

YA05031Y

Keynote Address for AARE Focus Conference on 'Quality in Education Research: directions in policy and practice'

Cairns July 2005

Is impact a measure of quality?

Producing quality research and producing quality indicators of research in Australia

Lyn Yates¹

Abstract

Neither impact nor peer review nor 'addressing national research priorities' are unproblematically appropriate criteria for quality assessment in the field of education, but each is a player in current debates about quality research, and each is indicative of some competing agendas that thread through current quality research re-assessments in Australia. This paper discusses some conflicting agendas and trajectories within recent Australian research policy and funding mechanisms for education, as well as the broader context of the status of education research and how this impacts on debates and strategies with regard to quality. It is argued that the education research community do need to develop appropriate quality indicators in the field of education, but to do this effectively requires attention also to the contextual pragmatics and politics of how such assessments will be enacted; and it is also important that the current focus on measures of quality assessment be re-coupled with more attention to contexts of production of education research and the issue of how quality research can be developed.

¹ Address: Professor Lyn Yates, Faculty of Education, LED Room 410 Alice Hoy Building, University of Melbourne, Vic 3010. l.yates@unimelb.edu.au

AARE since its inception has been seriously concerned with a role of promoting quality in education research. However, we are here this year in a specific context, the DEST Research Quality Framework exercise. An initial *RQF Issues paper*² was circulated by DEST in March 2005. This gave the rationale for the development of a new framework for assessing research quality in Australia as the need to provide ‘transparency to taxpayers and governments about results and benefits of public spending on research’, and the desire to adopt practices that would lead institutions such as universities to focus better on quality and relevance of their research, that would encourage positive behaviours on the part of researchers, and that might inform future research funding distribution. Following the release of the *RQF Issues Paper*, a round of consultations was conducted by the Allen Consulting Group, and submissions about the potential form of a Research Assessment Exercise were made by AARE³ as well as some universities and faculties. In May 2005, the Expert Advisory Group set up to advise the government on the form of a Research Quality Framework for Australia produced a second discussion paper that followed this first consultation phase. This second paper, the *RQF Approaches Paper*⁴, has signalled a number of likely directions of an RQF for Australia on which further response is now being gathered. These include that such an exercise will include attention both to quality and to broader impact or use of research, that it will include a peer assessment component in some form, that it is likely to require a mix of quantitative and qualitative measures, that it should allow international benchmarking, and that it should incorporate some acknowledgement of disciplinary differences.

The eventual outcomes of the RQF will impact on all of us in terms of the conditions we work in, and the types of behaviour we will very likely be driven to adopt. It is a development that will impact on individual careers, and it will also impact on how institutions operate and the resources they have, and more generally, on what types of research and writing and related activities begin to get emphasized and de-emphasized in the field of education. So we are here at this conference to talk about quality, educational research, and the RQF.

² DEST, *Research Quality Framework: assessing the quality and impact of research in Australia Issues Paper*, Commonwealth of Australia, March 2005.

³ AARE, *Research Quality Framework: Response to the Issues Paper*, May 2005.
<http://www.aare.edu.au/exec/rqf.pdf>

As *researchers*, quality is something that we should be concerned with. As *education researchers*, impact or relevance or some other descriptor of relationship to the field of education practice is also something we should be concerned with. But what type of discussion are we trying to have here about ‘quality’ and ‘impact’? One way we might take up the agenda is as a normative discussion about worthwhileness, a discussion in which we are voices arguing a case about why a certain criterion or descriptor or assessment process is appropriate. Another way we might take up the current issues, in line with some arguments I made in my recent book (Yates 2004), is to consider ‘quality’, ‘impact’ and the RQF developments as events and texts and empirical objects of study. In the first case, the normative discussions about worth, we can engage in the RQF discussions by elaborating and defending and assessing our own aspirations for the work we do: elaborating criteria of methodological quality, for example; or defending conceptions of appropriate relationships between research and a field of professional practice; or arguing about who or what should be appropriate to judge whether work has in fact been of quality or has had impact. In the second case, we can take issues such as quality and impact as enacted practices, and we can engage in discussions that seek to examine these events, and bodies and textual artefacts in a more pragmatic and empirical way. We can ask what agendas and experiences are being brought to a particular development; what is the situation, context, constraints that particular field of discussion is occupied with; what are the tacit as well as explicit terms of the agenda.

I think at this conference we should be trying to do something of both of these activities (and in this paper I too will be trying to set up some aspects of both lines of discussion). That is, we are here in part to revisit our own normative conceptions of what matters in education research and its conditions for operating; and in part to examine and appreciate the agendas and constraints at work, including what may be possible to influence, and what are some likely outcomes of the present activities, and how we might respond to those both as individuals and as an association. Of course the two discussions, the normative and the contextual assessment, are not unrelated to each other, and I want to begin with some examples that might help us think about ‘quality’, ‘impact’ and situations within education research.

