



‘All the Evidence Shows ...’: reasonable expectations of educational research

TONY EDWARDS

ABSTRACT *After years of being criticised in Britain as irrelevant to the task of raising educational standards, educational researchers have been invited to join in the improvement of policy and practice. Welcome though the invitation is, some views of how they should deserve influence reflect unrealistic or narrow definitions of really useful research. This paper argues against exaggerating the prospects for a science of teaching; against tying the national research agenda tightly to what are currently identified as ‘central issues’; and for a broader approach to relevance than direct applicability to improving educational practice.*

INTRODUCTION

Many years ago, John Furlong and I argued against a science of teaching constructed from the outside, and for more classroom research to be done by teachers (Edwards & Furlong, 1977). Despite its theme, the journal illustrated the article by showing a teacher looking for coins in front of a staffroom dispenser of tea, coffee, and research data. That cartoon might have been more a prediction than a joke if recent calls for education policy and practice to be informed by evidence had not been accompanied by such strident complaints about the persistent failure of researchers to provide evidence worth using. As demands on education systems have risen, so has criticism of the irrelevance of most research. There are countries, Australia and Singapore for example, where researchers have been actively involved in shaping policy. This has not been the case in Britain. Although the Hillage report (1998) regretted that researchers and policy-makers in this country had ‘drifted apart’, it is hard to see a time when they have been ‘together’—at least since the 1960s when, for example, the re-organising of secondary education was well informed by research which told the Labour Government what it was ready to hear about the social class bias of early selection and the consequent ‘waste’ of talent (Crosland, 1966).

Now researchers are being urged to exchange irrelevance for influence by contributing to ‘the improvement of government’. As the Secretary of State for Education and Employment phrased the invitation, and defined his own Government’s ‘undogmatic’ approach to policy-making:

We need to be able to rely on ... social scientists to tell us what works and why and what types of policy initiatives are likely to be most effective. And we need better ways of ensuring that those who want this information can get it easily and quickly. I commend this vision to you and invite all of you to work with us to achieve it. (Blunkett, 2000, closing words)

The invitation is irresistible. So in a climate where target-setting is pervasive and 'excuses' for failure are unacceptable, what realistic but challenging expectations should educational researchers try to meet? In this paper, I argue that the prospects for being 'really useful' will be improved if some prescriptions for how research should deserve influence on policy and practice are diluted, or at least taken with care. I argue that aspirations to a 'science of teaching', as that term would commonly be understood, are inappropriate; that the obligation to contribute to 'the real debates which affect people's life chances' (Blunkett, 2000, para. 2) extends beyond issues which policy-makers may currently define as 'central'; and that the usefulness of educational research cannot therefore be measured by what it has to say directly about 'what works' [1]. Illustrative references reflect my own research interests, which is also why they relate to schooling rather than to the much wider range of activities and settings which constitute education.

POTENTIAL IMPACT OF RESEARCH ON PRACTICE

Compared with the main charge of irrelevance, that of reporting research findings in inaccessible language may appear venial. But it is part of a larger complaint that educational researchers have written largely for one another about topics of interest largely to themselves, that they have preferred a hermit-like withdrawal 'from the messy world of short-term practical problems into intellectual obscurities masquerading as profundities whilst dreaming of ultimate recognition' (Hargreaves, 1999a, p. 243). In similar vein, Alan Smithers dismissed academic journals as 'not the right vehicles for saying the right things to the right people', and recommended instead the quick transmission of results 'in disposable form' to those too busy doing things to 'wait for research to catch up' (quoted in the *Times Educational Supplement*, 22 January 1999). It is reasonable to expect 'crisp summaries written in plain English' which emphasise the practical implications of research findings (Millett, 1997), but such publications are complementary not sufficient. Although researchers should bear in mind how their work may be selectively cited and interpreted, and not step aside after publishing it in esoteric journals as though their responsibilities are now fulfilled, 'jargon' is an indiscriminating term of abuse because what is deliberate or incompetent mystification in one context can be necessary technical clarification in another. Drawing out practical implications from research evidence is inseparable from establishing the dependability (or trustworthiness) of the evidence on which any conclusions are based. This is why writing for other researchers is integral to the research process, and why the overlap between 'academic' and 'user' reviews of evidence should be substantial [2]. Science by press release is not less objectionable in the social sciences, even though (as I will argue later) avoiding it is likely to highlight far-reaching differences between researchers and research-users about what counts as reliable and valid evidence, and about levels of confidence that an innovation is demonstrably 'working'.

