

[Alcorn, N. (2009). Evaluating Quality in Educational Research: Issues and Challenges of the PBRF. *New Zealand Annual Review of Education*, 18, 65-87]

## Evaluating Quality in Educational Research: Issues and Challenges of the PBRF

NOELINE ALCORN

### *Abstract:*

*This article argues that we, as a community of educational researchers, should take up the challenge posed by Hattie (2005) to create a robust public debate about what constitutes quality in educational research. It also identifies specific issues that warrant urgent attention. The paper outlines the methodology used within the Performance-Based Research Fund (PBRF), in particular with regard to its Education Sector ratings, and goes on to discuss ways of raising achievement and increasing capacity. A major concern is to enhance the status of applied or practice-based research and to foster discussion of appropriate ways to evaluate and assess its impact and ethical basis, without losing sight of the need for clarity and rigour. The article calls for education academics to engage in the debate over this evaluation not just for the sake of raising ratings but in the interests of greater understanding and recognition of the purpose and processes of educational research more generally.*

The first Performance-Based Research Fund (PBRF) evaluation exercise in 2003 generated vigorous public debate among members of the education academic community. A special seminar organised by the New Zealand Association for Research in Education (NZARE) led to the publication of a book entitled *Punishing the Discipline: The PBRF Regime* (Smith & Jesson, 2005). A seminar jointly sponsored by the Tertiary Education Commission (TEC) and the Institute of Policy Studies at Victoria University of Wellington, resulted in a second book, *Evaluating the PBRF: Framing the Debate*, which included further evaluations from educationists (Hall & Morris Matthews, 2006; Middleton, 2006). Much of the literature was critical both of the PBRF

66 *Noeline Alcorn*

methodology and of the ways it was thought to disadvantage education academics and to downgrade the importance of teaching. Members of the Education panel collectively wrote a paper describing the process and offering guidance to individuals and institutions for the next quality assessment round in 2006 (Alcorn et al., 2004). There has been less response to the 2006 round. This article argues that we, as a community of educational researchers, should take up the challenge posed by Hattie (2005) to create a robust public debate about what constitutes quality in educational research. It also identifies specific issues that warrant urgent attention.

Education has fared poorly in the PBRF statistical stakes, scoring close to the bottom of the disciplinary quality scores in both 2003 and 2006, but this overall result gives a false picture of research in the field. Education is both an established and an emerging research community in New Zealand. Useful masters' theses were being produced in the 1920s and 1930s, and the establishment of the New Zealand Council for Educational Research in 1934 provided an impetus for a wide variety of research studies. University Departments of Education became increasingly research active and produced internationally recognised work. The many and varied contributions of Ralph Winterbourn in special education, Ivan Snook in the philosophy of education, Roy Nash in sociology, Marie Clay in our understanding of the reading process, Warwick Elley on literacy and assessment, and Graham Nuttall on the nature of classroom learning are just a few, admittedly selective, examples. The PBRF results in 2003 and in 2006 indicated that there are a substantial number of educational researchers whose work is rated as internationally or nationally significant (numbering 128 in 2006). However, there are also a larger number of education academics who are unable as yet to meet the quality standard.

A significant proportion of education academics are comparatively new researchers who began their careers as tertiary teachers without a research imperative, and whose former working conditions and hours reflected their heavy teaching load. While all institutions have taken steps to help those staff adapt to new ways of working, it is likely to be some time before the field is "saturated". In the interim, as Middleton (2005) observed, staff who had hitherto been regarded as professionally active and successful found their assessment in the first PBRF round destabilised their sense of professional identity and caused them to feel (and sometimes to be labelled by their institutions) failures. A further problem for education is that many education researchers are engaged

in applied or practice-based research, which many writers claim is seen as less prestigious than other forms of research (Hall & Morris Matthews, 2006; Morton & Gordon, 2005; Higgins, 2005; Haigh, 2005). Middleton (2005) suggested that part of the problem was a perceived lack of fit between the PBRF and the field of education, noting that some academics felt the PBRF definition of research was too narrow.

## Context

### *Intent and impact of the PBRF*

The PBRF was designed “to ensure that excellent research in the tertiary education sector is encouraged and rewarded” (Tertiary Education Commission, 2009). As the initial Working Group document, *Investing in Excellence* makes clear, that “involves measuring and making judgement about quality” (Performance-Based Research Fund Working Group, 2002, p. 7). A methodology was developed and disciplinary panels set up to assess evidence portfolios from staff employed as teachers and researchers in degree programmes. All academics whose roles included teaching and/or research, who were employed on contracts of a year or longer, and whose hours of work were 0.2 or above, were eligible for consideration. The PBRF unit of assessment differed from that of the British Research Assessment Exercise (RAE) in that individual rather than group or departmental portfolios were to be submitted and results based on an aggregation of individual scores. Individuals submit four nominated research outputs they consider their most significant in the six year period under review, provide a list of up to 30 other publications, describe their contribution to the research environment (CRE), and give examples of peer esteem (PE) for their research. They are informed of their own quality rating, which may be an A, B, C or R which has been derived from their quality score weighted 70%, CRE weighted 15%, and PE 15%. These quality scores, though extremely important for the individual, when aggregated provide the basis for 60% of the PBRF research funding of the institution, and thus in turn have a strong bearing upon its reputation. Also taken into account are research degree completions and external research funding.