⁴ DEST, *Research Quality Framework: assessing the quality and impact of research in Australia Advanced*

Is impact a measure of quality?

Part 1: how does AARE enact quality?

First, consider AARE itself. When this association began in 1970, its objective was ‘to encourage and improve research in education and the application of its results in education’ (Bessant and Holbrook 1995), in other words, ‘quality and impact’ were on the agenda from the outset. At that time it set out to enact those principles by restricting membership of the association to those who had postgraduate education qualifications and who were recognized researchers, either by a record of research publication, or by holding certain recognized positions or roles. Bessant and Holbrook’s history of the association notes

From the earliest days there was strong support for maintaining a rigid surveillance over the admittance of new members. All applications were circulated among the members of the Membership Committee with instructions to each member to indicate acceptance or refusal.

The authors comment

[The procedures] were designed to restrict the membership to an elite of education research workers after the manner of other research organisations and in tune with the general approach of professional bodies.

(Bessant and Holbrook 1995, p46)

In other words, in its original conception (although it was not a unanimous position of the founding group), AARE took training and qualifications, peer recognition, and restrictive professional association membership as markers of quality, as ways to develop quality within the field, and as ways to signal the legitimacy and quality of education research to outside bodies. (Note that last point: quality is not just what you claim it to be, but what others accept about your claims.) Then as now, there was also some discussion about what form of research should be acknowledged as proper research in education: some members wanted only psychological research to be included, but that was not agreed to, though that type of research did dominate until the later arguments about case-study and ideology in research that led to some very lively debates about what education research looked like at conferences in the late 1970s and early 1980s.

Over time, AARE has moved to a rather different conception of how to produce and legitimate quality. In place of restricting membership and sanctioning who can count as a researcher, it has developed an agenda of inclusiveness. This is a different conception of

promoting quality: one which tries to get all those concerned in any way with research in this field to engage in discussions, journal exchange, training activities about research. AARE's claim to speak on issues of education research is now based more on the size and breadth and comprehensive scope of the membership than the eliteness of the members; and its attention to quality has focused more on processes of support and development to produce better research (research training, research directors meetings, special attention to postgraduate students and new researchers) than on setting hurdles and sanctions for who can be an education researcher and what can count as education research.

Consider, for example, the recent issue about whether the annual conference should move to having fully refereed papers, given DEST policy which allows such papers as publications in the accounting of annual publications, while non-refereed papers are not counted. The policy finally agreed was that refereed papers would be allowed and processed, but as one option among others, an option that did not preclude other people coming with non-refereed papers. This policy was agreed only after considerable debate over a number of executive meetings and after widespread member consultation, and when it was clear that not allowing refereed papers would prevent some people from coming because of their university's policies on conference attendance.⁵ Those speaking against adopting refereeing of papers as policy were concerned not only about whether this would discourage postgraduate students and interested teachers and newer researchers from attending, but argued too that the practical conditions in which refereeing must be done (that is, the sheer numbers of papers involved and the scale of the work required from referees) makes such refereeing an imperfect and unreliable measure of quality and also that it encourages papers which recycle past and finished work, rather than encouraging the conference to be a place where research is presented and debated as it is being developed. Those speaking in favour of refereeing as policy argued that, whatever we as individuals might think, for practical purposes, refereeing is taken by DEST as the enactment of quality and we need to work with that, but they also argued that the refereeing processes do encourage higher quality presentations. In this case, both sides of the argument were concerned with what practices did promote quality papers and quality research, and also with practical consequences of acting in certain ways. The debates about this within the executive went on for about a year and eventually produced the current compromise policy, because both positions have some merit.

In the original conception of AARE, education researchers were a special group who were employed in special places (primarily universities and ACER, and sometimes in education departments). In more recent conceptions there is less boundary maintenance about who is and who is not an education researcher. Annual conferences now set up strategies to encourage more involvement from schools and school teachers and there is more recognition of the very extensive contribution made to education research by people who work in other parts of education, often in the context of undertaking higher degrees. At the same time, there has been more effort to make AARE part of a global network of national associations of education researchers, an acknowledgement both of the international interchange that is seen to be part of building quality research, and of the global sweep of developments such as the quality research agendas and political value of networking and learning from each other in relation to those.