These points may seem a complacent evasion of the main charge—that educational research has failed to have visible, significant and beneficial effects on classroom practice. There has been no educational equivalent to the process by which 'the investigation of the symptoms and causes of disease' leads to 'the development and application of therapeutic interventions which bridge the gap between the basic sciences and the practitioner' (Hargreaves, 1998a, p. 2). Without it, critics claim, teaching remains too largely a matter of teachers' personal experiences and preferences; being well-informed about research remains neither a professional requirement nor a promo-

tional advantage; politicians feel fully entitled to prescribe on professional matters; and there is no progression towards more effective teaching and learning.

So what might an equivalent to clinical research look like? It will not come through spectacular ground-breaking investigations, those 'single pieces of research' with direct, immediate effects on educational standards which Her Majesty's Chief Inspector has claimed to have looked for in vain over the past 30 years (*Sunday Telegraph*, 21 November 1999). Even in the familiar unfavourable comparison with medicine, accumulating evidence of 'what works' (and that some established methods do not work) first changes 'background' understanding before new (or newly well-grounded) knowledge is gradually incorporated in professional discourse and in a redefinition of good practice (Hargreaves, 1998a). It is a process of gradual accretion. Unfortunately, rather like the 'permeated' themes of the National Curriculum, it may be so hard to detect as to raise doubts about whether it has happened at all. Yet according to the Chief Inspector, in his Annual Report for 1999 (page 3), 'we' already 'know what constitutes good teaching' and researchers merely waste resources on complicating 'what ought to be straightforward' and 'obfuscating' the 'classroom realities that really matter'. It is unclear whether the 'we' is royal, or whether it refers to Ofsted's collective wisdom, or to the education profession minus irrelevant obfuscators. More contentious, because explicitly research-based, is the claim that a 'codified, scientifically established body of knowledge' is already available for direct application (Reynolds, 1998). From this perspective, the objective of 'an educational technology that will deliver failure-free schooling' is already within reach, embodied in the slogan of the High Reliability Schools Project which David Reynolds co-directs – 'Schools get it right first time, every time. Pupils succeed every time' (Project Newsletter, autumn 1998).

This seems a remarkably strong version of the Engineering Model of research, in which technical knowledge of what works should replace the present diversity of professional experience, personal preference, and pedagogic fashion. Such knowledge would presumably consist of generalisations about effective practice so robust that teachers would waste their own time trying to rediscover them and their students' time by doing anything else. Some theories, of course, have no pretensions to being robust in that sense. They offer ways of thinking about a researchable topic, to be assessed as more or less useful rather than as true or false. Bernstein's socio-linguistic codes are a notable example (Edwards, 1987). But critics of educational research want theory-as-end-product, sets of 'inter-related substantive statements' grounded in evidence which, while always open to further empirical testing, can provisionally be taken as 'known' (Mouzelis, 1995, pp. 1–3). Mouzelis goes on to argue, however, that in all the social sciences, the lack of context in theories of this kind 'invariably leads to substantive conclusions which are either trivial or wrong'. What kinds of non-trivial, workable, generalisations might educational research seek to produce?

