The publication game: acceptable and not-acceptable in the British REF exercise

TONY TOWNSEND

Every five years, universities in the UK submit themselves to an exercise that produces league tables to determine the levels of funding support provided for research infrastructure by government, based on the previous five years of research activity. In 2013, the next exercise, called the Research Exercise Framework, will see in each university, for each eligible staff member, up to four pieces of research output 'briefly defined as a process of investigation leading to new insights, effectively shared' submitted and reviewed by a national panel established for each of 36 specialist discipline areas.

Judgement of the quality of the research work being done by each university, and the subsequent government funds that flow from this process, is based on the decisions made by these discipline groups. This article looks at the issues that are implicit in judgements made about quality research output in educational leadership and how such an exercise might restrict researchers in educational leadership to certain types of research activity.

Introduction

In 2013, universities in the UK again will submit themselves to an exercise that identifies league tables based on research outputs. This next exercise, called the Research Exercise Framework (REF), will determine the levels of funding support provided for research infrastructure by government, based on the previous five years of research activity. Around 10 billion pounds will be allocated to support universities to do research, to provide the equipment, support personnel and some local funding to kickstart young researchers, and the level of funding that each university gets will be based on the previous five years of outputs, generally assessed through publications in quality journals and other outlets, the amount of research funding that the university was able to attract and the number of research students the university has supported, and how they did these things now will be included a recognition of the impact that this research has had on the social and or economic development of the community beyond academe.

Since the First World War, when many of the universities that existed at the time were decimated, and increasingly since the Second World War, when the same thing happened, universities in the UK have received government support. Margery Fry, who was the only continuously serving member of the UK University Grants Committee at the time of writing her paper for Higher Education Quarterly in 1948, identified nine propositions to describe relations between universities and governments:

- (1) The provision of adequate teaching, study and research on a university level is essential to the well-being of a modern state.
- (2) The government of each state has the responsibility of seeing that such provision is made.
- (3) In Britain the cost of this provision has now far outstripped the total of endowments and of what students can pay, or local authorities or private benefactors be expected to give.

- (4) The state must therefore aid the universities on an increasingly generous scale.
 - (5) Public expenditure must, in the long run, be subject to public control, i.e. to parliament.
 - (6) Universities should not be subjected to interference in the interests of any political party.
 - (7) Public money should not be used to compete against itself.
 - (8) Universities are the best judges of their own activities, whether as to subjects to be studied, standards to be maintained, staff to be appointed or research to be encouraged. They must therefore control the expenditure of their own funds.
 - (9) Every necessary subject should be taught at some university, 'rare' subjects should not be wastefully duplicated.
- (Fry 1948: 221–222)

This post war expansion of university education, described in Fry (1948), continued until the late 1970s, when the Thatcher government changed the direction of government funding. Since research accounted for around 40% of the funding allocated to higher education institutions (HEIs), it received higher levels of scrutiny. It was argued that universities were not accountable for how money was spent and that the money that was devoted to research was spread too thinly, across too many institutions, to produce world-leading research (Johnston 1995, Mace 2000).

The conservative government's solution was to introduce competition into university education in the same way that it introduced it to the school and health systems. Of the nine propositions listed above, most of them remain almost intact to this day. The one that can probably be removed from the list now is number 7 Public money should not be used to compete against itself.

Competition was promoted by government through the separation of teaching budgets (T) from research budgets (R) and

... from 1993 the government were to purchase T on the basis of student numbers (quantity) and R on the basis of research volume and quality. (Williams 1992, in Curran 2000: 389)

The REF, the latest version of the government's use of competition to fund university research, is used to fund the universities, not research itself, and is determined based on a university's previous research outcomes. The REF is different to past assessments, as Bekhadradnia (2009: 3) reports:

Each Research Assessment Exercise has built on and developed its predecessor, some with significant changes. For example, the 1992 exercise reduced the number of panels to 69 from more than 150, and evaluated applied research separately. In 1996, the publication count was dropped (in the 1992 exercise the number of publications per researcher was required to be submitted, giving rise to concerns that the exercise was giving rise to a 'publish or perish' culture).