A further aim of the PBRF Working Group of 2002 was to ensure that PBRF procedures were comprehensive and covered the “full range of original investigative activity that occurs within the sector, regardless of its type, form or place of output” (p. 8). It insisted that research was not only about the production of knowledge but also its application and dissemination, as well as supporting current and potential colleagues.

The PBRF definition of research, already broad, was strengthened in 2006 to ensure that the contribution of the performing arts was fully realised. The PBRF guidelines are explicit and catholic, the Education panel for both assessments reported consistency among the ratings ascribed by panel members, the panels reviewed their decisions exhaustively, and they were subject to moderation processes. Nevertheless there will inevitably be an element of professional judgment involved, and some individuals who feel a sense of disappointment and lack of appreciation of their work. There has been a wider disquiet, shared by academics in other applied and practice-based fields, that pure or theoretical research is privileged and that high-status journal articles are seen as a gold standard, making other forms of publication less valued.

### *Changes to PBRF education ratings between 2003 and 2006*

The 2006 PBRF quality evaluation was a “partial round” in which academics whose portfolios had been assessed in 2003 could opt to have their score carried forward. A high proportion across all disciplines chose to do so. All new academics were required to submit a portfolio and many who hoped to raise their score also submitted. The Education panel reviewed 418 portfolios in 2006, as compared to the 493 reviewed in 2003. (Eligible staff in 2003 numbered 1077, and in 2006 were 1057). Like those in other disciplinary areas, education academics in general improved their scores. The percentage of R scores dropped from 73.7% to 65.03% (Tertiary Education Commission, 2007) and the average quality score increased from 1.02 to 1.31. While in part this indicated that institutions had taken greater care to complete the evidence portfolios consistently and fully, it also reflected considerable effort to develop a research culture. But there were also examples of what might be termed “game-playing” by institutions. A case in point was where two universities reached an agreement with the TEC that scores of former staff of Colleges of Education, now merged with the university, would be reported separately.

A significant development implemented in the 2006 quality assessment round was the establishment of new quality categories: C (NE) and R (NE). This was in response to feedback from panels after the 2003 round. Concerns were expressed that:

The assessment criteria could provide disincentives to staff who had recently completed a PhD, which required intensive concentration on research. Some of these staff were building a research platform

but achieved an “R” quality category since they were unable to demonstrate sufficient peer esteem (PE) or contribution to the research environment. (Tertiary Education Commission, 2008a, p. 5)

The added categories enabled new and emerging (NE) researchers to be awarded a C grade on the strength of a completed PhD and two quality assured publications without needing to provide evidence of PE or CRE. Overall the TEC found that the change was successful: the development of an assessment pathway specifically for new and emerging researchers was a very significant improvement to the assessment framework of the quality evaluation (Tertiary Education Commission, 2007, p. 284).

This judgment was based on the significant number of new and emerging researchers whose research was acknowledged and recognised by the award of a funded research category in the 2006 PBRF assessment. In the 2006 Quality Assessment 20.4 percent of all researchers (1,768 researchers) claimed NE status. However, there were considerable differences between subjects, from Physics with no “NE” researchers to Theatre and Dance, Film Television and Multimedia with 40.4 percent. Almost half (46.7 percent) of “NE” staff received “C (NE)” scores, enabling them to achieve funded status. This score equated to just over one quarter of all those who received a “C” quality category. (Tertiary Education Commission, 2008a, p. 7)

However, the results for Education were discouraging, with only 14% of those who claimed NE status being awarded a funded category. Education remained one of the three lowest rated subject areas. Clearly those qualifying as NE researchers in education did not fit the pattern the category demanded. Education academics rarely complete PhDs as fulltime students and move on to postdoctoral fellowships. Instead they are appointed from fulltime professional posts and pursue doctoral studies part-time on top of demanding academic employment. They are unlikely to complete a PhD in less than five years and after graduation have less time to make the transition to publishing their work quickly in quality assured publications. Often they lack mentors who can facilitate the process. Like their colleagues in the humanities and social sciences they normally find doctoral work a lonely process rather than being part of a research group that aids publication and offers well-defined career progressions. Those who have not completed a PhD in the assessment period may have been working on the degree for the entire time and eschewed other publications as distractions from their major work.

## Raising the Status of Educational Research

### *Improving overall ratings*

The key issue is how Education can improve its overall rating. One possibility is to implement an assessment process like that in the UK, where the Research Assessment Exercise (RAE) measures the excellence of the work submitted but does not take into account the work of academics whose outputs are not submitted. Thus excellence can be rewarded without public attention being drawn to those whose work is less highly rated. But any system will be subject to game playing, and in the RAE many institutions exclude the work of those staff judged internally to be achieving satisfactorily but not “at an international level”. As a result, education results look impressive. In 2008, 75% of work submitted for the RAE was judged to be at international standard or above, 40% of international excellence or better and 15% world class. Five departments had a quarter or more of the work submitted judged world leading (Research Assessment Exercise, 2009). However, the proportion of academics submitting work continued to drop. For the 2008 round in the UK, only 1903 research-active staff had work submitted, a small number when we consider there were 978 eligible staff in the 2006 PBRF quality evaluation in New Zealand. One result of this process is to concentrate research funding in fewer institutions. After the 2001 round, only 33 of the 83 UK universities that entered education academics were funded, and only one post-1992 university (Furlong and Oancea, 2006, p. 90). As a result, a large number of institutions teaching Education received no research funding from RAE sources. In a profession that sees itself increasingly as evidence-based, this is problematic. It is unclear at this stage to what extent the 2008 RAE assessment, which identified “pockets of excellence” across a wider range of institutions than before, will affect education research funding.