As that task was defined by David Hargreaves in his 1996 Teacher Training Agency lecture, they would be 'conclusive demonstrations' that changing from one teaching method to another would bring 'a significant and enduring improvement in teaching and learning'. He has insisted since then that he was referring to probability, not certainty; that research can only 'inform' practice because it can never replace other knowledge which teachers bring to bear on practical problems; and that even the best research evidence is not available as 'fixed, universal relationships' between methods and outcomes, but as 'local, context-sensitive patterns' which have to be interpreted by practitioners within their particular working environments (Hargreaves, 1998a; 1998b). I agree with those qualifications. Their combined effect is to resemble Carol Fitz-

Gibbon's (1996) concept of 'distributed research', in which data about students' performance is fed back to schools to be interpreted in the light of their local knowledge and educational priorities before being used to re-examine what they do. There is certainly a considerable distance between 'conclusive demonstrations' of effective practice, and statements of what has demonstrably worked for whom, in what conditions, with what effects, and might be used or adapted by other practitioners in other specific circumstances (Hargreaves, 1998b, p. 50).

That second formulation seems not very different either from Michael Bassey's (1998) 'fuzzy generalisations'. Too much fuzziness of course is unlikely to tempt teachers to try out something new, which is why they need enough detail to assess the fit between their own situation and the research setting from which potentially useful findings have come, and why promising investigations need to be replicated in different sites so that fuzzy generalisations can be gradually sharpened (Bassey, 1999). But teachers are likely to find the notion of fuzziness more realistic than prescriptions of effective practice which depend on artificially tidied up, or 'controlled', or over-generalised, versions of what they know to be complex, fast-moving, considerably unpredictable, learning environments. And contrary to Reynolds' (1998) complaint that systematic research into effective teacher behaviours has gone on almost everywhere but here, there is substantial British evidence about most items on his list of what 'we need to know', even though many of those involved in producing it would not want to claim that they had been investigating 'teaching technologies'. For example, a review of research in mathematics education includes among its list of evidence-based conclusions that although the level of pupil attainment '*can* increase' if teachers wait longer for answers to their questions, questioning '*may* not always be the most effective way to generate discussion' (Askew & Wiliam, 1995, pp. 16–17; italics added). The tentativeness is deliberate, but it may be too cautious. Classroom 'discussion' in which pupils' ideas are actively explored is rare because the 'normal' teaching exchange (teacher question–student response–teacher evaluation) is so hard to escape. The evidence for its pervasiveness is extensive, ranges across age-groups and subjects, and can be applied in quite specific suggestions about what teachers might do instead of asking questions on those occasions when they want something closer to dialogue than to interrogation (Edwards & Westgate, 1994, pp. 124–133).

Generalising beyond the settings investigated depends on identifying necessary or favourable conditions for something to work elsewhere. But although this requires some consistency in relevant findings, 'the facts' are unlikely to point only in one direction. Sherlock Holmes' warning that 'the temptation to form premature theories upon insufficient evidence is the bane of our profession' can apply to researchers eager to get their findings noticed, but the temptation appears more often to afflict policy-makers whose usual disregard of research is interrupted by drawing conveniently selective conclusions from what it has 'shown'. For example, an Ofsted report on *Setting in Primary Schools* (1999) attributed 'spectacular improvements' in mathematics to that form of pupil grouping. The evidence for the assertion was promptly described by the head of the National Numeracy Programme as 'flimsy and inconclusive', and the fact of improved standards as open to very different explanations (Anita Straker, reported in the *Times Educational Supplement* 16 April, 1999). A wide-ranging review of relevant evidence concluded that 'it is the provision of differentiated learning experiences which is important', and that ability grouping *within* classes was certainly a way of providing them. The lack of evidence confirming the effectiveness of streaming or setting *between* classes, however, was attributed not to deficiencies in the research, but to the presence

of so many other potentially significant variables within schools and to the difficulty of grouping pupils in ways which do not 'reinforce the social divisions found outside the school' (Harlen & Malcolm, 1999, pp. 54–57).