In 2001, the number of research outputs submitted was reduced to 4, and the most significant change occurred in 2008 with the new profiled grading system, which represented a big improvement on the sharply delineated step scoring that had preceded it, but which led to some reduction in the concentration of research funding

because it more sensitively recognised research strengths that were previously hidden.

Once again, universities, colleges, schools, faculties, departments and individual academics are gearing up to play the game, one aptly described by the television programme *Yes Minister*:

Minister: ... how do I show I manage our national research programme? The PM says he wants vigorous leadership.

Sir Humphrey: Well, we could try a spot of competition. You know, make each university compete for funds against all the others. We just need a way of funding the universities unequally.

That should do the trick! (Sugden 1997: 367)

It is clear that the sentiments expressed above are translated into the real world as well. As Sir Howard Newby, Chief Executive of Higher Education Funding Council for England (HEFCE), put it;

However, when we look at higher education as a whole, we cannot possibly sustain world class research in the 90 universities in England let alone the colleges of higher education elsewhere. It is simply again not practical to think in those terms. What we have to do instead is to encourage institutions to identify their real strengths and focus on them and, if necessary, to encourage more collaboration between institutions in order that collectively they are delivering on these agendas even if individually they may be focusing on particular areas.

(Select Committee on Education and Skills 2003: question 445)

It could be argued that this is a game where the rich get richer and the poor get poorer, where the old established universities that have done well previously and have been rewarded by their ability to establish excellent research infrastructure and to attract the highest quality academics from around the world, are competing with newer, smaller or less established universities, using the same rules, as if it was a level playing field. However, it seems as if it might be anything but this. Paul (2008: 324) reported:

Analysis of the funding allocations shows the usual Pareto type distribution of funding, with the four universities with the highest income (Cambridge, Oxford, University College London and Imperial College) receiving nearly 30% of the total (Cambridge received £107 million), the top 10 receiving over 50% of the total, the top 20 receiving over 70% and the bottom 80 out of 130 or so institutions receiving less than 6%.

But it is not only the winners and losers within the game that might draw questions, it is how the game is played as well. It is a game where 'we are judged primarily by other academics, on the basis of publications read only by other academics and research grants awarded by academics to academics!' (Piercy 2000: 31) which 'captures department reputations and sets them out in public form ... This exposure has left some (institutions and individuals) winners and some losers – whilst others are still learning to play the game!' (Henkel 1998: 109). However, the game is a somewhat strange one because those 'doing the judging are from other institutions in

the same sector, essentially competing for the same resources' (Bence and Oppenheim 2004: 363).

Boaden and Cilliers (2001:6) argue that: Academic research quality is traditionally measured through peer review. The quality of research can therefore be assessed by considering the number of publications in journals, and the quality of those journals.

The nature of the exercise becomes one where a peer review process determines the level of research support a university can generate, and is one where the rules of the game have altered every five years since the exercise began. In 2013, the new game commences, with the major new feature being impact. Quality is no longer the only criterion, although it still accounts for 65% of the scoring of an institution. Now 20% is allocated based upon the social and economic impact that the research produced in universities is judged to have.

The REF

The Research Excellence Framework (REF) is the new system for assessing the quality of research in 36 different units of assessment (sub-panels) within four main panels of assessment in higher education institutions in the UK. It is conducted jointly by the HEFCE, the Scottish Funding Council, the Higher Education Funding Council for Wales and the Department for Employment and Learning, Northern Ireland and replaces the Research Assessment Exercise (RAE), last conducted in 2008. The REF is to be used for three main purposes:

- (a) to inform the selective allocation of the funding councils' grant for research to the institutions which they fund, with effect from 2015–16.
- (b) to provide accountability for public investment in research and to produce evidence of the benefits of this investment.
- (c) to provide benchmarking information and establish reputational yardsticks, for use within the higher education (HE) sector and for public information. (REF 2011: 5)

The REF describes a single framework for assessment across all disciplines, with a common set of data required in all submissions, standard definitions and procedures and assessment by expert panels against broad generic criteria. The panels of experts are researchers who are currently or have recently been active in high quality research or its wider application and they are expected to use appropriate quantitative indicators to support their professional judgement. However, it is argued that expert review remains paramount and since there are differences in the types of research undertaken across various discipline areas, each panel will have some flexibility in the criteria they use to assess the level of quality associated with each HEI.