I have begun with that reminder of how AARE itself has taken up different conceptions and enacted practices concerning quality because the different conceptions and aspirations, and losses as well as gains evident in the different conceptions are part of what we need to grapple with here. First, education as a research field and as a social arena is marked by diversity, by different values about what is good. These are both disciplinary-framed perspectives but also differences of broader social political values (is competition and benchmarked international achievement more important than ameliorating disadvantage, for example?). Whatever mechanisms are to be set up to produce or judge quality and impact, how well do they deal with this lack of homogeneity as to what values are good?

Secondly, what counts as quality in research to governments and outsiders is never simply a technical measure. In the debates surrounding the setting up of the Institute of Education Sciences in the US, frequent appeal was made to public *perceptions* that education research was of low quality. (Please note in passing, this appeal is based on anecdotal evidence at the same time as it produces policies critical of such evidence, and many similar examples can be found in recent Ministerial statements in Australia.) Perceptions of the status and standing of education researchers are one influence on the extent to which government bodies allow judgement of quality to be made from within this community compared with the extent to

⁵ These events and decisions took place in the late 1990s during my own period on the Executive and as

which they insist on other forms and sources of judgement (for example in the composition of assessment panels, and in the extent to which there is recognition of education research as a differentiated discipline or field). The founding fathers of the AARE took one recognized approach to building status and right to speak: they established markers of professionalism, hurdles, processes that marked out some form of eliteness for those who were part of the community. Given the debates of recent years about quality of education research, one issue for this conference is what can be done in these particular times that would improve our ability to speak and be heard and be represented in relation to discussions and assessment of research quality.

A third issue the reflection on AARE history bring to mind is that the research quality issue is explicitly both about *developing* good practice ('driving good behaviours' as the policy documents say) and about having *criteria or means of making assessments* about what is good and what does not measure up. The recent history of AARE has emphasized a certain approach to the first: giving support and training and encouragement, drawing people into an active research community; but it has, I think, been reluctant to take on the second, to make judgements about what is not quality research. Recent policy driven approaches both inside and outside Australia tend to take a different approach to the developing quality research issue: one that works by mandating elements of research training, or having external assessments of performance; that uses targets and consequences rather than support and encouragement as the main mechanisms. If these are not sufficient or appropriate mechanisms, we need to elaborate why and what else needs to be done. And we need to think about the gains and losses involved in emphasizing inclusion over standards.

Is impact a measure of quality?

Part 2: Revisiting recent history

Next, I want to take a few roles I have been in, over the past five or six years, and look at the form 'quality and impact' agendas took in those situations. My aim is twofold. I want to draw out some tensions and contexts that will continue to be part of whatever is developed, and might usefully be part of our explicit deliberations on the RQF in process at this point. And I want to refresh our minds about problems with the status quo. Personally I find change

difficult, and find it easy to think about problems with proposed new developments. So it is salutary to also be thinking about the problems of not making changes.

First consider my past three years experience on the ‘college of experts’ of the ARC, making assessments, reading assessments by others, observing assessment practices across disciplines, being asked to give advice to applicants, being asked to account for why education applications were not being more successful. My experience on this committee generated the opening sentence of the abstract I gave for this talk: “Neither impact nor peer review nor ‘addressing national research priorities’ are unproblematically appropriate criteria for quality assessment in the field of education...”. My concern was that ‘peer review’ by itself does not allow sufficiently for the patterns of work good researchers in education might need to be involved in; but both ‘impact’ and ‘national priorities’ judge researchers, at least in part, in terms of decision-making and practices un-related to research; they elide quality and effectiveness of research with measures of fiat and fashion.

In the current ARC processes, a large part of the assessment of quality research is an assessment of the quality of the researcher, and a large part of the assessment of the quality of the researcher is their production of peer reviewed publications, their achievement of academic distinctions, and their past success in similar assessments as indicated by their track record with previous grants from the same source. That creates potential problems for education researchers, in that time they give to professional engagement and dissemination is not recognized as part of their quality; and my suspicion is that even the cross-disciplinary assessment of peer reviewed publication is tainted by suspicion of publications whose nomenclature signals some closer relationship to the field of application as distinct from ‘pure’ academic interests. (That is, Psychology rates higher than Educational Psychology and that rates higher than journals without a traditional social science discipline nomenclature, such as journals of curriculum.) So in that sense, drawing a sharp hierarchy between ‘peer reviewed’ and impact or professional relevance work seems to distort what might be a quite appropriate and legitimate agenda of education researchers. On the other hand, I have a suspicion that ‘impact’ (or potential impact) is already a common element in how assessors, especially assessors outside education, rate the ‘significance’ of a project application. But, unlike with the publication data, and unlike with methodology and design discussions where although there are competing paradigms, there is large area of common

(or more transparent) criteria, the issue of ‘significance’ is not at all transparent, and indeed does bring in the issue of substantive value diversity (Yates 2004).