For education researchers concerned to raise the status of their field there appear to be two major strategies: working to stress the importance of professional research and the measurement of its impact; and working with academics who are new to research to help them develop as writers and researchers as well as teachers.

### *Research capacity building in education*

Tertiary institutions and education academics can help new researchers in a variety of practical ways through mentoring them into publication, helping them learn to write and/or present findings in a variety of forms

and to develop an understanding of academic conventions. New researchers need advice about starting small, but setting their work in a context of what is already known. They also need time and resources to establish a research programme. Institutions can give feedback on portfolios, on identifying the significance of research projects and their practical and theoretical impact, including the development of capacity among the wider profession. Many tertiary institutions have established research officers to provide practical support to researchers and/or developed mentoring programmes. It is important that highly rated researchers are not only recruited to help boost academic ratings but also expected to develop a research programme that will include less-experienced staff and assist their development as researchers.

The Education panel in 2006 noted in its report that evidence portfolios (EPs) were much better prepared and presented than in 2003. Seven of the eleven panel members had been part of the 2003 panel, which aided consistency of scoring and enabled anecdotal comparisons to be made. There were a number of EPs which demonstrated that while some research was being completed it was insufficient to demonstrate a platform. The panel made recommendations to assist individuals and institutions prepare for the 2012 round, by:

- encouraging individuals to move beyond presenting at conferences by using the feedback gained to turn the presentation into a scholarly work for publication;
- clarifying what quality assurance a publication had been through. This was of particular importance where in-house and new publications were cited;
- checking the robustness and comprehensiveness of the information submitted, especially in clarifying the “my contribution” section;
- clarifying the context for the research and the reason for publication in particular outlets. For example, Pacific research might require a local publication;
- exercising judgment in determining the four most significant publications; and
- providing additional support for Māori and Pacific researchers.

### **Developing A Robust Appreciation of Applied and Practice-Based Research**

In addition to enhancing research capacity in general, education academics need to work towards enhancing the status of applied or practice-based research. They need to foster discussion of appropriate ways to evaluate its impact and ethical basis without losing sight of clarity and rigour in its reporting. Such discussion is more advanced in the UK than in New Zealand; hence this article draws heavily on UK work. However, Adams (2008) noted disquiet about the perceived status of applied research in the 2006 PBRF round, and the Tertiary Education Commission Sector Reference Group (2008b) produced a consultation paper on the evaluation of applied research. These papers indicate the issue is a significant one for New Zealand.

#### *Stressing the importance of practice-based research*

For many education academics, maintaining contacts and dialogue with schools, professional groups and government agencies is a key priority. Dialogue with these groups helps to define areas for further investigation; problems of practice, and theoretical underpinnings for policy are all important. Finding ways to ensure that professional advice is research-based and that innovation is tested and evaluated is an essential aspect of their work. Many researchers regard teachers as partners in their research not as subjects to be studied. Some of this work is local and specific and raises issues of validity and generalisability, though it can also be comparative or national in scope.

In Britain the need for applied research in education has been highlighted and debated hotly ever since Hargreaves' (1996) speech to the Teacher Training Agency (TTA) in which he lamented the lack of relevance of much research to the policy and professional communities, and suggested that the medical profession provided models for developing evidence-based practice from applied research. His views and their endorsement by Tooley (1998) and Woodhead (1998), then Head of Ofsted, provoked spirited exchanges of views within the research community itself (Hammersley, 2004; Pring, 2004; MacLure, 2005). There were also attempts to restrict what counts as research to quantitative and replicable studies, advocated by Slavin (2003), and endorsed by the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre) in London.

Thirteen years on in the UK there seems to be a growing consensus that many forms of research can contribute to both practice and policy

making. In the 2008 RAE the Education sub-panel included “user representatives” in its membership. These panelists added an appendix to the main report endorsing the assessment process and attesting to the strength of the field. They called for a broad approach to defining applied research:

We would like to take this opportunity to say that as users we often need and value ideas and theories as much as data and evidence; sometimes it is a new and cogent way of conceptualising a problem, of discarding guiding assumptions, that is needed, at others it is an explanatory underpinning theory, at yet others it is the different perspective offered by combining disciplinary and/or methodological approaches. These can be as influential on policy as analyses of large/longitudinal datasets are assumed to be. (Research Assessment Exercise, 2009, p. 5)

#### Defining applied research

The OECD Frascati Manual defines applied research as “original investigation undertaken in order to acquire new knowledge ... directed towards a specific practical aim or objective” (OECD, 2002, p. 78). Within the scientific community, Stokes (1997) demonstrated through his Pasteur’s quadrant model that use-inspired research might provide basic results.

		Consideration of use	
		No	Yes
Quest for fundamental understanding	Yes	Pure basic research (Bohr)	Use inspired basic research (Pasteur)
	No		Pure applied research (Edison)

Figure 1 Pasteur’s Quadrant (Stokes, 1997, p. 8)

Furlong and Oancea (2006), writing of applied research in education, adopt an inclusive definition, “seeing applied and practice-based research as *an area situated between academia-led theoretical pursuits* (e.g., historical research) and *research-informed practice, and consisting of a multitude of models of research explicitly conducted in, with, and/or practice* (p. 94).