The REF is guided by three main underlying principles:

- Equity: All types of research and all forms of research output across all disciplines shall be assessed on a fair and equal basis.
- Equality: HEIs are strongly encouraged to submit the work of all their excellent researchers, which may reduce the number submitted for individuals who were constrained by circumstances during the assessment period (2008–2013). Panels will assess work on an equal basis without any penalty for reducing the number of submitted outputs.

- Transparency: The criteria and procedures that will be applied in the assessment will be published in full, well in advance of institutions making their submissions. The outcomes will be published in full and decision-making processes will be explained openly. (REF 2011: 6)

In October 2013, HEIs will be formally invited eligible HEIs to make submissions, due on the 29 November 2013. The submission will contain:

- (a) Information on staff in post on the census date, 31 October 2013, selected by the institution to be included in the submission.
- (b) Details of publications and other forms of assessable output that selected staff has produced during the publication period (1 January 2008–31 December 2013).
- (c) A completed template describing the submitted unit's approach during the assessment period (1 January 2008–31 July 2013) to enabling impact from its research, and case studies describing specific examples of impacts achieved during the assessment period, underpinned by excellent research in the period 1 January 1993–31 December 2013.
- (d) Data about research doctoral degrees awarded and research income related to the period 1 August 2008–31 July 2013.
- (e) A completed template describing the research environment, related to the period 1 January 2008–31 July 2013. (Adapted from REF 2011: 7)

Institutions will be judged on:

- (a) Outputs (assessment weighting 65%): The quality of submitted research outputs will be judged in terms of their 'originality, significance and rigour', with reference to international research quality standards.
- (b) Impact (assessment weighting 20%): The 'reach and significance' of the impacts on the economy, society and/or culture that were underpinned by excellent research will be assessed, as well as the approach to enabling impact from its research.
- (c) Environment (assessment weighting 15%): The institution's research environment in terms of its 'vitality and sustainability', including its contribution to the vitality and sustainability of the wider discipline or research base will be assessed. (Adapted from REF 2011: 8–9)

The REF documentation considers impact in the following way:

For the purposes of the REF, impact is defined as an effect on, change or benefit to the economy, society, culture, public policy or services, health and the environment or quality of life, beyond academia ... [It] includes, but is not limited to, an effect on, change or benefit to:

- the activity, attitude, awareness, behaviour, capacity, opportunity, performance, policy, practice and process or understanding
- of an audience, beneficiary, community, constituency and organization or individuals
- in any geographic location whether locally, regionally and nationally or internationally.

For the purposes of the impact element of the REF:

- Impacts on research or the advancement of academic knowledge within the higher education sector (whether in the UK or internationally) are excluded.
- Impacts on students, teaching or other activities within the submitting HEI are

excluded.

- Other impacts within the higher education sector, including on teaching or students, are included where they extend significantly beyond the submitting HEI.

Impact extends beyond the institution and it involves changes in the ways in which people think about or do things. It is upon consideration of the first two items above that the rest of this article will turn.

The Impact of the REF on research

The way in which the REF is structured and undertaken has a huge impact on the research undertaken by universities and academics. As Bekhadradnia (2009: 1) argued: 'Something like £10 billion will be allocated as a result of the evaluations that are conducted, and this money is allocated highly selectively according to quality; moreover, the nature of the assessment process inevitably has a strong impact on the behaviour of universities, the academics who work in them and the nature of the research that they do'. Each piece of work submitted is judged to be at one of five levels identified below in Table 1.

Table 1. Criteria and definitions of starred levels

The criteria for assessing the quality of outputs are 'originality, significance and rigour'

- Four star Quality that is world leading in terms of originality, significance and rigour
- Three star Quality that is internationally excellent in terms of originality, significance and rigour but which falls short of the highest standards of excellence
- Two star Quality that is recognized internationally in terms of originality, significance and rigour
- One star Quality that is recognized nationally in terms of originality, significance and rigour
- Unclassified Quality that falls below the standard of nationally recognized work. Or work which does not meet the published definition of research for the purposes of this assessment

The issue of quality

The major impact of exercises such as the REF is not national, but local, and is associated with the term 'quality' and how universities interpret this along the way. For some, the number of citations associated with a published piece of research represents a quality perspective. For others, the reputation of the publication outlet, usually journals, is a key indicator.