Now let us consider the current ARC processes in relation to those issues for our discussion I raised earlier: how to deal with substantive value diversity; the issue of voice and recognition of education researchers in making judgements; and the issue of sticks versus carrots, or assessment versus support as a mechanism of producing quality research. There is no perfect way of dealing with value diversity, but I think the ARC had developed relatively robust processes in relation to this, and a key element of that was using a process of multiple assessments, and multiple assessors. I think the concerns of Education researchers about relative success rates and how the quality of education research was being judged were much more a consequence of the size of the pool, and the problem of ‘impact’ being a tacit presence but something that was insufficiently brought to explicit consideration in terms of how assessors were using it. For example, in terms of judging ‘significance’, approaches that speak to a national policy question are widely understood; approaches that seek to influence professional practice of teachers in other ways can be read as more trivial. And ‘significance’ as a descriptor that might be applied to an advance in an *abstract theoretical* framework is widely accepted in areas such as Economics, but for Education it is less clear whether assessors allow significance to projects that does not have an obvious shorter-term applied outcome. Judgements of significance are also a consequence of relative representation: who is making the judgements and the balance of the social science committee.

Taking up the assessment compared with support spectrum I discussed earlier, the current ARC processes are directed to producing some hierarchies of eliteness; of differentiating out and providing more support for those doing the best work. In that context, the fact that people might be working in very different conditions (for example a teaching and research academic versus a research only academic; or in universities with significantly different workloads and equipment and funding for travel) was not part of the assessment of quality. Currently, at the competitive judgment end, the assessment is an assessment of what has been achieved; it is someone else’s problem how to allow that end point to be reached.

Another part of my own work over the past five or six years, has been to be one of those someone else’s. As a senior academic and faculty research associate dean, it has been part of

my work to be responsible for the research development and achievement of staff or students within a unit or within the faculty as a whole. This is a role which forces full-on attention to both upwards and downwards: to developing quality researchers and to demonstrating quality outcomes. Here one of the most evident problems was how to arrange workload and how to distribute the faculty budget across the various calls on it, in terms of teaching, administration and research. ARC grants for example are something universities are increasingly concerned to win, but from a faculty budget perspective they are a net cost to the faculty. How much do we allow or encourage specialization of activities and support those who achieve the most, and how much do we try to insist that it is everyone's responsibility to be producing both teaching and research quality? We know the rhetoric about doing both, but how does this translate in terms of workload policies, timetabling, resource allocation in a faculty? In practical terms, how many people should we employ in the teaching practice support office, and how many in research support roles? Around the country I know there have some very different approaches being taken to these decisions.

Now that benchmarking and performance indicators are so favoured, it is more evident that both individually and collectively a faculty support for one activity or direction *is* often at the expense of another. For example, a focus on increasing numbers of grant applications leads to a reduction in the number of publications in a given year. Or a focus on increasing research publication output leads to complaints from students and schools in the teacher training programs that academics are not spending enough time with them.

So one issue faced by Deans and Associate Deans (Research) in faculties and strongly felt in education as a field are the competing demands about types of activity to be produced, especially in terms of teaching and professional preparation as compared with demonstrating research eliteness. These middle managers can often be found caught in intense discussion about whether to develop research quality as a pervasive imperative for all members of the faculty, or to concentrate resources, specialize and reward stars.

Another issue in this role is that 'impression management' concerning the faculty as a research entity has become much more important both within the university and outside it; and this does not bear a 1:1 correspondence with the benchmark indicators. For example, if impact and relevance matters, there are prior questions about how do you indicate to potential partners that you are an appropriate one. In the competitive grants arena, the

demonstration of concentration and prior grant winning achievement rank high. But for government and other industry partners, different considerations matter. Personal relationships, time devoted to the task, efficiency of delivery, understanding the politics and values of the partner, and being able to communicate effectively with that partner are all important. Refereed publication, far from being a requirement or sign of quality, is often at best a distraction but sometimes a potential conflict of interest, a type of publication the partner is wary about. The other type of impression management that in my experience matters very much is how the education faculty or school is seen by the senior levels of the university, particularly the DVC Research. This non-transparent aspect of how quality gets built often influences what resources are effectively differentially distributed to different types of research enterprise; what types of collaboration opportunities are set up; what is promoted by the university to outside bodies, and what then feeds back into those wider indicators of research performance in terms of quality and impact. Research DVCs and Research Offices are influenced by performance indicators, but also by who they are themselves, their own biographical and disciplinary history, and by who they encounter from education, in terms of the gender, style, age, disciplinary foundation of those embodied individuals. These things do influence how education is assessed and how it is treated. I have seen examples of universities where a faculty judged good in the global research field is belittled in its own university environment; and I have seen poorer universities with a smaller component of individual high achievers nurturing its research, selling the quality of that group to important state partners, and eventually producing returns that enhance the indicators taken into account in peer reviewed performance as well.