As Oancea and Furlong (2007) point out, not all practice-based research in education is instrumental or strategic. Proponents of action research (Carr & Kemmis, 1986; Elliott, 2001) claim an emancipatory focus, and demonstrate “how research may contribute to theoretical knowledge while at the same time being part of changing practice in a way that is intrinsically worthwhile” (p. 123). Schon (1983, 1991) advocating reflexive practice, and Loughram (2000) working with self-study, both illustrate the complexity of the interactions between theory and practice.

#### Making distinctions between professional practice and practice-based research

Education is an applied discipline and much (though not all) of the research of education academics is applied. The PBRF definition of research is a broad one, embracing pure and applied research, creative works and Māori knowledge. But in both PBRF quality evaluation rounds a number of academics from a range of applied fields have felt their work was not appropriately measured. A commissioned report by Jonathan Adams, an independent British researcher, found that “some (but not all) people feel that the outcome of assessment tends to devalue practice-based activity” (Adams, 2008, p. 48). He commented “the problem is that much professional activity does not readily qualify as ‘original research’, but the boundary between strict professional practice and professional innovation (essentially a form of translational research) is inevitably blurred” (p. 49). A report on the 2008 RAE assessment suggests that this is a problem for the UK as well as New Zealand:

Given that many researchers in education come to research from a teaching career or an academic career outside the discipline of education, we think that it might be advisable in any future exercise to make an even clearer distinction between pedagogical research in higher education and descriptive or anecdotal accounts of teaching developments and evaluations. Several outputs fell into the latter category: they were often very interesting and worthwhile in their own right (and would probably have a strong appeal for practitioner readers), but did not meet the definition of research for the RAE or made only a limited contribution to their area. (Research Assessment Exercise, 2009, p. 8)

PBRF guidelines provide a number of examples of differences between research and professional practice. For example, preparing for teaching, routine data collection, providing advice or opinions (except where

consistent with the PBRF definition of research) and any other routine professional practice (e.g., in arts, law, architecture or business) that does not comply with the definition (Tertiary Education Commission, 2005, pp. 20-21) are specifically excluded. The education panel offered further guidance:

Descriptive reports of classroom practice are not research. But an analytic account, set in the context of other research, can be the basis of research. Curriculum documents are not of themselves research. However, a paper examining the intellectual processes involved in their development and the consultation of other research literature may be research. (Tertiary Education Commission, 2005, p. 86)

The engineering and management panels also provided detailed and specific advice about quality assurance for non-traditional research studies, and exemplars of measuring impact.

It may well be that the guidelines need to be addressed further and made even more explicit. A TEC paper notes that some new and emerging researchers, for instance, would welcome exemplars (Tertiary Education Commission, 2008a). But for the exercise to have real benefit, the education academic community needs to engage in robust debate about the nature of applied research in the field and the ways in which it should be evaluated. This would help inform both the TEC and future PBRF Education panels and provide guidance to individual academics and research groups.

#### *The evaluation of practice-based research – UK experience*

The evaluation of any research attests to its quality. But quality is a multi-faceted concept. Yates (2004) in her book *What Does Good Education Research Look Like?* suggests that thesis examiners, journal editors, research grant funders, teachers and other practitioners, and book publishers all have different views of what good research looks like. It is difficult for panels in a research assessment exercise to take account of this diversity, and critics continue to assert that current definitions of research favour publication in traditional outlets where the perceived reputation of a journal is a form of quality proxy.

Both the Roberts report (2003) and the Lambert report (2005) in Britain recommended that further attention be paid to ways in which applied research could be more fairly and generously assessed. Roberts was later invited to chair the Expert Advisory Group (EAG) for a Research Quality Framework (RQF) in Australia. The report of this group (2005) recommended that both academic quality and the wider

impact of research should be assessed. Many of those consulted suggested that assessment of both economic and non-economic impact should incorporate the views of end users. The EAG report acknowledges the difficulties inherent in this approach because of time lags, and the problems of establishing direct causal relationships between research and changes in practice.

Following the lead of the Roberts report in the UK, the EAG believed that one way in which impact could be taken into greater account would be through the involvement of end users in the assessment process. It also noted that hard and fast metrics are not a suitable way to assess wider impact.

#### *Developing a theoretical framework for the assessment of applied and practice-based research*

In 2004, Furlong & Oancea were commissioned by the Economic and Social Research Council (ESRC) in the United Kingdom to undertake a project “to develop an understanding of quality that could assist subsequent development of quality criteria appropriate for different types of applied and practice-based research” (Furlong & Oancea, 2006, p. 91). They carried out a literature review, interviewed researchers and policy makers, and discussed their draft findings at three conferences and a special consultation day. They note the plethora of attempts to define what good educational research is, following the challenges laid down by Hargreaves (1996) and the increasing interest in research by policy makers.

Applied and practice-based educational research is of increasing importance for governments and many other bodies. Within the UK, some of the most high profile examples of applied and practice-based research initiatives are the Best Practice Research Scholarship Scheme, the Scottish Applied Educational Research Scheme, the Networked Learning Communities initiative supported by the National College for School Leadership, and the ESRC’s Teaching and Learning Research Programme. (p. 89)

In New Zealand, the Ministry of Education has sponsored and published a number of Best Evidence Syntheses, provided funding for the Teaching and Learning Research Initiative (TLRI) and promoted evidence-based practice in schools.

