

## **Of Models and Metrics: The UK Debate on Assessing Humanities Research**

**Michael Worton, UCL (University College London), UK**

### **Context**

In the UK, research in universities receives government funding from two sources. In both cases, the allocation of funding is determined on a selective – and highly competitive – basis, although the criteria for assessment are somewhat different, as, indeed, are the modes of assessment.

Higher Education Funding Council Quality Related (QR) Funds are allocated on a basis of past achievements as measured and assessed by the Research Assessment Exercise (RAE). In other words, this allocation process is based on a retrospective evaluation of research performance.

On the other hand, Research Councils' funds are distributed on the basis of evaluations of individual grant applications, where the funding decisions are prospective and based on promise, on the quality and potential of the proposed project.

QR funds are allocated to the Higher Education Institutions (HEIs), enabling Vice-Chancellors and their Senior Management Teams to decide on how to allocate the funds; QR funding thus gives them flexibility in decision-making regarding funding allocations within their institutions, since they do not have to allocate funds earned by one particular unit of assessment to that unit, but can choose which areas to strengthen and invest in. Research Council funds, however, must be used for the specific project to which they are allocated. In the Dual Support system, as it is known, QR funds can be used to improve infrastructure and/or salaries, and to build research areas in a strategic way, whereas Research Council funds go to support the individual projects.

The fact of having the Dual Support mechanism for research is important for the long-term sustainability of a world-class research culture in the UK, since

it guarantees the integrity and autonomy of research conducted in HEIs. Furthermore, there is a crucial interaction between the funding councils' QR funding streams and the research councils' funding streams, in that many of the excellent projects supported by research councils have been initiated and developed using QR funding. The success of research councils funding is therefore to a great extent dependent on the health of the Dual Support system.

In the *Science and Innovation Investment Framework* (July 2004), the UK Government emphasised its commitment to the Dual Support system. Two years later, in a follow-up policy statement, *Science and Innovation Investment Framework: Next Steps* (March 2006), the UK Government reiterated this policy, announcing early action towards the greater use of metrics in allocating QR funds through the higher education funding bodies, in particular. In the accompanying consultation paper, there was a recognition that metrics are generally less well developed and less straightforward in the arts and humanities, and in the social sciences than in the scientific, technological, engineering and medical (STEM) disciplines. It was therefore proposed that work should be undertaken to develop a more differentiated approach to recognising and awarding research excellent in the former group of disciplines, which might even be allowed to continue with a form of the established RAE, rather than moving towards metrics, as would the STEM subjects. The significant difference between the STEM subjects and arts and humanities is the much greater proportion of research funding allocated to research in arts and humanities through QR (80%) than through Research Council funds (20%).

In June 2006, the Chief Executives of the Arts and Humanities Research Council (AHRC) and the Higher Education Funding Council for England (HEFCE) jointly established an expert group to explore the use of metrics in the assessment of arts and humanities research.

### **The Debate and the Group's Recommendations**

Membership of the group was chosen to reflect the diversity of the arts and humanities research community. In its first meetings, the group's members discussed the distinguishing characteristics of research in the arts and humanities, and how these might be recognised and reinforced through assessment, they discussed the use of metrics, and how these might best be applied across the disciplines, and they deliberated on how to ensure equal opportunity for all those assessed, at whatever point of their career. Despite the diversity of the group, it soon became clear that wide consultation would be necessary to reflect properly the varied disciplinary perspectives of the research assessment across the subject areas.

In September 2006 the group undertook a wide-ranging consultation exercise with major groups of academic stakeholders. Reviews that emerged from these consultations involved many of the key principles and proposed operational features of the assessment framework for research in arts and humanities that the group put forward in its report.

One of the key arguments of the report (and which was a view shared by all those consulted) is that there is no fundamental difference in the nature of the research enterprise in the STEM disciplines on the one hand, and the arts and humanities on the other. Rather, all of these disciplines represent a continuum of research endeavour, along which methods and resource requirements vary in ways that do not map easily onto the current subject divisions. The demand for research inputs varies along the spectrum from resource-intensive disciplines, such as chemistry and archaeology, to non resource-intensive disciplines such as mathematics and philosophy. The disciplines that make up the arts and humanities are certainly distinctive in their approaches and concerns but they should not be considered exceptional.

In the group's view, it should be possible to provide a broad framework of assessment that applies to all disciplines. However, the nature and scope of the elements of that framework should be sensitive to the distinctive characteristics of each discipline, such as the size of the research community,

its demand and need for inputs, the various inputs available to it, publication patterns, and the nature and organisation of the research process.

A key element in the debate was the fact that the research landscape in the UK has evolved enormously and in ways which could not be anticipated when the RAE was established twenty years ago. This is particularly true for arts and humanities, since the creation first of the Arts and Humanities Research Board in 1998, and then its successor, the Arts and Humanities Research Council in 2005. There has been significant growth in collaborative and interdisciplinary research; increasing use is made of information technology in all aspects of the production and organisation of research; greater emphasis is laid on the dissemination of research outputs and outcomes beyond the research community itself; doctoral students are now much more frequently included in teams working on research projects. All of this means that a more holistic approach to research assessment is now required.