Treating pedagogic research as a matter of identifying reliably effective teacher behaviours is especially objectionable when these are detached from the kinds of learning they are intended to promote. Disputes about educational purpose, and changes over time in which purposes have priority, make the notion of accumulating evidence of effective practice much more problematic than it is sometimes taken to be. Thus the Chief Inspector's claim to know 'what constitutes good teaching' raises questions not only about where his knowledge comes from, but also about the kinds of learning his prescriptions are intended to promote. They are questions which illustrate the artificiality of separating practice from policy. For example, as Paul Black (1995) has argued, a sequence of generalisations began from the disputable 'fact' of falling educational standards, proceeded to the empirically dubious explanation that progressive teaching methods were to blame, and reached the empirically dubious conclusion that a return to traditional methods would 'therefore' bring certain improvement. That diagnosis contributed hugely to the derision directed at educational theory and research for undermining standards. From another ideological direction and with far-reaching practical effects, the neo-liberal reforms of 1988 and 1992 were designed to subject schools so strongly to the discipline of market forces that their practices would be irresistibly constrained, and thereby improved, by having to respond and be accountable to consumer demand.

THE POTENTIAL IMPACT OF RESEARCH ON POLICY

Despite the familiar charge of irrelevance, it is difficult to see recent educational research as largely detached from issues 'central and directly relevant to the political and policy debate' (Blunkett, 2000, para. 7). Although the Hillage report (1998) concluded that funding had been 'insufficiently weighted' in favour of empirical research directly applicable to educational policy and practice, it also acknowledged that a preoccupation with evaluating successive government reforms had produced a great deal of work which was short-term, atheoretical and parochial. Complaints from elsewhere that much of this research had been barely disguised political opposition certainly raises issues of objectivity, but is more easily reconciled with the charge of being too responsive to, than of ignoring, the 'real world' (Bridges, 1998). For example, 'successive Conservative governments in the UK announced that they were going to marketise education, and research took up that statement of intent and documented how it was being done, and how it was and was not working' (Seddon, 1997, pp. 171–2). As a result, Seddon argues, policy researchers were themselves 'captured' by a politically dominant discourse, narrowed their horizons accordingly, and tended to over-generalise Thatcherism as the 'defining expression of international neo-liberal reform' (p. 167). It is also true that the claims of market enthusiasts were mainly investigated by researchers disposed to disbelieve them. But 'documenting' the effects of market reforms would not have been commissioned or even encouraged by Conservative Governments which appeared to regard almost all educational research as inherently subversive.

The climate is now very different. The merged Department for Education and Employment (DfEE), already inclined towards the Employment tradition of taking research seriously, responded to the Hillage report by declaring its commitment to

policy-making informed by 'publicly available research evidence', and its intention to establish centres dedicated to providing a 'readily available source of advice and expertise ... closely related to our key policy information needs' (DfEE, 1998). The Secretary of State's recent invitation to social scientists to meet those needs included enthusiastic references to the new Centre for Evidence-informed Policy and Practice (based at the Institute of Education), to following the example of the Cochrane Collaboration in Medicine by establishing a data base of useful evidence, and to his regret that an 'anti-intellectual seam' running through government 'at the political level and among officials' had contributed to making research 'too inward looking' (Blunkett, 2000, paras. 7 & 35). Yet to the extent that he assumed an eventual consensus about what research needed to be done which took his Government's priorities as setting the agenda, he was evading fundamental differences in the frames of reference of researchers and policy-makers.

When Michael Barber, head of the DfEE's newly-created Standards and Effectiveness Unit, told the 1997 conference of the British Educational Research Association (BERA) that he hoped for 'a dialogue with researchers' about the new Government's approach to raising standards, he added the proviso that while constructive criticism was welcome, negative comments 'arising from cynicism' were not. His stipulation may have confused cynicism and scepticism, or expressed an understandable dislike of carping from the sidelines at a Government already remarkably busy setting new achievement targets and advising on how to reach them. The implication however was that researchers are there mainly to serve government purposes. David Blunkett was careful to balance his own criticism of research 'driven by ideology paraded as intellectual inquiry or critique' and by the sole aim of proving government policy wrong, by recognising an obligation on Government to 'give serious consideration to "difficult" findings' which questioned a policy's success. The risk nevertheless is that a policy-maker's view of really useful research will resemble Margaret Thatcher's view of a really useful Minister (embodied at that time by David Young at the Department of Employment)—that it takes the problems which government brings to it, contributes to their solution, and refrains from adding new problems or further complicating old ones.