Furlong and Oancea are adamant the first consideration is the robustness of the theoretical or methodological framework, its fitness for purpose, its originality or new insights, its justification of argument on

the basis of research evidence or argument, its validity and trustworthiness and the researcher's use of other thinking and research in the field. In a second paper, after noting that a "powerful response" to their surveys was that traditional concerns about knowledge and methodological rigour were seen as fundamentally important in the evaluation of applied research, they went on:

Nevertheless, in the case of applied and practice-based research, it was argued that considerations of worth could not stop there, but needed to be balanced by further concerns, related to the relationship between research (its processes and outcomes) and practice and policy. Ignoring this in research assessment, it was argued, would be a missed opportunity to understand the theoretical and methodological concerns of applied and practice-based research in their own terms. (Oancea & Furlong, 2007, p. 125)

They rejected a technician approach and the production of checklists, choosing to reframe the discourse on research quality in more philosophical terms.

Our project was therefore an attempt to reframe the problem, by interpreting "application" as a complex entanglement of research and practice, "assessment" as deliberation and judgment, and "quality" as excellence or virtue, in a classical (Aristotelian) sense of the terms. (2007, p. 121)

They identified three domains *theorosis*, *poiesis*, and *praxis*, which are different yet complementary, claiming "it is possible to identify different forms of excellence or virtue in each of these different domains" (2007, p. 124). Identifying excellence thus becomes a matter of recognizing the distinctiveness of each domain (Figure 2).

Excellence in *Theorosis* or demonstrable knowledge may be characterized by trustworthiness, contribution to knowledge, transparency and explicitness in design and reporting, and paradigm-dependent considerations. It has similar characteristics to basic or scientific research and would normally work from the general to the particular and from abstraction to the concrete.

*Poiesis* or technical skill can be assessed on its "value for use" but this is defined much more broadly than providing solutions to problems posed by users. Oancea and Furlong note that Nussbaum (1986) argues for universality, teachability, precision and concern for explanation, and that Weiss (1987) describes "knowledge creep" to explain the impact of research on changing our assumptions about the world. This is a

concept taken up by Gardner (2004). Assessing such research may involve determining fitness for purpose, specificity, concern for enabling impact, operationisability and propriety. Some forms of "techne" may be assessed on their strategic and economic value.

As to *Praxis* (the Phronetic domain), Oancea and Furlong assert that "because the roots of phronetic engagement of research and practice are in ethical concerns and in tacit, situated knowledge, it is extremely difficult to capture in the research appraisal process. The best we can hope for is probably to identify some features that might in some way be connected with these aspirations, and try to take them into consideration whilst evaluating it" (2007, p. 132). Ethical action is essential. They suggest that features that could be considered include deliberation, reflexivity and criticism, plausibility from a practitioner's perspective, receptiveness and dialogue which entails welcoming the difference in perspective and jargon of others, engagement, including partnerships.

They categorise their framework as "a discursive tool for catalyzing the ongoing conversation about quality" (2007, p. 127).

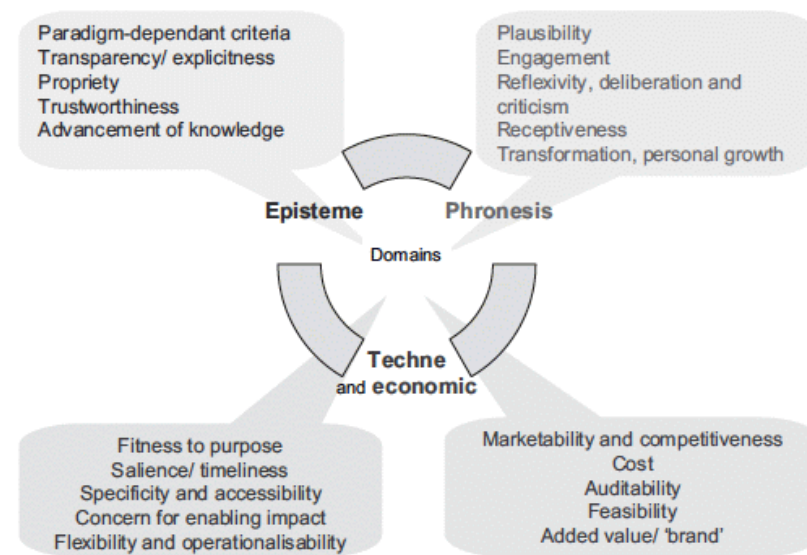


Figure 2. Domains of quality  
(after Oancea & Furlong, 2007, p. 133)

### **Implications for Educational Research in New Zealand and its Assessment in the PBRF Process**

#### *Insisting on high quality applied research and publicizing exemplars*

In an earlier paper (Alcorn, 2007) I pointed to the work of three New Zealand education academics whose practice-based research has received international recognition. This research, grounded and contextualized in practice, was rigorous and theorized. The report of the RAE panel in 2009 suggested that the importance of rigour and theorization was one that a number of researchers paid less attention to than they might. They commented that research needed to be contextualized in a number of ways, both in the messy realities of practice and also in the research literature, building on what has gone before, taking issue with earlier analyses and offering fresh viewpoints. This could help avoid inflated claims for small-scale studies and also enable synthesis of such studies. Research leaders and mentors of education academics in New Zealand institutions need to insist on this dual contextualization and perspective.