In the first instance, that of citation numbers, some disciplines of research are more likely to generate citations than others. To make judgements based solely on the number of citations, without taking into account the potential number that might be expected, is to be unfair. For instance the number of citations that a single academic attracts in a major medical field such as cancer research should not be compared with those of an academic who specializes in an area such as a specific but now extinct language, since the former has many people in the field and many outlets for research publications, whereas the latter might only have a handful of researchers

worldwide and a single journal in the field. Yet, many areas of the forthcoming REF will choose to use these types of metrics to determine quality.

In the second instance, the education panel of the forthcoming REF, as with the education panel from the previous RAE, has declared that pieces submitted to it will not be judged based on which journals they had appeared in, but on the level of quality identified by the panel through its assessment processes. However, such determinations at the national level are not always heeded at the local level. Comments such as those by Newby (Select Committee on Education and Skills 2003: question 445) make it hard for universities to see this exercise as developmental, in any sense of the term. He argued that:

... there is a logic to, as I say, starting with the most excellent research, making sure that is properly funded and working our way down. So, we have started with the 5*. We have moved onto the 5s and I remind the Committee that we are now fully funding the 5s on a 2001 basis. Then we come to the Level 4s. Are we going to say that we are going to switch money out of truly world class excellent research into research which is less than truly world class excellent? I really think that is not a good way to go.

University administrators making decisions based on these criteria, leads to situations such as that of the new Vice Chancellor of the University of Malaya, who gave existing professors a three-year period to publish three publications in ISI recognized journals or face a change of title from professor to senior researcher or of the University of Glasgow library circulating a table that identified the numbers of articles published by different university departments in journals identified as being at various levels of quality and exclusivity (from 1-star to 4+ stars) that had been drawn up by an Australian research authority, even after the Australians themselves had discredited the list.

Universities around the country are currently undergoing the process of determining who within their staff will be put forward, and what the minimum criteria for this will be. Since it has been made fairly clear on the national level that the amount of funding tied to publications that are judged by the panels as being '2-star' (which is 'Quality that is recognised internationally in terms of originality, significance and rigour') is minimal, staff who are 'REFrable' may have to have an average of 2.5 or even higher to get submitted by the university in the first place. To get this, staff will have to submit at the minimum, two '2-star' publications and two '3-star' publications (Quality that is internationally excellent in terms of originality, significance and rigour but which falls short of the highest standards of excellence).

If 3-star publications are identified as being 'internationally excellent' but that this still 'falls short of the highest standards of excellence' one wonders what the highest standards of excellence might be. It suggests that being internationally excellent is no longer enough. The concept of 'world-leading' suggests that there are very few pieces that might be considered within this category. As Paul argues (2008: 325) 'What would a world leading edge piece of research consist of? Is it Einsteinium in its impact? If so, there is not a lot of that about. Does it change research directions, or is

it heavily referred to by all and sundry (but a good review paper will achieve this), or what?’

If research funding is to be determined mainly on ‘world-leading’ research, then the chances of junior researchers being included in the REF becomes small. In the longer term, it may also shape the type of academics that are appointed and the disciplines that are supported. But there is also a real concern about the validity of the exercise in the first place. McNay (2002), in reporting on the 2001 RAE exercise, indicated that how the exercise was undertaken and was left in the hands of individual panels, and that these varied in their interpretations of how the exercise should be conducted. He noted how: Panels differed in: the wording in relation to criteria for national excellence; the elements included in criteria; how they constructed a final grade; use of judgements on output; and the account taken of the balance of quality across the four outputs. (McNay 2002: 50)

This report provided one example of how panels interpreted and used the concept of ‘international excellence’. For economics and econometrics (HEFCE 1999): International excellence will be interpreted within a wide international context. This will be defined by reference to high quality research activity, where it has been identified, around the world. That is tautological and almost meaningless in its vacuity. Philosophy was somewhat better, but very demanding: International excellence will be defined as ‘work which is or ought to be a primary point of reference in its field, i.e. a contribution of whose general theme every serious worker in the field is or ought to be aware’. That could be a very short list, yet a higher proportion of units got a 5* grade in philosophy than in economics and econometrics. (McNay 2002: 49)