From that perspective, research is most likely to be noticed when it is 'helpful'—that is, when its findings broadly support and may even improve what government is already determined on other grounds to do. It is understandable that the political capital invested in major initiatives makes it tempting to declare them a success before they have been evaluated, even when (like the Technical and Vocational Education Initiative) they were first introduced in explicitly pilot form (Bell & Raffe, 1991). The Assisted Places Scheme, for example, was designed to demonstrate the incoming Thatcher Government's commitment to traditional academic standards and a traditionally selective ladder of educational opportunity, and later re-interpreted as a significant first step towards a school system founded on 'choice and diversity'. These were political acts of faith. Although competing predictions about the beneficiaries and wider effects of assisted places could be empirically 'tested', successive Conservative Governments considerably extended the Scheme without making or responding to any evaluation of its value-for-public-money (Edwards *et al*, 1989). Summative assessment of student performance in the National Curriculum is a prime example of politically driven over-simplification (Daugherty, 1995). Yet it is a firmly research-based conclusion that no other way of raising educational standards matches formative assessment systematically and sensitively carried out (Black & Wiliam, 1998). That review of evidence even anticipated David Blunkett's request for indications of the relative

effectiveness of different educational interventions by estimating the learning gains thereby produced (which were greatest for lower-achieving pupils), and the effects on England's position in international league tables of pupil performance in mathematics and science (from the middle to the top five). It was also careful to specify factors most likely to facilitate or inhibit those beneficial effects, and it presented an action plan for bringing 'demonstrable improvements in the quality of classroom learning'. It will be interesting to trace the influence of those evidence-based conclusions on a continuing political reluctance to trust teachers, and a continuing political preference for giving educational consumers the apparently hard information which summative testing provides.

Of course policy cannot be read off from what research 'has shown', nor can policy decisions be delayed until sufficient evidence is available. There are too many other pressures, interests, priorities and commitments to take into account, and there are value-positions to be sustained and displayed which are not open to empirical verification. The examples just cited are not of evidence-informed policy, but of the 'heroic model' in which Ministers 'who know what they want, set out to get it without recourse to supporting or opposing evidence' (Kogan, 1999, p. 11). This was why quasi-markets were created, in this country and elsewhere, as a matter of ideological conviction (Whitty & Edwards, 1998). For a Labour Government which does not share the conviction, but which inherited some consequent innovations which it appears eager to extend, there is surely an obligation to explore systematically for whom, in what conditions, at whose expense, 'choice works' and private provision or public-private partnerships create enhanced opportunities. In the more open policy-making climate which David Blunkett has described, this should be a key area where researchers are encouraged to 'speak truth to power' (Coffield, 1999a). It is not that theirs is a superior 'truth', rather that they bring a different kind of knowledge which is diminished by being tightly confined within the bounds of what is currently taken to be politically practical and politically desirable.