These issues are identical, whether the research is for a university qualification, a report to an external agency, or submission to a prestigious journal. Before publication (or the award of a passing grade for a thesis), the piece of work should be reviewed by academics experienced in the field. This is another area where academics need to remain vigilant. We have had little public debate about eroding tertiary standards in this country, but in both the USA and the UK there is ongoing disquiet amongst academics who have been pressured to award higher grades than they thought appropriate for research theses. We must ensure that in our efforts to assess less traditional research outputs fairly and according to their intentions, we do not lose sight of traditional virtues.

#### *Initiating public debate on the evaluation of applied research in education*

For those who are evaluating applied or practice-based research, the assessment of research impact is crucial. For some disciplines, impact is measured largely through a citation analysis. This may work well for some disciplines but it is of dubious value in many of the social sciences and humanities. It assumes that impact is measured entirely by other academics. It is generally Euro-American in scope, ignoring publications in other parts of the world. Educational practice tends to be culture- and context-specific, and while we can and do learn from each other, policies

and practices do not necessarily translate well. At a seminar jointly sponsored by the TEC and the Institute of Policy Studies at Victoria University of Wellington in 2008, participants were warned by several speakers that such an approach should be treated with caution. Debate on the new metric system that will replace the RAE in Britain has stressed the absolute imperative of peer review in the humanities and social sciences.

For disciplines like education, which include applied and practice-based research, assessing research impact is considerably more complex than citation analysis. The impact of research on policy and practice is important but problematic as causality is difficult to determine. Roberts and the EAG in Australia recommend adding research users or members of the profession to research assessment panels like those in the PBRF, to ensure that such issues are not ignored. This practice is already followed in New Zealand by the engineering panel, which also includes applied research.

Education is about people, and practice-based research raises a number of ethical issues that need to be sensitively addressed. Consideration of how these ethical issues are handled is an important part of research evaluation. The extent to which research partners and participants benefit directly and indirectly, as well as the imperative to ensure that no harm is done, must both be assessed. In our New Zealand assessment system, though PBRF panelists read up to 50% of the Nominated Research Outputs (NROs) they assess, nevertheless they must also rely on the reputation of the journal editors and referees on whose advice a paper was published, especially if the journal is prestigious or well-known and respected. This may mean less direct attention being paid to ethical and other concerns.

#### *Ensuring quality assured applied research is submitted to the PBRF*

There is a considerable amount of practice-based research, often action research or evaluation research, funded by external agencies and carried out by New Zealand academics. It is unclear to what extent the findings from these projects are submitted as nominated high quality research outputs for the PBRF. Commentary from the UK shows that educational institutions may be wary about how highly they would be rated.

Much of that government investment has explicitly been directed towards attempting to secure a stronger evidence base for decision-making in policy and practice, with a concomitant investment in research access and dissemination activities. We also



know that the Teaching and Learning Research Programme (TLRP) and Applied Educational Research Scheme (AERS) have supported and funded a wide variety of research-based activities and outputs for different user audiences, but we noticed that only a small proportion of these outputs (and those arising out of the plentiful research funded by government) seem to have been submitted to this RAE. We accordingly wonder whether clear signals about the relationship between quality in research as measured by RAE and the greater UK investment in research-for-use have yet to be given. At this stage, it appears that the risk institutions perceive of submitting research-based user-focused outputs has led to limited submission of this work. As the Panel took pains to set out the criteria in such a way as to encourage submission of this work (see p. 31, paragraph 19 of the RAE panel criteria for UOA45), we think it is a very important issue to pursue. (Research Assessment Exercise, 2009, p. 8)

This is an important issue for New Zealand academics to debate now, and also when the panel-specific guidelines for the next PBRF are sent out for consultation with the sector. If institutions take a conservative approach and self-censor research outputs to continue privileging international journal articles, without assessing the impact of more practice-based contributions, then they risk playing the game in a way that militates against the recognition of this form of research.

#### *Widening outlets for research dissemination*

The issues of how and where research is disseminated pose further challenges for the assessment of quality in applied research. Rigorous and robust studies can be reported in a number of ways. The dominance of the refereed journal, preferably published outside New Zealand, is both promoted and lamented. It has assumed even greater dominance since the advent of the PBRF and academics are sometimes urged by peers and/or management in their institutions to aim for higher quality international publications and avoid New Zealand journals. The percentage of research books has been dropping, at least in the UK. But there are other ways in which findings can be communicated. One of the two British universities to be awarded a grade of 5\* in the 2001 RAE entered a website as an output. A key criterion is clarity. The RAE report on the 2008 quality assessment in education notes:

We read outputs in which sophisticated and complex ideas and data were vividly and eloquently communicated by their authors, whilst in others valuable material was expressed in an unnecessarily arcane and convoluted way. (Research Assessment Exercise, 2009, p. 8)

#### *Engaging in the PBRF consultation process*

In assessing EPs, PBRF panel members make use of published guidelines, which are available to researchers in completing their portfolios. The guidelines, which provide statements of what is expected at each level of A, B, C, C (NE), R and R (NE), are both generic and panel specific and are reviewed for each quality assessment. Suggested changes are subject to community consultation through the TEC website before each quality assessment round. Hattie (2005) lamented that to date there has been little public debate about these criteria, and challenges educationalists to engage in powerful and focused debate about what should count as quality in educational research, what signifies peer esteem for research and to what extent researchers contribute to the wider research environment.