The actual number of items to be considered by the panel in order to make their judgement also varied widely, from panel to panel. Criteria ranged from ‘will collectively examine in detail at least one cited item per member of staff’ (accounting and finance) to ‘at least a minimum of 5% from each submission, but the Panel will examine as much as necessary to make an informed judgement’ (education), through ‘will examine in detail a minimum of 10% of the items of research output’ (business and management studies) and ‘will examine in detail at least 25%’ (sociology), to ‘collectively will examine in detail an absolute minimum of 50% of the output cited for every member of staff submitted’ (history) and ‘Virtually all works cited will be examined in detail’ (theology, divinity and religious studies) (McNay 2002: 50).

It seems almost as if it is the case of the larger the discipline area, the lower the level of examination that takes place.

Further differences emerge when one considers the mandate given to the panels in how their judgements are made:

Economics and econometrics looked at the ‘overall quality of nominated published outputs’. The text from the Education panel seemed to give a weighting to context as much as to content:

... the Education Panel will reach its judgements in ways that recognise the department’s overall research performance, strategy and culture. It will take account of the relative size of

research groups and the coherence of the elements within their individual activities, in deciding how each will influence the department's individual grade. (McNay 2002: 51)

Given this, it seems that some research areas and indeed some institutions seem to hold a privileged position when it comes to the judgements being made.

Another concern is the issue of applied or practical research, especially in education (Furlong and Oancea 2005). The concern about applied research could well have emanated from a comment from Sir Howard Newby when he expressed concern about the quality of educational research in general.

It is not as good as it might be and I speak as a former Chairman of the Economic and Social Research Council (Select Committee on Education and Skills 2003: question 448). He went on to say that... there are quite substantial resources available for education research in the normal way, it is simply that, on the indicators I mentioned, there is a large quantity of researchers, staff and graduate students engaged in education research. We do have concerns about the overall quality (Select Committee on Education and Skills 2003: question 449). Research in education is considered inferior to research in some other discipline areas and has been for some time. A number of critiques of educational research in the past have expressed a concern about the level of relevance (Hargreaves 1996, Hillage et al. 1998, Tooley 1998). The relationship between pure research and applied or practice-based research is an area that needs further exploration.

Furlong and Oancea (2005: 4–5) argue that although 'The 2001 RAE stated that it supported the "equal consideration" of basic, strategic and applied research and the consideration of practical or practice-oriented research as a matter of principle', that subsequently two government reports 'implied that this in fact had not happened'. However, in recent years the need for "evidence informed" policy and practice has rapidly gained support and is coming to shape research agendas throughout the country' (Furlong and Oancea 2005: 6). This has been accompanied by concerns on both sides. From the policy and practice side, there are concerns about the 'lack of accountability of researchers' but from the researchers there is a concern that this is the start of 'a "marketisation" of the research process' (Furlong and Oancea 2005: 6) - a move away from pure research and theoretical knowledge to the more pragmatic concern about what works in terms of new policies and new practices.

Furlong and Oancea argue for a consideration of what quality applied or practical research might look like while recognizing that 'defining applied and practice-based research with any precision is difficult because there are many competing views with only some areas of overlap between them' (2005: 7). They propose for a multi-dimension model, with four dimensions - quality, epistemic, technological, capacity building and value for people and the economic, as a means of assessing quality in applied research. Within each are a series of sub-dimensions, as in Figure 1.

As seen in this figure, it is suggested that the impact that research has should be part of its quality assessment. This has been taken up in the REF 2013, where impact, not only of applied research, but of pure research as well, will generate 20% of a

university's total score.