In the normative world in which researchers traditionally worked, 'control should rest with the researcher' because their independence is a necessary condition for advances in knowledge. In the normative world of the policy-maker, however, the direction and even at times the outcomes of research may quite properly be controlled 'in the public interest' (Bell & Raffe, 1991, pp. 134-8). Of course, researchers may be so far removed from what is possible that their findings remains irrelevant. Hargreaves (1999a, p. 240) cites Robert Lynd's (1939) distinction between research 'scholars' at risk of being disengaged from real social problems, and research 'technicians' at risk of merely providing managerial solutions to whatever problems policy-makers choose to bring them. But while Lynd's defence of applied social science rejected any self-indulgence in perpetual criticism, he also insisted that social scientists had a duty to be 'troublesome, to disconcert the habitual arrangements by which we manage to live along, and to demonstrate the possibility of change' (1939, p. 181). This is not a role likely to be welcomed by policy-makers. It is sometimes the case, as Alison Wolf has argued in relation to her research on General National Vocational Qualifications (GNVQs), that researchers work more constructively (because more 'realistically') from inside a policy programme and that the investigation of practical problems may provide the most fertile opportunities for basic research (cited in Ecclestone, 1998, p. 692). But to the extent that they are then incorporated within the 'official' view, their contribution may lose its distinctiveness. Reynolds (1999) recognised that danger in relation to the 'central influence' on government policy which he claimed for research into school

effectiveness, in particular that researchers might be drawn away from those 'wider' explanations for educational failure to which I turn in the final part of the paper. The GNVQ example is interesting because although Wolf's own publicly-funded research certainly 'troubled' then dominant policy assumptions about (for example) core skills, the official research programme supporting the development of GNVQs inhibited any fundamental questioning of the competence-based assessment to which policy was then committed (Ecclestone, 1998). Nor was notice apparently taken of research into those obstacles to any parity of esteem for vocational qualifications which are entrenched in the traditional prestige and market-value of A levels (Edwards, 1997).

Implicit in the preceding argument is how the funding of educational research should be balanced between research initiated from within a field of inquiry with the traditional academic aim of 'advancing knowledge' with little overt reference to prospective usefulness, and research which is either commissioned for specific practical purposes or selected in relation to thematic priorities identified as being in the national interest. Amid criticism that there has been far too much of the former, it is worth noting again the high proportion of recent research which has been responsive to demand (Bridges, 1998; Hammersley, 1998). An 'interactive balance between creative autonomy and the public agenda' (Amann, 1997) is surely preferable to a large-scale conscription of education researchers seeking grants. In this context, the National Education Research Forum was advocated as the means of breaking into an 'exclusive club where an élite decides which members shall have the money to do what' (Hargreaves, 1998, p. 15; 1999b), a description of recent times past which many grant-seekers will not easily recognise. As the Forum takes on its tasks of identifying serious gaps in what we 'need to know' and encouraging research to fill them, how active should it be in drawing funds towards those priorities? Will it be the setting for dialogue between the different interests represented, or will it drift (or even move purposefully) towards becoming a National Research Board setting an extensive national agenda of what needs to be done? Even after ten years a similar forum in Scotland was still exerting pressure for 'quick, direct, simple answers' to immediately urgent questions while apparently failing to reach any agreement about longer-term research objectives (Brown & Harlen, 1998). David Blunkett's insistence that 'to have practical influence, conclusions from research must be *realistic and achievable*' (2000, para. 49; original emphasis) reflects an understandable impatience with findings which take no account of financial or political constraints or of contending pressures on government. Yet I have argued that for researchers to anticipate deferentially what is currently practical politics may contradict their obligation to 'disturb'—in his words, to question 'fundamental assumptions and orthodoxies' in ways which may have 'big policy effects much further down the road'. In the final section of the paper, I argue that defining useful research too narrowly to discovering 'what works' works against both a longer view and the 'bigger' effects.

RESEARCH SHOULD BE DIRECTED TO WHERE POLICY AND PRACTICE CAN BE IMPROVED

It may be salutary to remind researchers, sometimes inclined to investigate everywhere but there, how much difference the teacher in the classroom can make to pupils (Millett, 1997). It is reasonable for user-reviews of research to 'focus on issues that directly influence pupils' achievements ... and matters that teachers might wish to address as part of their work' (Gillborn & Gipps, 1996, p. 7), and even to exclude evidence which is contradictory or entirely inconclusive on the grounds that 'unless

research can change what actually happens in classrooms, then it is very little use to the busy classroom practitioner' (Askew & Wiliam, 1995, p. 1). My objection is not to focusing on what 'can be done' by those to whom a research report is primarily addressed, but to measuring usefulness largely by that criterion. In particular, if research into school effectiveness 'resonates with much of the contemporary educational discourse that is associated with New Labour' (Reynolds, 1999, p. 66), it is because it supports confidence that school-level improvements and more effective teaching technologies can overcome 'outside' influences on pupil attainment.