Two procedural problems in the evaluation process make it more difficult to engage with the nuanced and sophisticated processes advocated by Furlong and Oancea (2006). Both are related to the form of the Evidence Portfolio. In the British RAE, academics submit the four research outputs they consider their best during the assessment period, the last of which covered eight years. In New Zealand, on the other hand academics nominate their four most significant publications from the six-year qualifying period, but may list up to 30 more. This means that quantity may well be seen as an additional indicator of quality and a portfolio with only four publications may be seen as less meritorious than one with a much larger number. The other is that British panels are expected to make judgments about the work submitted and to read them all. In New Zealand, panel members read up to 50% of NROs and must rely on the reputation of the publication outlets as a quality proxy for at least some of their assessment. Not to do so, it is argued, would be to query the judgments of blind referees and an editorial panel.

Discussion about criteria could be fostered through the New Zealand Association for Research in Education (NZARE), with involvement from the University Deans of Education or nominees. Developing more nuanced and rigorous understandings of applied and practice-based research in education has a much wider application than the PBRF. Journal editors, website developers, examiners and thesis supervisors, research commissioners, as well as neophyte and experienced researchers would all benefit. Oancea and Furlong (2007) suggest the task is crucial but difficult. However, they believe that developing a quality checklist would not cover the complexity of applied educational research and that a broader philosophical

framework needs to be developed. It is important that academics engage with the TEC and with research leaders in their own institutions to debate the knotty issues of standards for evaluating applied and professional research.

### Conclusion

The article has argued that it is vital to increase capacity and achievement in educational research in New Zealand. There are many strategies in place to assist this process within institutions and the PBRF has sharpened awareness of its importance. At the same time official policy discourse has stressed the need for educationalists to engage in evidence-based practice and to develop an enquiry stance towards that practice. In a tertiary institution this will include research as well as teaching practice.

I have argued also that education academics need to engage in debate about the nature of quality and impact in the applied and practice-based research that forms a significant part of the research quantum in the field. To be effective, such research must be reported in ways that maintain rigour, show an awareness of wider thinking in the field, engage with the work of others yet remain firmly contextualized in a specific situation. We need to make this debate public, become aware of international thinking, take part in consultation by the TEC, and insist on the importance of theory in determining standards.

I end with a caveat. The PBRF has sharpened our awareness that as a professional group in New Zealand, education academics still rate poorly against other disciplinary fields, some of which are themselves largely applied and practice-based. We need to engage in measures to increase research capacity and raise the status of applied research. We do so not just to raise our standings but for a larger moral purpose which we share with educationalists around the world. We carry out enquiry in education for a range of purposes: to ask questions about the aims of education; to study our educational history and set it within a social and cultural context; to understand how human beings learn; to seek to theorise and improve our classroom practice based on evidence; to provide a considered basis for systemic change to provide greater equity; to evaluate reforms and review practice. What drives these enquiries is a search for greater understanding. We must not lose sight of the bigger picture or the spirit of genuine enquiry that drives us to learn more about our human condition.

### References

- Adams, J. (2008). *Strategic review of the Performance-Based Research Fund: The assessment process*. Wellington: Tertiary Education Commission.
- Alcorn, N., Bishop, R., Cardno, C., Crooks, T., Fairburn-Dunlop, P., Hattie, J., Jones, A., Kane, R., O'Brien, P., & Stevenson, J. (2004). Enhancing educational research in New Zealand: Experiences and recommendations from the Education PBRF panel. *New Zealand Journal of Education Studies*, 39(2), 275-303.
- Alcorn, N. (2007). Evidence and education: The braided roles of research, policy and practice. *New Zealand Annual Review of Education*, 16, 5-24.
- Bakker, L., Boston, J., Campbell, L., & Smyth, R. (Eds.). (2006). *Evaluating the Performance-Based Research Fund: Framing the debate*. Wellington: Institute of Policy Studies.
- Carr, W., & Kemmis, S. (1986). *Becoming critical: Education, knowledge and action research*. Waurun Ponds, Victoria: Deakin University.
- Elliott, J. (2001). Making evidence-based practice educational? *British Journal of Educational Research*, 27(5), 555-574.
- Expert Advisory Group. (2005). *Research Quality Framework: Assessing the quality and impact of research in Australia*. Canberra: Commonwealth of Australia.
- Furlong, J., & Oancea, A. (2006). Assessing quality in applied and practice-based research in education: A framework for discussion. *Review of Australian Research in Education: Counterpoints on the Quality and Impact of Educational Research. Special Issue 6 of Australian Educational Researcher*, 89-104.
- Gardner, H. (2004). *Changing Minds: The art and science of changing our own and other people's minds*. Boston, Mass.: Harvard Business School Press.
- Haigh, M. (2005). Outputs from research: Points for the PBRF? – potential for professional impact. In R. Smith & J. Jesson (Eds.), *Punishing the Discipline – the PBRF regime: Evaluating the position of education – where to from here?* (pp. 112-123). Auckland: AUT University & the University of Auckland.
- Hall, C., & Morris Matthews, K. (2006). Impact of the Performance-Based Research Fund on teaching and the research-teaching balance. In L. Bakker, J. Boston, L. Campbell & R. Smyth (Eds.),