Figure 1. The dimensions and sub-dimensions of applied and practice-based research quality (from Furlong and Oancea 2005: 15)

Quality sub-dimensions	Dimensions of quality			
	<i>Epistemic: methodological and theoretical robustness</i>	<i>Technological</i>	<i>Capacity development and value for people</i>	<i>Economic</i>
	Trustworthiness	Purposivity	Plausibility	Marketability and competitiveness
	Builds on what is known + contribution to knowledge	Salience/ timeliness	Partnership, collaboration and engagement	Cost-effectiveness
	Explicitness	Specificity and accessibility	Reflexivity, deliberation and criticism	Auditability
	Propriety	Concern for enabling impact	Receptiveness	Feasibility
	Paradigm-dependent criteria	Flexibility and operationalisability	Transformation and personal growth	Originality
	<i>Scientific robustness</i>		<i>Social & economic robustness</i>	

The use of Impact as a measure of excellence

Perhaps the most significant change in the current research exercise is the introduction of the notion of impact. Wilson et al. (2010: 1) suggest that there is 'renewed interest and emphasis on the gap between research and policy and practice, both nationally in the UK and internationally'. The suggestion that research should be something that helps to define policy and practice is not new, but the current assessment of impact goes further in that it expressly excludes the use of academic impact as an acceptable form. Bekhadradnia (2009: 1) reports:

The biggest innovation in the proposals for the future is the introduction of 'impact' as an explicit and separate element in the grading system. Whereas previously panels were able in their assessments of overall quality to take into account the impact that the research that was being evaluated had had, they are now required to assess this separately, and the funding bodies' initial proposal is that 25 per cent of the final score that will be used for determining funding will be based on the impact score. Impact is here defined as impact on the economy and on society more generally, but explicitly excludes academic impact.

It is the last sentence that becomes critical in thinking about what research might be seen as taking priority and how this might itself have a tremendous impact, especially on those involved in educational leadership.

It seems that dissemination of research needs to not only move beyond academic circles in order for it to be considered first rate, but it is only research that has societal or economic impact that is first rate. This notion will constrain both universities and

academics to certain kinds of research in order to receive the highest ranking, which in turn will receive the highest funding. Junior researchers, who may not have had time to create an impact, and more senior colleagues who have focused on research that may not have policy or social implications but might help to change our understanding of an issue or concept, might now be excluded from the exercise because their work does not address the right areas to create impact. This is articulately explained by Bekhradnia (2009: 6) when he suggests that:

As they stand at present, the historian that does good but not extraordinary work on Henry VIII could be valued more highly, if he presents a television series on the subject, than his colleague who spends 17 years producing a book that changes the way that historians see their subject, but which does not have wider impact beyond the academic community.

However, some research is more amenable to the types of impact currently being proposed than others. Boaden and Cilliers (2001:6) note how:

Scientific research, for example, has far more potential than social science research for its outputs to be developed in saleable 'products', often through 'spin-off' companies linked to universities. Social science research results in less tangible 'products' although it may spawn a substantial service industry in terms of consultancy to support the implementation of the findings for individuals and organisations.

When impact factors play such a role in shaping what is worthy of research, then we are moving towards a time where only those areas that can be deemed as contributing to 'the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia' (REF 2011: 40) are those that universities might support, because those are the research activities likely to be funded. It could be argued that such a move is similar to judging the quality of a school by referring only to literacy and numeracy tests. Like schools, there are many other reasons for universities to exist, even within the dimension of research activity.

Educational leadership

Educational leadership research is considered for the REF by the education panel. So many of the concerns expressed by Newby above about educational research in general might also be directed at research into educational leadership. In the special issue of *Educational Management and Leadership* (Vol. 33, 2005) a range of concerns were raised about the state-of-the-art in educational leadership research. Gorard (2005: 156) reports that 'Education had one of the weakest RAE results of any subject in 2001, comparable in quality only to nursing, environmental sciences, performing arts and media studies (Tysome 2004)' and that 'too much evidence comes from poorly designed or advocacy driven studies'. Muijs (2011: 116) reports on his analysis of over 500 articles from six international leadership journals that 'it is clear that the predominant modes of research in the field are case study and survey research methods. There was no significant change in the proportion of these types of research over time'. This leads to a perception that there is 'a deeply worrying lack of quantitative skills' (Commission on the Social Sciences 2003: 8) and perhaps is the reason why there is a tendency in the USA 'to deny federal funding to many traditional

researchers in schools of education who do not have the skills needed, and to increase funding to private consultancy companies who have a staff able to conduct several RCTs in parallel' (Gorard 2005: 156).