In this politically and methodologically contentious context, *Fifteen Thousand Hours* was a rare instance of educational research which was immediately cited and acted upon because its conclusion that schools 'make a difference' challenged the then dominant view that they merely reproduced existing inequalities (Rutter *et al.*, 1979). There was concern, shared by Michael Rutter and his colleagues, that identifying characteristics of more effective schools should not be turned into simple recipes. And there were wide-ranging arguments about whether research which concentrated on within-school factors might not be over-emphasising their significance, just as previous research into 'external' determinants of educational achievement had appeared to demonstrate that schools did not matter by failing to investigate differences between them. Opposing positions in that debate are greatly over-simplified by being presented in stark terms of either-or, as they were when David Blunkett attacked 'cynics who say that school performance is all about socio-economics and the areas that these schools are located in' and who identify children as 'pre-ordained to fail by class, or gender, or by ethnic group, or by their home life' (speech at a National Union of Teachers conference, quoted in the *Times Educational Supplement* 3 March, 2000). I know no social scientific research claiming that level of determinism, even when it describes the heavy odds against escaping from social disadvantage (Bourdieu *et al.*, 1999), or which would use aggregated data to predict the fate of individuals (Power *et al.*, 1999). But if research would fail teachers by concentrating largely on 'outside-school determinants about which we can do little' (Reynolds, 1998, p. 26), it would also fail them by endorsing the view that more effective schooling can do almost 'everything'.

Even the Chief Inspector of Schools appeared to concede this when, amid well-practised jibes about 'complex and abstruse macro-level explanations' far removed from practice, he recalled a time when sociologists of education used to 'engage with serious issues in a humane and accessible way'. They did so, he suggested, when their 'classical terrain' extended far beyond 'schools as social systems' to 'social class and educability', and thereby provided salutary warnings against believing that 'the latest managerial interventions will suddenly and miraculously transform the realities that define classroom life' (Woodhead, 1998).

That major sources of 'definition' lie outside schools is why a political refusal to accept excuses for educational failure should not lead to 'designing out' those patterns of social disadvantage which do much to explain it (Gibson & Ashani, 1998), or to treating a few schools' exceptional success as sufficient evidence that structural inequalities can be overcome (Mortimore & Whitty, 1997), or to making the obligation to become and remain employable in a fast-changing labour market an almost entirely individual responsibility (Coffield, 1999b). This is why I believe that comparisons with medical research should extend beyond the familiar references to clinical sciences, in particular to randomised controlled trials which have few obvious educational equivalents unless advocates of the neuro-sciences as the way ahead for educational research are proved correct. Epidemiological research into the conditions associated with high

incidence of particular illnesses, and with general ill-health, is much closer to the terrain apparently admired by Woodhead. This was exemplified in the political arithmetic tradition of research into the inter-relationships of social origins and social destinations as a preliminary to political reform (Halsey *et al.*, 1997, pp. 37–38). Continuing research into the ‘distribution of life chances’ (a ‘classic’ term from the 1960s) has highlighted the increasing costs to individuals of leaving school with few or no qualifications, non-participation in education and training between the ages of 16 and 18 being closely associated with being unqualified, unemployed, and poor, ten years later (Bynner, 1999; Glennerster, 1999; Social Exclusion Unit, 1999).