- Evaluating the Performance-Based Research Fund: Framing the Debate* (pp. 419-476). Wellington: Institute of Policy Studies.
- Hammersley, M. (2004). Some questions about evidence-based practice in education. In T. Gary & R. Pring (Eds.), *Evidence-based practice in education* (pp. 133-149). Berkshire: Open University.
- Hargreaves, D. (1996). *Teaching as a research-based profession: Possibilities and prospects*. London: Teacher Training Agency.
- Hargreaves, D. (1997). In defence of research for evidence-based teaching: A rejoinder to Martyn Hammersley. *British Educational Research Journal*, 23(4), 405-420.
- Hattie, J. (2005). Foreword: The PBRF – its effects and some questions for consideration for educational research. In R. Smith & J. Jesson (Eds.), *Punishing the discipline – the PBRF Regime: Evaluating the position of education – where to from here?* Auckland: AUT University and the University of Auckland.
- Higgins, J. (2005). What counts as research under the PBRF environment? Considering theoretical-applied distinctions. In R. Smith & J. Jesson (Eds.), *Punishing the discipline – the PBRF regime: Evaluating the position of education – where to from here?* (pp. 101-108). Auckland: AUT University & the University of Auckland.
- Lambert, R. (2005). *Lambert Review of Business-University Collaboration: Final Report*. London: HM Treasury.
- Loughram, J. (2000). *International handbook of self-study of teaching and the educational practitioner*. Dordrecht, London: Kluwer Academic.
- MacLure, M. (2005). "Clarity bordering on stupidity": where's the quality in systematic review? *Journal of Educational Policy*, 20(4), 393-416.
- Middleton, S. (2005). One flew over the PBRF: Disciplining the subject of education. In R. Smith & J. Jesson (Eds.), *Punishing the Discipline – the PBRF regime: Evaluating the position of education – where to from here?* (pp. 27-35). Auckland: AUT University & the University of Auckland.
- Middleton, S. (2006). Researching identities: Impact of the Performance Based Research Fund. In L. Bakker, J. Boston, L. Campbell & R. Smyth (Eds.), *Evaluating the Performance-Based Research Fund: Framing the Debate* (pp. 477-500). Wellington: Institute of Policy Studies.

- Morton, M., & Gordon, L. (2005). What counts as research? Practitioner research as counter discourse. In R. Smith & J. Jesson (Eds.), *Punishing the Discipline – the PBRF regime: Evaluating the position of education – where to from here?* (pp. 95-100). Auckland: AUT University & the University of Auckland.
- Nussbaum, M. (1986) *The fragility of goodness. Luck and ethics in Greek tragedy and philosophy*. Cambridge: Cambridge University Press.
- Oancea, A. & Furlong, J. (2007). Expressions of excellence and the assessment of applied and practice-based research. *Research Papers in Education*, 22(2), 119-137.
- OECD (2002). *Frascati Manual: Proposed standard practice for surveys on research and experimental development*. Paris: OECD.
- Performance-Based Research Fund Working Group. (2002). *Investing in Excellence*. Wellington: Ministry of Education & Transition Tertiary Education Commission.
- Pring, R. (2004). Conclusion: Evidence-based policy and practice. In G. Thomas & R. Pring (Eds.), *Evidence-based practice in education* (pp. 201-212). Berkshire: Open University Press.
- Research Assessment Exercise. (2009). *RAE 2008 UOA 45. Subject overview report*. Education. Downloaded from: <[www.rae.ac.uk/pubs/2009/ov/](http://www.rae.ac.uk/pubs/2009/ov/)>.
- Roberts, G. (2003). *Review of research assessment*. Report by Sir Gareth Roberts to the UK funding bodies. London: UK funding bodies.
- Schon, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schon, D. (Ed.). (1991). *Reflective turn: Case studies in and on education*. New York: Teachers College Press.
- Slavin, R. (2003). A reader's guide to scientifically based research. *Educational Leadership*, 60(5), 12-16.
- Smith, R., & Jesson, J. (Eds.). (2005). *Punishing the Discipline – the PBRF regime: Evaluating the position of education – where to from here?* Auckland: AUT University and the University of Auckland.
- Stokes, D. (1997). *Pasteur's Quadrant: Basic Science and Technological Innovation*. Washington: Brookings Institute.
- Tertiary Education Commission. (2005). *Performance-Based Research Fund: Guidelines 2006*. Wellington: Tertiary Education Commission.

- Tertiary Education Commission. (2007). *Performance-Based Research Fund: Evaluating research excellence*. Wellington: Tertiary Education Commission.
- Tertiary Education Commission. (2008a). *Performance-Based Research Fund Sector Reference Group review: New and emerging researchers*. Wellington: Tertiary Education Commission.
- Tertiary Education Commission. (2008b). *Performance-Based Research Fund Sector Reference Group review: Applied and practice-based research*. Wellington: Tertiary Education Commission.
- Tertiary Education Commission. (2009). Retrieved on March 14, 2009 from: <[www.tec.govt.nz](http://www.tec.govt.nz)>.
- Tooley, J., & Darby, D. (1998). *Educational Research: An Ofsted critique*. London: Ofsted.
- Weiss, C. (1997). *Evaluation: Methods for studying programs and policies*. New Jersey: Prentice Hall.
- Woodhead, C. (1998, March 20). Academia gone to seed. *New Statesman*.
- Yates, L. (2004). *What does good education research look like?* Maidenhead and New York: Open University Press, McGraw-Hill Education.

### **The author**

Noeline Alcorn is Emeritus Professor of Education at the University of Waikato where she was formerly Dean of the School of Education. She now works part time in the Wilf Malcolm Institute for Educational Research at the university. She was a member of the Education Panel for the 2003 quality evaluation and chaired the panel for the 2006 round.