Further, both Gorard and Muijs identify the leaning that researchers in the educational leadership field have towards 'Position Papers' or 'think-pieces', which are 'articles that are neither based on empirical research nor systematic literature reviews, but contain position statements on factors such as ethical leadership or introduce new leadership concepts. A lot of these take the form of "tips for leaders"' (Muijs 2011: 116). Both authors refer to 'the uniformity of the methods used, mainly small-scale, qualitative work, with little transparency and no comparison groups' (Gorard 2005: 158). This leads to the conclusion that 'Research in educational leadership and management, therefore, has a relatively weak quality profile within the already relatively weak quality profile of educational research' (Gorard 2005: 159) or worse, 'that educational research is a fashion not a science' (Foskett et al. 2005: 246).

One difficulty facing researchers in educational leadership is the multiple audiences that might use the research. These include practitioners, policy-makers and academics. However Galton (2000) found that 'of the 523 research areas of influence listed, only 39 concerned ELM, despite the fact that the majority of the 302 respondents, 84 percent, were heads 5 or deputy headteachers and, therefore, one might assume, leaders' and that 'the UK and the US governments for example have implied a commitment to evidence-based policy but complained that this intention is frustrated by the unsuitability of available research (Blunkett, 2000; US Department of Education, 2002)' (Foskett et al. 2005: 247). It seems then that the educational leadership and the management research might not be reaching two of its three most important constituents, which leaves, well, academics talking to academics. When one considers the REF's future commitment to impact beyond the academic world, it suggests that there is quite a bit of work to do.

The way forward

There are a number of things that might need to be considered as a mechanism for moving educational leadership research towards the REF gold standard of 4-star research, and perhaps becoming acceptable outside our own academic circles. The first of these is to change from looking inward to looking outward. Gorad (2005: 162) argues that educational leadership research is 'at present, generally inward looking, researching the impact of changes on management rather than testing the impact of management on anything else. Put bluntly, we have to demonstrate convincingly what practical difference educational management and leadership makes to anything but itself'. But having done this, we need to develop rigorous, theory-driven research to enable us to build a better understanding of 'what works', because however, we view that term, the nature of educational research in general, in the current economic and political climate will only survive, if we answer some of the questions being asked. As Foskett et al. (2005: 249) argue: 'The autonomy of researchers to identify their own priorities has been displaced by an increasing prioritization on funded research in all accountability measures, accompanied by a narrowing of the range of potential funders and their almost unanimous emphasis on the contribution of research to practice'.

We need to put more work into methodology. Muijs (2011: 229) calls for more longitudinal research, both quantitative and qualitative, although sometimes quality measure such as the REF mitigate against such longitudinal work, as it cannot be completed and published within the five-year cycle. He also calls for 'more studies where leadership is integrated within a model of school effectiveness, which is theorised and takes into account the ways in which leadership interacts with other key school factors'.

There is the need to clarify concepts and to have a better measurement of these to lessen the level of unreliability currently existing in many areas of study. Foskett et al. (2005) identify four priorities for educational leadership research: Interdisciplinarity, an increase of quantitative and evidence-based studies, the use of a range of methodologies rather than a few, and an increase in the size of the populations used.

Educational leadership is seen as a weak link in the already weak link of education research, when it comes to the REF. In this article, I have shown how preferences are given to certain disciplines over others, certain institutions over other and certain publication outlets over others. I have not touched upon the perceptual differences held when one considers books or book chapters, as opposed to journal articles. The REF exercise is designed to give certain groups further opportunities in the funding game at the expense of those who would like to learn how to play. The rules are very specific, but the concepts contained within them are fuzzy.

There are certain players who are excluded from their own team because they do not fit the profile required and others because they have not yet had sufficient experience to be able to compete. The game might be conducted using numbers and quality coefficients and the owners of the game might indicate that the rules are fair for all. Yet, in the case of the UK, we know which four institutions will get to be recognized as the 'top 4' (same as last time), we know which 30 or so institutions will pick up the vast bulk of the available funds and we know that the game is rigged. But we keep playing it because it is the only game in town.

