Speculation that a Poppleton sociology don had secured a top mark in the new measurement of research impact has been dashed by a spokesperson for the Economic and Social Research Council ... [who] insisted that Doctor Catchpole had created what he called ‘the wrong kind of impact’. ‘In general’, he went on, ‘the ESRC measures impact by the extent to which research findings agree with assumptions which have already been made by policy makers. In this respect, Doctor Catchpole has failed to make any impact whatsoever’.

(Laurie Taylor, THE, 30 July 2009)

However HEFCE deals with the responses to its recent REF Consultation, much of them critical of its proposals on impact, the assessment of impact is clearly going to be a major part of any future compact between academic researchers and their paymasters. However, there remain important unanswered questions about what is meant by impact and how it might be measured or even described in the field of Education. The focus of this article is the assessment of impact in the light of the forthcoming Research Excellence Framework (REF) although the importance of demonstrating impact (and hence the issues raised in this article) is relevant in a range of contexts.

Background
Since the Warry Report (Research Council Economic Impact Group, 2006), there has been increased pressure on public bodies to demonstrate the impact of the research they undertake or commission. As an ESRC report notes ‘It is no longer assumed, as it may have been in the past, that research expenditure will eventually and on its own dynamics lead to social and economic benefits’ (Molas-Gallart and Tang, 2007, p. ii).

In the past, the ranking and funding outcomes of the various Research Assessment Exercises (RAEs) depended on judgements of quality made largely by senior academics who tended to emphasise the need to contribute to ‘pure’, theoretical knowledge rather than policy makers’ concerns about ‘what works’ (Furlong and Oancea, 2005). In this respect, RAEs often acted as a disincentive to conduct applied research and appeared relatively unsupportive of practitioner inquiry and collaboration between academics and teachers.

Well before current proposals for the REF (HEFCE, 2009a), the situation was changing. Thus, in line with a drive for increased relevance in research, a 1998 review of the RAE recommended that research users be represented on the peer
review panels for the next exercise in 2001. However, even where the presence of research users on panels was subsequently strong, as in Education, their own impact on judgements of research quality seems to have been limited.

Two further reports followed (the Roberts Report and the Lambert Report), this time criticising the outcomes of the 2001 RAE on the grounds that it had failed to recognise the importance of knowledge transfer between universities and industry or to encourage innovative forms of collaborative university-led research (UK Funding Bodies, 2003; HM Treasury, DfES and DTI, 2003). Reflecting this, the changes made for the 2008 RAE represented a much stronger push in the direction of rewarding research on the basis of relevance and impact. Although peer review was retained alongside the representation of research users, there were explicit criteria in each subject to enable the assessment of applied, practice-based and interdisciplinary research. The Education sub-panel criteria included clear reference to the provision of evidence for, and impact on, policy and practice and a recognition of applied and practice-based research which was of relevance to the needs of the public and voluntary sectors and commerce and industry:

20. The sub-panel will take an interest, as appropriate for the area of research, in the impact and potential impact of an output on policy or practice, as well as in the academic environment, as part of its consideration of the significance of research.

(HEFCE, 2006, p. 31)

Whether or not this criterion was fully reflected in the outcomes of RAE 2008, the direction of travel was well under way before the current controversy over the REF.

At the launch of the REF, it was clear that the government was determined that its ambitions should no longer be subverted by what it perceived to be the conservatism and prejudices of academics (Denham, 2008), though there was little by way of consensus about how this was to be done and many elements were left for later consideration (Eastwood, 2008). The issues seem now to have been largely resolved. Metrics have been downgraded in all disciplines and look likely to have little, if any, role in the assessment of research quality in the arts, humanities and social sciences.

**The Research Excellence Framework (REF)**

The main assessment mechanism to be used in the REF (as currently planned) is that of expert review. The weightings which are proposed for the three components of research outputs, impact and research environment are 60 per cent, 25 per cent and 15 per cent respectively, for all disciplines.

The proposals for the assessment of quality of outputs still use the criteria of ‘originality, rigour and significance’ as in RAE2008. It is worth highlighting that ‘significance’ is defined as ‘the capacity to make a difference either through
intellectual influence within the academic sphere, or through actual or potential use beyond the academic sphere’ and so research which ‘makes an impact’ will rate highly here as well as under the impact assessment.

HEFCE proposes that impact be assessed at the unit rather than the individual level. It is also deemed important that the impact is based on high quality research. It is anticipated that impact will be assessed through expert panels which will include research users. The criteria used will be ‘reach’ (breadth) and ‘significance’ (depth). Acknowledging the time lag between a research output and some forms of impact, it is proposed that the impact assessment will be made on research which has had impact during the current assessment period but may have been undertaken up to 10 or 15 years previously. The evidence submitted to the sub-panel will be in the form of an impact statement for each submitted unit and a number of cases studies. HEFCE are currently running a pilot exercise (see HEFCE Circular 19/2009) to trial the proposals for assessing impact; this is due to conclude in summer 2010.