It was recognised, that for the moment, research outputs (monographs, edited volumes, journal articles, exhibitions, performances, etc) remain the most reliable indicators of research quality. However, it was equally recognised that with time, experience, and further research, credible quantitative methods could emerge. Indeed, while it would be necessary to retain the application of human judgement through peer review process as an element in the overall assessment framework, we could already move to using some metrics.

There remain some powerful background anxieties in the research community. First of all, there is a prevalent assumption that the term “metrics” refers only to measurements of either research income or bibliometric data – and it is undoubtedly true that these latter data do not as yet fully capture the range of research activity. For instance, bibliometric data are seen as rarely able to recognise new and innovative contributions by researchers at an early stage in their career. Furthermore, there are fears that the use of external research funding as a metric would tend to privilege empirical or laboratory-based studies over desk-based, theoretical work. Another anxiety is that large-scale collaborative projects would be privileged over small-scale “lone

scholar” projects. However, our argument is that these anxieties are unfounded if one takes a holistic look at the totality of the research process – from inputs to activity to outputs to outcomes, and within this holistic assessment, one would need to establish an appropriate balance of metrics and expert judgements to enable a robust and credible profile of research performance to emerge.

A proposed framework would consist of evaluation of the following: research outputs; spend on research infrastructure and other funding of the research environment; peer-reviewed external research income (from the research councils, but also from other peer-reviewed sources, such as charitable foundations, overseas funding agencies, etc); and evaluation of the wider social, cultural and economic significance of the research process; PhD completions per research-active member of staff; esteem indicators (such as election to national bodies; membership of editorial boards; invitations to give named lectures, large lecture series etc). Many of these can already be measured by metrics, whereas for others metrics are being developed that should, within a few years, be robust enough to be used in funding allocation processes. So while metrics alone will not allow the overall performance and quality of research to be assessed at individual or departmental level, metrics nonetheless have an important role to play in research quality assessment, both in themselves and in their value in providing the evidence to inform the expert judgement of reviewers.

While peer review processes specific to the assessment of research outputs should be retained for the [immediate future, it is urgently necessary to relieve the assessment burden on reviewers, since this burden has become unsustainable. This must be done in a way that does not weaken confidence in the process; indeed, any alleviation of burden must be seen actively to enhance the effectiveness of the peer review process as a whole. We therefore proposed that rather than having relatively small groups of peer reviewers to assess all submissions in a particular unit of assessment, the peer review should be much more distributed. One could, for instance, take the model of the AHRC peer review college, which currently has five hundred

peer reviewers who can be called upon to review individual research applications. Such peer review colleges include specialists with a much broader range of expertise than it is possible to represent on any RAE panel; they also include international assessors, and, crucially, relevant non-academic experts. One of the greatest advantages of using the established, standing, bodies of peer reviewers is that the uniform process of induction and training for such peer reviewers will lead to a greater consistency of reviews and outcomes.

A further recommendation of the group (and one which is somewhat controversial) to reduce the current assessment burden on reviewers is to sample the submitted outputs from individual researchers.

Our proposed changes in the approach to peer review are not designed simply to reduce costs or to alleviate the burden imposed by the assessment of outputs; rather, they reflect a shift of emphasis on what is being assessed. The twin aims of moderating the current level of concentration on outputs and of focusing also on other areas of the research process are advanced as means of enabling a balanced assessment framework that captures more fully the totality and sustainability of the research process.

On bibliometrics, the group's research uncovered the clear deficiencies of commercial citation indices in terms of their coverage of arts and humanities outputs, which make the use of bibliometric indicators for assessment purposes highly problematic at present. However, we signalled the promising developments in public-access bibliometric tools and public initiatives both nationally and in the UK and internationally. One such initiative is the European Science Foundation Member Organisations' development of a European Reference Index for the Humanities, which over the next few years should provide, for example, robust ways of comparing the research performance of different countries and thereby underpin further requests for further funding within their national communities. With regard to the UK's QR funding allocations for research, we recommended a funding cycle of between five to seven years in order to ensure stability of institutional planning. The

group also recommended that as appropriate metrics were developed, they should be collected and used annually within HEIs, since they would provide timely and accurate information for institutional managers on all aspects of research quality.

## **Conclusions**

The group's emphasis on the need to focus on the totality of the research process was welcomed by the UK research community, which also recognised that as a consequence of changing attitudes towards the breadth of the research process, a new framework of assessment needs to be developed. Any new framework must also take account of the fact that the research landscape has evolved considerably over the last twenty years and continues to evolve dynamically. For this reason, a process more holistic than at present needs to be established, Peer review (or expert judgement) will continue to play an important part in research assessment, but it should be increasingly distributed and should also be increasingly informed by metrics, which will enable judgements to be more robust.

Much work remains to be done, and the UK's Department for Education and Skills (DfES) and HEFCE are already working on drawing up the outlines of a new assessment process. The group has strongly recommended that a new metrics-informed process should be used to make a comparison with the results of the 2008 RAE and also be mapped against the finding of the 2001 RAE. In this way, confidence will be built in the research community in regard to metrics and their effectiveness as evaluation tools. Above all, by moving steadily to a more metrics-informed framework of research assessment, we shall ensure that expert judgements are increasingly based in objective evidence, thereby providing greater transparency of the funding decisions outcomes that play such a crucial role in encouraging and rewarding the world-class research done in the UK in the arts and humanities.

END