In terms of mentoring staff, some of the issues I faced were these: how much to persuade staff into research concentrations and types of research that were not their initial enthusiasm; how much to emphasize the need for ongoing performance indicator runs on the board, compared with attention to building a depth of research development over time (this is more of a conflict with some types of research than with others); whether to encourage or discourage more junior staff in relation to collaborations with researchers at other institutions where we would not be the first named institution.

In recent years another issue that has been felt very much by faculties is the increasing emphasis on postgraduate completions and on some auditing of quality processes in relation to postgraduate supervision. Both the earlier DEST commissioned study on *The Impact of*

Education Research (Australia. Research Evaluation Programme. Higher Education Division 2000), and AARE's submission to the current RQF enquiry emphasize that in education as a field, a large amount of the research of the field (and indeed the quality research of the field) is done by practitioners in the professional field, often in the form of undertaking formal research degrees. The 'impact' here is not just about something being found out and applied somewhere else (though it may take that form), but one of immediate process, a matter of the practitioners and leaders of a field operating with quality because they have taken on a research orientation to their work, and because they themselves become more highly skilled professionals. So it is interesting that the current interim *RQF Approaches Paper* seems to have largely sequestered a focus on postgraduate students as outside the scope of what they are going to deal with, as something that will be dealt with by AUQA and other mechanisms.

I have had mixed feelings about a number of the moves of recent years to focus on completions and quality processes (and the moves in UK to begin to nominate components of adequate research training). On the one hand, all of those who deal with part-time mature students who often hold senior positions in their workplace and have mid-life family crises to deal with, know how difficult it is to attain the efficient completion processes envisaged in the data. Often the very government that is calling for better progress by postgraduate students is also the employer of those students in their working life, and intensifying their work in ways that makes it hard for them to focus on their research. On the other hand, I think it is the case, that the new explicit attention to quality of supervision has driven some positive and more systematic attention to this area.

Finally, another role I have had over the past few years is as an individual researcher. Here the issues I face in developing 'quality and impact' include: time on task, given that the type of research I do is not one that can be delegated to others with myself only coming in at the design stage, and the output stage to put my name on a final piece of writing that has largely been done by someone else; secondly, the collaboration issue – as I get more senior, this is actually slightly more difficult, as a lot of the people I know are also relatively senior, and we are all wary of the implications for our institution of collaborations where we are not the first-named CI – and yet the government wants (and I think it would be desirable to have) more larger-scale collaborative projects than we have had – ones that allow work that is both larger-scale but also detailed at a qualitative level and in terms of team intellectual resources

brought to bear on the work. (At the same time, I think research by sole researchers and scholars has been unjustly belittled because too much thinking takes the science laboratory as the only model of good work.)

So as one background for our discussions over the conference about what behaviours may be driven by a new form of RQF, it is worth considering some of the behaviours that have been driven by the emphasis on ARC grants, overall research income, and postgraduate completions in recent years. Some that I have pointed out here include the way that ‘impact’ has been both present and hidden in relation to assessing grants and also overlooked in the ways that researcher quality has been judged previously, and that we might benefit by a more explicit consideration of this of the type that the RQF consultations are bringing forth. Another is that the current metrics on which university funding is based, and some considerable part of the competitive grant assessment process do tend to drive short-term rather than longer-term work; quantity of applications and publications over quality; and to under-fund the most prestigious quality research, in that ARC grants tend to be a net cost to institutions in terms of the infrastructure they must provide, and in terms of not allowing funding of the salary of principle researchers. I have also suggested in this quick overview of some recent processes that judgements about quality tend to draw on visible and reasonably transparent indicators, but that they also draw on networks and associations and histories that lie outside these, and I think this will be the case whatever RQF is chosen, though the relative balance of these components and their substantive form can shift significantly. And I have suggested that the focus on quality in recent years has been driven very much by attention to end-point assessments and judgements (benchmarks, performance indicators) with little overt attention to unintended effects or simultaneous developments that worsen the conditions for developing good processes and outcomes, though this latter problem is intensely felt on the ground.

Now I want to turn to the final section of this paper and consider ‘quality and impact’ in relation to the RQF developments this year, and to draw out a few concluding considerations from the issues I have been talking about here.

Part 3:

Is impact a measure of quality?

Benchmarks, indicators, assessment panels, quality research and the RQF.