The third aspect of the REF is the contribution of the research environment in terms of supporting high quality research and dissemination of such research. The specific areas intended to be covered are resourcing (research income, staffing, infrastructure and facilities), management (strategic planning, staff development, training for postgraduate research students) and engagement (support and structures for interaction between research and users/public, contributions to the knowledge base and interdisciplinary work) (HEFCE, 2009a).

The quality of research (as judged through outputs) remains the most important element as it is intended to make up 60 per cent of the overall assessment. In fact, the final proportion which will be due to research quality is likely to be higher as excellent impact requires excellent research. Impact is described as building ‘on excellent research to deliver demonstrable benefits to the economy, society, public policy, culture or quality of life’ (HEFCE, 2009a, p. 2).

Some key considerations and concerns
In principle, there is much to welcome in the proposal that impact should be rewarded and encouraged in the field of Education. After all, while part of the audience for academic research is the Academy itself, e.g. in improving teaching and directing further research, other stakeholders expect that a reasonable proportion of QR money will be used in a way that ends up being of direct benefit to users. Indeed, in Education, we suspect that few researchers would themselves disagree with Black and William (2003) that ‘the majority of research in education should be undertaken with a view to improving educational provision’ (p. 632), as long as this manifests itself in a broad interest in improving provision, not just work which is of direct relevance to government agendas, which is what some critics of current proposals fear. Objecting to the dominance of one particularly powerful end-user – government or even just the present government – does not require that we reject the role of external stakeholders altogether. Sometimes it is a technocratic
obsession with ‘what works’ that worries researchers rather than the notion that research should have some impact outside the Academy.

However, we would argue that there are important reasons why at least some research, even in Education, should be relatively uninterested in considerations of impact, not least because it is impossible to state with any certainty which research will make an impact in the future. Furthermore, the Education research community conducts a range of valuable work which it would make little sense to judge on its contribution to improving provision or raising standards – much work in the history of education being just one example. In such areas, the concerns of scholars in other fields are also applicable to our own (see for example BBC, 2009; Moriarty, 2009).

It is worth noting that impact was assessed in RAE 2008 through both the ‘quality’ element (in the ‘significance’ criterion) and the ‘research environment’ element (which used indicators for income from research users, influence on public policy advice and exploitation of new ideas and products). It does not, of course, follow that high quality work leads to impact (even when considering applied research) and research which has high impact is not necessarily of high quality. A welcome outcome of the REF would be to incentivise a stronger correlation between quality and impact for applied research. The challenge is whether there is a way of defining what is meant by impact and, crucially in terms of the REF, whether this can be validly assessed. Given that this remains an open question, a weighting of 25 per cent for impact in the first REF seems excessive.

Some international approaches

The issues we are facing are not just national ones and that there may be some lessons to learn from others’ experiences and approaches.

Before the current Australian Labor government came to power in 2007, the coalition government there had decided to introduce a Research Quality Framework (RQF) to assess research quality and research impact (Department of Education, Science and Training, 2006). The process designed for assessing impact had similarities to the proposed REF in the UK; for example, assessment was to take place through an expert panel via their judgment of submissions including an impact statement and case studies. It is also worth noting that the plans allowed research groups to apply to opt out of impact assessment where, for example, their research was not intended to have any impacts beyond enhancing their discipline, or where the research was too new to have had any determinable impact (Department of Education, Science and Training, 2007). However, Watson (2008) argues that the approach to measuring impact would have favoured instrumental research in education over more theoretical and critical work.

One month after coming into office the Labor government abandoned the RQF (Carr, 2007a). Senator Kim Carr (the Minister for Innovation, Industry, Science and Research) said that ‘The RQF is poorly designed, administratively expensive and relies on an “impact” measure that is unverifiable and ill-defined’ (Carr, 2007a).
Whilst it seems that the Rudd government did (and does) consider impact to be important, it felt that the methods of assessing it that had been developed for the RQF were flawed. In a speech before the election, Carr summarised Labor’s opinion of the RQF:

[The impact] measure seems to be a proxy for industry collaboration or community engagement. In our view, it should be funded explicitly. If applied, it will also have the effect of imposing the language and processes of science on the humanities. The humanities, arts and social sciences do not need to adopt the language of science to legitimise its participation in the innovation system.

(Carr, 2007b)

However, having celebrated the abandonment of that exercise by Kevin Rudd’s Labor government, researchers in Australia are certainly not going to be left alone. Just like the UK’s forthcoming REF, Australia’s new Excellence in Research for Australia (ERA) initiative (first taking place in 2010) will assess research quality using a combination of metrics and expert review. As in the UK, this exercise will inform the government’s funding compacts with universities (Carr, 2008).