References

- Bekhradnia, B. (2009) Proposals for the research excellence framework - a critique. Report to HEPI conference on Research Funding and Assessment: the Future, held at the Royal Society on 14 October 2009.
- Bence, V. and Oppenheim, C. (2004) The role of academic journal publications in the UK Research Assessment Exercise. *Learned Publishing*, 17, 53–68.
- Boaden, R. and Cilliers, J. (2001) Quality and the research assessment exercise: just one aspect of performance?. *Quality Assurance in Education*, 9(1), 5–13.
- Commission on the Social Sciences. (2003) *Great Expectations: The Social Sciences in Britain* (London: Academy of Learned Societies for the Social Sciences).
- Curran, P. J. (2000) Competition in UK higher education: competitive advantage in the research assessment exercise and Porter's diamond model. *Higher Education Quarterly*, 54(4), 386–410.

Foskett, N., Lumby, J. and Fidler, B. (2005) Evolution or Extinction? : Reflections on the Future of Research in Educational Leadership and Management. *Educational Management Administration & Leadership* 2005 33: 245-253.. DOI: 10.1177/1741143205051056

Fry, M. (1948) The University Grants Committee, an experiment in administration. *Universities Quarterly*, 2, 221–230.

Furlong, A. and Oancea, A. (2005) Assessing Quality in Applied and Practice-Based Educational Research: A Framework for Discussion (Oxford: Department of Educational Studies Oxford University).

Galton, M. (2000) *Integrating Theory and Practice: Teachers' Perspectives on Educational Research*. Accessed online September 22, 2011, <http://www.tlrp.org/pub/acadpub/galton2000pdf>

Gorard, S. (2005) Current Contexts for Research in Educational Leadership and Management *Educational Management Administration & Leadership* 2005 33: 155-164. DOI: 10.1177/1741143205051050

Hargreaves, D. (1996) Teaching as a Research-based Profession: possibilities and prospects. *British Educational Research Journal* 23, 141–161.

HEFCE, Higher Education Funding Council for England (1999), *Research Assessment Exercise 2001: Assessment panels' criteria and working methods* (HEFCE, Bristol).

Henkel, M. (1999) The modernisation of research evaluation: The case of the UK. *Higher Education*, 38: 105–122.

Hillage, J., Pearson, R., Anderson, A. and Tamkin, P. (1998) Excellence in research in schools. Research report 74 (London: Department for Education and Employment).

Johnston, R. J. (1995) Geographical research, geography and geographers in the changing British university system. *Progress in Human Geography*, 19, 355–371.

Mace, J. (2000) The RAE and university efficiency. *Higher Education Review*, 32(2), 17–35.

McNay, I. (2003) Assessing the assessment: an analysis of the UK Research Assessment Exercise, 2001, and its outcomes, with special reference to research in education. *Science and Public Policy*, 30(1), 47–54.

Muijs, D. (2011) Researching Leadership: Towards a New Paradigm in Townsend, T. and MacBeath, J. (Eds) *The International Handbook for Leadership for Learning*, Springer, New York. 115-125

Paul, R. (2008) Measuring research quality: the United Kingdom Government's Research Assessment Exercise. *European Journal of Information Systems*, 17, 324–329.

Piercy, N. (2000), 'Why it is fundamentally stupid for a business school to improve its RAE score'. *European Journal of Marketing*, 34 (1/2): 27–36.

REF (2011), *Assessment framework and guidance on submissions*. Available online at <http://www.ref.ac.uk/pubs/2011-02/> (accessed online September 29, 2011)

Select Committee on Education and Skills (2003) Examination of Witness (Questions 440–451). Available online at: <http://www.publications.parliament.uk/pa/cm200203/cmselect/cmeduski/425/3030504.htm> (accessed 27 September 2011).

Sugden, D. (1997) Reflections on the research assessment exercise. *Area*, 29, 367–368. 5

Tooley, J. (1998) Educational Research - an OFSTED Critique (London: OFSTED).

Tysome, T. (2004) Next RAE may Demote 5*s. Times Higher Educational Supplement, 2 April, pp. 7.

Williams, G. (1992) An evaluation of new funding mechanisms in British higher education: some micro-economic and institutional management issues. Higher Education in Europe, 17, 65–85.

Wilson, P., Petticrew, M., Calnan, M. and Nazareth, I. (2010) Does dissemination extend beyond publication: a survey of a cross section of public funded research in the UK. Implementation Science, 5, 61. Available online at <http://www.implementationscience.com/content/5/1/61> (accessed 2 September 2011).