Earlier I mentioned that to be in *education* research is to name a relationship to a particular arena of activity, and that past processes of assessing quality which have no regard to that relationship, that is, that do not allow evidence of contribution to that field as part of the assessment of quality, would seem to be distorted. To my mind, that approach has in practice had some unfortunate effects in relation to such things as assessing track record, or posing the required work of those who work in education faculties as competing demands rather than ones that are coherent and relevant to both teaching quality and research quality agendas. Research, I argue, is an activity that is defined by most people as a practice with both technical/methodological parameters (it investigates something appropriately and systematically) and also substantive parameters (it makes a contribution to knowledge). The problem with peer review processes that left out impact or relationship to the field is that they could confine methodology and what counts as a contribution to knowledge in overly narrow terms. But the problems of trying to include ‘impact’ are not just the ones the *RQF Approaches Paper* is trying to come to grips with: how do you measure it? And who should assess it? The problem is also, what, out of a number of different possibilities, is this component trying to get at?

Hitler had impact. His research publication, *Mein Kampf*, had impact. My US-based colleague, Fazal Rizvi, has drawn my attention to a recent listing of influential books by a US conservative weekly, *Human Events Online*⁶ from a poll of ‘conservative scholars and public policy leaders’. This was their listing of books which had most impact *negatively* on economic and social life, their ‘Ten Most Harmful Books of the 19th and 20th Century’. No.1, predictably, was *The Communist Manifesto*, just edging out *Mein Kampf* which was in second place. *Quotations from Chairman Mao* was at No.3, *The Kinsey Report* was No.4, and, proudly standing there at No.5, was John Dewey’s *Democracy and Education*.⁷

This suggests some questions. Does ‘impact’ only count when you approve of its effects (and who gets to judge this)? Does it count whether or not other ‘quality’ criteria (for example, adequate methodology and peer recognition) underpinned the research? Do you

⁶ <http://www.humaneventsonline.com/article.php?id=7591>

⁷ For the curious, the remaining list was: 6. *Das Kapital*; 7. *The Feminine Mystique*; 8. Auguste Comte’s *Introduction to Positive Philosophy*; 9. Nietzsche’s *Beyond Good and Evil* and 10. John Maynard Keynes; *General Theory of Enlightenment, Interest and Money*. Honorable mentions included books by John Stuart

score more Brownie points if you can bring in a government and army on your side to make sure your work does indeed have the impact it so richly deserves? I am exaggerating of course, but these are real issues for fields of social research which have some normative component in relation to directions and aims, and in which the extent of ‘impact’ is tied to arrangements of funding and politics. To give one very practical example, consider Linkage research projects in Australia. Frequently the partner is the state government or an agency of this, and state government education policies in Australia have had very different histories, in their arrangements and structures and cultures, including histories of not wanting to credit or copy what another state has done. My experience on the ARC reading applications and also assessments of applications is that the local specificity of education poses some difficulties for scoring significance and broad contribution, when the funding scheme is self-consciously national and preferably international in its interests and benchmarks. It is clear too that state governments in Australia do spend differently from each other on education research, do set up different and differential possibilities of how research might demonstrate impact.

What ‘impact’ is about, in a sense is not a transparent issue. So let us now consider briefly the current RQF in Australia, and the agendas that are driving it.

Research is a key element of an innovative and economically prosperous nation and as such should be conducted within a sustained culture of excellence.

The Australian Government is committed to ensuring that resources provided to carry out research are directed to areas of research excellence and the impacts and benefits that result from the public investment in that research. A clear rationale for examining the quality and impact of research is that high quality research has the best chance of success in a global market and that will ensure a further deepening of Australia’s innovation base.

The Australian Government has therefore decided to develop a Research Quality Framework (RQF) that will provide a more consistent and comprehensive approach to assessing the quality and impact of publicly funded research. An RQF, when developed, may be the basis for future allocations of research funding.

Executive Summary, *RQF Issues Paper, March 2005*.

In their mid-year *RQF Approaches Paper*, the Executive Group summarize the original rationale for an RQF as being about developing ‘a more consistent and comprehensive approach to assessing the quality and impact of publicly funded research’ (p7) and that its aims are to:

- be transparent to government and taxpayers so that they are better informed about the results and benefits of the public investment in research;

Mill, Charles Darwin, Michel Foucault, Margaret Mead, Antonio Gramsci, Simone de Beauvoir, Ralph Nader and Sigmund Freud.

- ensure that all institutions are encouraged to focus on the quality and relevance of their research; and
- inform future research funding distribution.

RQF Approaches Paper, May 2005, 1.4.