There are other approaches being explored in other countries in terms of assessing impact; the most notable of these are the systems used in New Zealand and the Netherlands (CHASS, 2005). In New Zealand, for example, indicators such as information on policy advice given and measures of outputs produced specifically for users are taken into consideration.

**Recognising and defining impact**

As part of the preparation work for the RQF, the Department for Education, Science and Training in Australia commissioned the Council for the Humanities, Arts and Social Sciences (CHASS) to explore different measure of quality and impact for publicly funded research in the humanities, arts and social sciences (HASS) sector (CHASS, 2005). The research team explored different approaches taken to assessing impact (and quality) by different publicly funded research agencies (these included, for example, the Australian Institute of Marine Science) and research funding agencies (the Australian Research Council, the National Health and Medical Research Council) often working in science-related areas. They found that there was a much lower level of consistency in approach to assessing impact than quality and that a range of different approaches were used including research commercialisation, technology transfer, non-academic citation of research funded, advice given to government, actual policy influence, measures of international collaboration and community engagement and measures of community awareness of research (including changes in attitudes due to the research).

In the UK, a report commissioned by the British Academy found that HASS research was often undervalued by government and the wider public. The report aimed to
‘show that the impacts of the HSS disciplines can in future be better tracked and recognized. This in turn is the first and most essential step to ensuring that the significance of HASS research is appropriately recognized for funding support and better valued in society at large’ (LSE Public Policy Group, 2008, p. 5). The research team found that researchers were often unaware of opportunities for enabling impact through engaging with policy makers and the policy making process. Researchers felt that opportunities were being missed and that there was considerable potential for increased contribution (Wilson, 2008). One aim of the REF should be to incentivise this activity which should enable an increased appreciation of the contribution of social science research and allow academics to invest time in impact-related activities. It should also be noted that increasing impact is a two-way process; policy makers must also increase engagement with social science academics (LSE Public Policy Group, 2008).

In the UK, of course, Research Councils are under increasing pressure to assess and monitor the societal and economic impact of the research they fund. The ESRC now monitors the impact of the work it funds through building impact into all final evaluations. The ESRC takes a broad view of impact (local and global; scientific, societal and economic etc.) and is keen to stress that it is not important for all research to have impact beyond the academic community as there are many areas and types of worthwhile research were this is not appropriate. The ESRC is currently experimenting with different approaches to assessing impact which employ both quantitative and qualitative techniques.

As part of the government’s HE review, a report by Prof Paul Wellings (Wellings, 2008) recommended that HEIs produce an annual statement describing their contribution to the economy and society in order to demonstrate these public benefits to government, research users and the wider public. HEFCE are currently running a pilot which is aiming to improve how HEIs can demonstrate where they add public value, a wider concept than research impact (HEFCE Circular 16/2009).

Potential problems with the current REF proposals

Relative importance of economic impact

Although the current HEFCE consultation states, encouragingly, that ‘There should be a wide definition of impacts, including economic, social, public policy, cultural and quality of life’, one notes the privileging of ‘economic’ (first author status) and one wonders what the effective relative weighting of these various understanding of impact will be when the chips are down.

1 See for example the new grading scale for ESRC projects – ‘the new grading scale allows ESRC researchers to achieve the highest evaluation grade through either outstanding academic or practical impact, or a combination of both’ (Sooben, 2009, slide 8).
**Validity of impact assessments**

It seems likely, particularly the first time that it is used, that there will be less validity in the assessment of impact than in the assessment of research environment. It is therefore somewhat surprising that it is proposed that the REF component most informed by objective measurements (research income, PGR numbers and completions, etc.) should contribute only 15 per cent to the overall REF quality profile, whereas the as yet untested impact assessment should contribute 25 per cent. There is a danger that the case studies may become (or, perhaps as seriously, be seen by some to be) exercises in creative writing.

Corroboration is an aspect highlighted by HEFCE in the consultation. The consultations states that the approach ‘includes scope for third party corroboration’ (HEFCE, 2009b, p. 15) in terms of claims made regarding impact in the submission. There is the potential here for personal contacts and relationships to have in some instances inappropriate influence on determining research impact. The actual process which leads to policy impact is often messy and opaque and often the actual influence may not be openly acknowledged. For a number of reasons it is likely, therefore, to be hard to assess the validity of impact claims.

In the REF it is proposed that academic or internal impact will be assessed only in the ‘outputs’ element which assesses research quality. There could be incentives to move away from disciplinary towards more applied research. At the same time, the explicit intended weighting for impact in the REF presumably requires a narrower conception of ‘significance’ than obtained in the 2008 RAE, otherwise double counting will occur.

