer.

The government is considering a decision this month on whether to repeat the assessment and expand it to all disciplines.

The assessments are made public, but will not be used to distribute funding.



ITALY

Name: National Agency for the Evaluation of the University System and Research (ANVUR)

Next assessment: Unknown
Researchers submitted three selections of their work — six if

they had no teaching commitments — published between 2004 and 2010.

The research outputs were evaluated by 14 subject panels. Science panels made extensive use of bibliometrics. Large, medium and small universities were ranked separately, as were research agencies and inter-university consortia.

The results were made public and were used to distribute around €540 million (US\$729 million) as part of the 2013 university budget.



AUSTRALIA

Name: Excellence in Research for Australia

Next assessment: 2015

Universities track every piece of research output from their academics; more than 400,000

pieces were submitted in 2012.

Output is reviewed by expert panels, using metrics such as citation counts and patents filed, as well as research funding and signs of prestige including researchers' membership of learned academies.

The results are released publicly to allow comparisons between institutions, but just Aus\$68 million (US\$64 million) is distributed according to the outcome.

rankings determined the disbursement of just Aus\$68 million (US\$64 million). The outcome is mainly used to give institutions an idea of where they stand in terms of national and international quality, says Aidan Byrne, chief executive of the research council.

The exercise has added benefits, he says. For example, it helps to verify that the council is distributing its Aus\$800-million competitive-grants portfolio in a reasonable way. With a round of assessments costing Aus\$4 million,

the UCU's survey is the stipulation by many universities that researchers must have produced four high-quality publications between 2008 and 2013, says Stefano Fella, a national industrial-relations official at the union. Of the academics polled, 67% felt that they could not produce the required output without working excessive hours — and 34% said that the stress was affecting their health. Many have reported changing how they approach their work, says Fella — for

“You should do good science, and not think in this appallingly strategic way.”

says Byrne, “it’s a very efficient method of quality control”. Although there is no formal connection between the ERA and the grants process, the academics who peer-review grant applications are aware of ERA outcomes, and that feeds into their decisions, he says.

GROWING PAINS

It is too early to know how the newer assessment efforts in Italy, Australia and other countries will affect the research environment there (see ‘Stand and be counted’). But researchers say that they have seen enough of the long-lived UK programme to know some of the downsides.

One of the main worries that came up in

example, some might have rushed to get a publication in the assessment period, even if the work might have benefited from more time. “They don’t think about the best way to present the work,” says Fella, “but what would be best for the REF?”

Frederic Lee, an economist at the University of Missouri–Kansas City, has studied how the UK research-assessment system has affected his discipline. He experienced two rounds of assessments first-hand while working at De Montfort University in Leicester in the 1990s. He says that economists who study alternative theories such as Marxism have been squeezed out because the assessment has consistently favoured mainstream work at elite institutions,

published in a small subset of journals. “There has been a lemming effect that has led to a homogeneity of research topics,” he says.

Lee says that he was never pressured to abandon his research on the history of heterodox economic theories in the United Kingdom, but was encouraged to submit his work to particular mainstream journals, where it stood a slim chance of getting accepted. Other academics have told him that they have been pressured to switch to more conventional research topics, and some had been squeezed out of departments at major institutions. *Nature* spoke to one economist at the University of Manchester who studies alternative theories, and who left the department in part because the focus on RAE-friendly theories meant that prospects for advancement seemed essentially non-existent.

Academics are particularly worried about the move to assess the impact of research in the REF. They fear that this signals a preference for short-term, applied work over basic research that has no obvious, immediate public benefit. “As far as I’m concerned, you should do good science, and not think in this appallingly strategic way,” says Bishop. “Some good science takes a long time to do well.”

The time, effort and money being spent on submissions are also a major concern: preparations for the 2008 RAE cost universities £47 million, according to a 2009 review of the exercise. Even smaller universities such as Lancaster asked several academics to spend months reviewing applications for mock REFs. The time burden can be even worse for administrators, who might have to hire extra staff to work on the REF, says Bishop. University College London, for example, has recruited four editorial consultants to work on the impact portion of the assessment.

McMillan says that it is natural to spend a bit more time and money when preparing to tackle a new criterion. “It’s a dimension that we’re not used to.” He adds that administrators at Lancaster are hiring external professional editors to help with only the final part of the process: polishing the case studies and impact statements that are written by academics and the university’s research support office. Still, McMillan himself is currently spending two to three days a week tweaking Lancaster’s submissions. “I think the REF is probably taking up more time than previous exercises,” he says. “The shift to the impact agenda has seen a big increase in the workload.”

But some universities have seen the benefits of all that work. The vast improvements made by Royal Holloway’s psychology department demonstrate how much periodic evaluations can help, says Rastle. “Having the REF hanging over our heads makes sure we take all the steps we can to get the best out of our people.” ■

Brian Owens is a freelance writer based in New Brunswick, Canada.