Summing up, the rationale for introducing a new Research Quality Framework is set up in terms of accountability and driving better outcomes. The accountability which is spoken is in terms of fairness of process in allocating resources between fields of research and institutions and individuals; and also of justifying expenditure of funds to taxpayers. The outcomes that are spoken are of producing evident economic, social and environmental benefit for Australia; of increasing efficiency and effectiveness in terms of research resources; and of enhancing Australia's international standing in quality research performance.

What is not explicit is that the mechanisms being developed here, the forms these mechanisms will take, are not abstract and timeless ones, but ones that will be framed within a set of current, historically specific, taken-for-granted about how fairness is demonstrated, how outcomes are measured, how expenditure is justified. Producing quality indicators set against international benchmarks is part of a contemporary global zeitgeist of government. When I began in universities, a different zeitgeist reigned. Fairness included distributing funds equally to each department and funding departments so that time could be devoted to research as well as teaching as part of normal business. Governments were more concerned with their accountability for access to universities than with accountability for the internal processes or outcomes of universities. A lot more could be said about historical changes, and there is not some golden age where all good things were done, but rather times that produce different benefits and losses, and which, as they say today, 'drive different behaviours'. What I want to do here is simply recognize the context we are in, that it does have specific agendas, and it does have specific taken-for-granted about how performance and transparency are demonstrated. In this concluding section, I want to talk about some perspectives on the processes now being set in place.

Benchmarks and audits versus ad hoc encounters and prejudice

Central to both the current research funding distribution and to proposed forms of a research assessment exercise is an auditing culture, one that sets criteria and indicators and a formal regulated process by which performance is judged. We are all conscious of the increase in

the audit culture at all levels of our activities, and both as an individual academic, and also as someone who has been responsible for the performance of others, I have been weighed down by the time-consuming work of compliance with these processes, and also by the substantive measures and the worry about falling short on the various criteria (and I still am). But it is also worth thinking about the other side of this story: about the power and potential power of networks, ad hoc encounters, and prejudices.

To take one example, the competitive grant systems currently set up formal ways of assessing quality, with a lot of attention to technical procedures and transparency of process. But many times, I heard anecdotally that the Prime Minister's or Minister's view of whether a program is working well is highly influenced by chance encounters they have – for example of a captain of industry at a cocktail party who speaks well of a particular Linkage collaboration; or someone not connected with the research field who speaks of being a beneficiary of that research. I've had similar experiences of how Vice-Chancellors perceive the quality of research of their education faculties in three different universities. In one case a faculty that was highly regarded in the field of education research and shows up well on benchmarks, was not highly regarded by the DVC Research who simply did not rate education compared with science no matter how well they seemed to be doing. In another case, the VC thought well of a faculty because his wife worked in a school where the principal thought highly of it. And among faculty deans and promotion committees too, it is not uncommon to find halos or crosses attached to various people based on some ad hoc incidents, or more longstanding prejudices about certain types of work of the people making the judgements.

These prejudices and ad hoc influences happen whether or not we are part of an audit culture, but given the irritation of that culture, it is useful to remember the potential role explicit criteria might have as a brake on reward by unhampered prejudice. In terms of disadvantage in education, I have long believed in Jane Roland Martin's idea that while we cannot dispense with some modes and attributes being favoured over others, what we can do is try to show those who are disadvantaged the hidden curriculum. Arnot, David and Weiner (Arnot 1999) made a similar case about how some of the technical processes of management introduced to education systems under Margaret Thatcher, did have some advantages in terms of women's opportunities in the face of old boy networks.

Whatever forms of management and governance are adopted in the Australian research arena, it will not be the case that everyone can get the support they require; or that comparative and competitive judgements about relative worth of individual researchers will cease to be made in appointment and selection committees. What is being set up is a formal system for doing this, but the absence of such a system, as previously, does not imply the absence of comparative judgement. One issue for our deliberations here then is the adequacy of the particular measures chosen and of the rules about how quality is to be demonstrated. A second is the fairness with which certain criteria and processes can be operated, relative to the onerousness and cost of working with them. And a third is the responsibility to work with newer researchers to allow them to understand the rules of the game.

Getting a voice at the table

An ongoing concern of education researchers has been representation. We have been concerned about both the numbers and characteristics of those who get to be assessors in current ARC schemes, and of the potential make-up of assessment panels in a mooted research assessment exercise. Within universities too, many have been concerned about whether education research gets taken seriously as a distinctive voice in discussions about quality, postgraduate performance and the like. As an association, AARE has lobbied for education to be given greater representation in funding mechanisms. But in this paper I have also been trying to remind us of the broader backdrop, especially politics and reputation, which also needs to be part of our concern.