**What counts as positive impact?**

There is little indication so far as to what might count as impact in Education. There is only one direct example for Education in the draft menu of impact indicators – ‘improved educational attainment among disadvantaged groups’. As the Poppleton quotation at the start of this paper humorously indicates, there are very considerable value judgements as to what constitutes desirable impact. In a field such as Education these are especially apparent. Indeed, a single HEI might, for example, have some colleagues who have worked with the DCSF, QCA and others to improve recorded pupil attainment on KS2 tests and others who work to bring such tests to an end by critiquing their validity or the assumptions of the regime within which they operate.

More generally, some research in Education may be located within an instrumental or technocratic frame of reference, while other research is highly critical of that approach. Although some education research is about ‘what works’, much current work in the field concerns whether we are doing the right things and what constitutes socially just schooling (Gale and Densmore, 2003).

The quality and impact of research in this context cannot be assessed only on its capacity directly to change official policy and practice. Such research may also be seen to have impact where it feeds into public debate and the discursive milieux
within which decision makers at all levels operate. Levin (1998) suggests that research may be used to strengthen the public mind on education and thereby increase ‘resistance’ to superficial but seemingly attractive policies. Acting as public intellectuals may thus come to be seen as a rather more legitimate form of impact for researchers than the questionable ‘quick fixes’ encouraged by a narrow ‘what works’ philosophy (Molnar, 2006).

**Time lags**

As HEFCE and others acknowledge, the problem of time lags between undertaking research and its impacts can be considerable. It is difficult to know how serious an objection this is. Gillies (2008) argues that assessments such as the RAE not only risk missing ‘pink diamonds’ – i.e. research of quite exceptional quality – but actually discriminate against such research. However, the examples he cites in support of his case (the Wittgensteins of this world) are such rare mavericks that one wonders how widespread a problem this is or whether any alternative assessment mechanism would deal with them better.

Whilst the proposals are clear that the research does not have to be undertaken in the assessment period, there are still issues regarding time lags. The consultation states that ‘the assessment of impact will focus on the submitted unit’s contribution to demonstrable economic and social impacts through activity undertaken within the unit during the assessment period’ (HEFCE, 2009b, p. 16). So whilst the research may be undertaken earlier, the dissemination or follow up work building on the research and the corresponding impact both have to happen in the assessment period. This adds an element of risk for academics and HEIs to invest in achieving impact as they would have to see the benefits of effort achieved within the same assessment period. It could also potentially incentivise more instrumental impact which could be more likely to be achieved quickly. It may also mean that investment in impact will be more likely to take place during the beginning of an assessment period. There is always the law of unintended consequences.

**User involvement**

There is an argument that says that academics should not be involved at all in assessment of impact, only users. Interpreted strictly, this would require an entire set of user panels to be established in parallel with academic panels, with consequent implications for cost, quality control, etc.. Furthermore, in view of some of the issues raised above, who these users should be is likely to be more contentious than who the academics should be.

**Attribution**

There is also the possibility for negative impacts on the research undertaken. For example, there may be less incentive to develop a solid knowledge base through corroborating others’ work as the impact trail would most likely consider the initial research project as the underlying research causing the impact.
There is additionally a potential problem with regards to assessing the contribution to any impact from different parties – either different researchers or intermediary organisations and third parties. Paragraph 68 of the consultation document states that:

We do not envisage that a unit could claim credit for impact which was based on research undertaken in the unit but which was exploited or applied through the benefits of others, without a demonstrable contribution by the unit to that exploitation.

(HEFCE, 2009b, p. 16)

It is clear that this provides disincentives to engaging with other organisations in terms of impact as the HEI may not be considered attributable for the impact. This does not appear to cohere with aims of collaboration and enhancing impact as third parties or intermediaries are often effective in terms of having impact, particularly on policy.

**Incentivising particular types of research**

Related also to attribution is the reality that the challenge of assessing causal impact is often harder in the applied social sciences than in the natural sciences. There may be a danger that certain methodologies, for example randomised control trials, will be favoured in impact assessment. Research which has the potential to show impact quickly and obviously (i.e. more instrumental research) seems to be incentivised. There is also the potential problem (which is present in the current publishing tradition) that impact is often easier to establish from positive findings.

**Conclusion**

The above account has pointed out a number of issues that remain to be addressed if impacts are to be identified validly in any discipline and in Education in particular. Nevertheless, funders and the public surely have a right to expect that much research does ‘make a difference’, provided this is not judged on overly narrow criteria. Indeed, our own experience is that many academics in the discipline of Education work in the field largely because they hope their work will have a positive impact for learners in particular and society in general. Those of us who work in Education can be critical of attempts to assess impact without being dismissive of them. But we should also insist, as HEFCE claims it will, that quality is the key consideration. Particularly in a field like Education, we need to beware of the danger of incentivising poor quality research that then has high impact (Gorard, 2008).
References


Levin, B (1998) An epidemic of education policy: (what) can we learn from each other? *Comparative Education* 34 (2) 131-41.