What influences whether education research as an association or as a field of research does get taken seriously? Relationships with other fields and other researchers; relationships with particular policy-makers and with the voting public through what is happening in schools and other arenas of education; what we can say about our own mechanisms of training and quality; what international heavy-hitters we can call on to attest to our status and impact, all might matter here. The global movements of recent years to define quality research as having a certain scientific form, and to directly and indirectly impugn the quality of the bulk of work that is done under the name of education research is an important issue to be addressed in today's context (though a more sensitive inclusion of 'impact' might actually help on this front). To deal with it involves more than simply reiterating the arguments that have been found convincing *within* the community of researchers.

Another issue in getting a voice at some important tables is one I mentioned in recalling the history of AARE at the beginning of this talk. One element of the current declared agenda of the government is to look for and nurture *eliteness* in Australia's research activity. Quality assessment exercises in the current era are about judgements of comparative standing. Such judgements have been at work in the past, but these developments formalize them and give them more force. As a research community, I think we have been trying to meet this agenda in a limited way by reference to evidence of international standing in publications and citations (and we could be making more of the numbers of key Australian researchers poached to high positions in other countries over the last ten years). But we have been more actively pursuing concerns about having research more widely embedded and nurtured across institutions involved in education. Responding to what seem inevitable moves to visibly rank and compare researchers and their institutions I think will be a significant challenge to AARE as well as to individuals and the different places we work in.

'Impact' and diversity versus homogeneity as quality research

One of my own biggest concerns about education research and where it is going is of a likely narrowing down in both form and focus of the work that will take place. Here I am not talking about the call for greater concentrations of research activity, aspects of which I think would be beneficial. I am talking about the material impetus over the past 15 years in Australia for education research to concentrate on topics designated by the government (state or commonwealth) and to take up approaches that bring clear short-term outcomes, and to use methods that can justify large calls on expenditure. In Australia, philosophers and historians have dwindled in number in education, though the contribution of these disciplines is acknowledged in the broader research circles (at least two of the very few federation fellows in the social science have been philosophers, and issues relating to citizenship and politics today have considerable scope for a contribution to knowledge from both philosophers and historians). It is widely understood in fields across the research spectrum that a serious research culture should include and nurture a wide range of activity, from blue-sky and fundamental, so-called 'pure' research; to activity focused on particular more accessible parts of the big agenda, or activity that is trying to solve a problem in the field as commissioned by someone else. Precisely because education is a normative field, and because politics and times change, I think it impoverishes the quality of resources available even to applied concerns, to only have research taking place that is positioned tightly within current concerns.

The foreshadowed moves in the present RQF discussions to extend attention to impact have potential to mitigate some of the concerns I have discussed earlier in this paper, but some potential too to heighten the trajectory of tying research to the short-term politics of the day, and to measurable outcomes as the only driver of quality. ‘Impact’ in education research is currently well understood in terms of research designed to speak to policy-makers’ and system questions, but the current RQF and associated submissions, and the resulting *Approaches* paper by the Expert Advisory Group I think has made some advance on that unexamined assumption, by beginning to note and think through some of the *different* ways research in the humanities and social science might have ‘impact’, for instance, they note from a number of submissions some reservation that it can easily be measured in any short-term way. Potentially here there may be a broadening of the ways research can make a case about its significance, because a criterion has been made explicit and examined, not just enacted. But the effects of course will depend on how the processes of assessment are set up: on who assesses as well as what criteria and evidence are considered legitimate.

Developing quality education research in Australia

To some extent, in recent years, AARE and the government have been working on different questions in relation to building quality research in Australia. AARE has been trying to deal with the questions, ‘how do we build, support and enhance a quality research culture and one that is appropriate to the field of education?’ and ‘how do we get more support for education research as a field?’ DEST and the ARC have been working on the questions, ‘How do we sort out better and worse research in Australia? How do we make Australian research more efficient and effective? How do we get better value from it? and How do we better establish its international standing?’ We do not have the power to ignore the questions the government is enacting, but nor should we give up on the questions we have rightly been concerned with.

References:

Annot, M., David, M. & Weiner G. (1999). *Closing the Gender Gap: postwar education and social change*. Cambridge, Polity Press.

Australia. Research Evaluation Programme. Higher Education Division, D. o. E., Training and Youth Affairs. (2000). *The Impact of Educational Research*. Canberra, Commonwealth of Australia.

Bessant, B. and Holbrook, A. (1995). *Reflections on Educational Research in Australia: a history of the Australian Association for Research in Education*. Coldstream, Australian Association for Research in Education.

Yates, L. (2004). *What does good education research look like? Situating a field and its practices*. Maidenhead and New York, Open University Press.