The evidence strongly supports the Labour Government’s emphasis on raising educational standards ‘for all’. But it has also shown persistently high correlations between social disadvantage and low attainment in conditions in which child poverty has risen three-fold since the late 1970s and in which the high concentration of children from the poorest families in particular areas and schools has increased the impact of socio-economic factors on *school*-level performance. From this perspective, government action to alleviate child poverty would do more to raise educational attainment than setting targets and holding schools almost entirely accountable for achieving them (Whitty *et al.*, 1998). Indeed, government recognition of the inter-locking effects of low income, health and housing is evident in the creation of Education Action Zones and the work of the Social Exclusion Unit. In research terms, those ‘bigger’ effects make unavoidably complex macro-level explanations for educational failure a necessary complement to more immediately practical research into more effective schooling. And the search for those explanations should not be contained within currently authorised versions of what the ‘central issues’ are. For example, research into who gains and who loses from the privatising of erstwhile public provision, or from the freedom allowed to popular schools to manage their intakes, may well produce findings which are ‘difficult’ to reconcile with continuing government assumptions about the benefits of diversity and choice. Useful research into ‘what works and why’ therefore extends beyond the evaluation of specific policies and practices to that wider function—defined by David Blunkett as contributing to ‘a coherent picture of how society works’ and of what is within the scope of government to improve—which I interpret in terms of the endemic social scientific dilemma of giving due weight to structural opportunities and constraints without making the actions of individuals appear either over-determined or unrealistically free.

NOTES

- [1] Earlier versions of this paper were given in 1998 at a conference on Educational Research in Wales, a Research Open Day in the Sheffield University Department of Educational Studies, and this year as part of a seminar series organised around contributions to this Special Issue by the Oxford University Department of Educational Studies. Although the title survives, the content has changed considerably in response to discussions on those occasions, to developments since 1998 (in particular David Blunkett’s recent appeal for useful research), and to helpful comments on a late draft from Michael Bassey, Bruce Carrington and Frank Coffield. I am also grateful to Martyn Hammersley for the stimulus to revising my own arguments which came from his unpublished paper, ‘Why research into practice does not go: some questions about the enlightenment function of social and educational enquiry’.

- [2] This wider public is recognised in the Education Panel's criteria for the next Research Assessment Exercise, which include a requirement to make explicit for 'each piece of research output ... its prime audience and educational significance'. The importance of addressing different audiences appropriately is also emphasised in the British Educational Research Association's guidelines on 'good practice in educational research writing' (Spring, 2000).

REFERENCES

- AMANN, R. (1997) A society which understands itself, *Social Sciences*, 36.
- ASKEW, M. & WILIAM, D. (1995) *Recent Research in Mathematics Education* 5–16 (London, OFSTED).
- BASSEY, M. (1998) Fuzzy generalisations and professional discourse, *Research Intelligence* 63, pp. 20–24.
- BASSEY, M. (1999) *Case-study Research in Educational Settings* (Buckingham & Philadelphia, Open University Press).
- BELL, C. & RAFFE, D. (1991) Working together? Research, policy and practice: the experience of the Scottish evaluation of TVEI, in: G. WALFORD (Ed.), *Doing Educational Research* (London, Routledge).
- BLACK, P. (1995) *Ideology, Evidence and the Raising of Standards*, annual education lecture, King's College London.
- BLACK, P. & WILIAM, D. (1998) Assessment and classroom learning, *Assessment in Education* 5, pp. 7–74.
- BLUNKETT, D. (2000) *Influence or Irrelevance: can Social Science improve Government?* (London, Department for Education and Employment); reprinted in *Research Intelligence* No. 71 (March, 2000), pp. 12–21.
- BOURDIEU, P. *et al* (1999) *The Weight of the World: social suffering in contemporary society* (Cambridge, Polity Press).
- BRIDGES, D. (1998) Research for sale: moral market or moral maze? *British Educational Research Journal* 24, pp. 593–608.
- BROWN, M. (1998) The tyranny of the international horse race, in: R. SLEE & G. WEINER, with S. TOMLINSON (Eds) *School Effectiveness for Whom?* (London, Falmer Press).
- BROWN, S. & HARLEN, W. (1998) A forum for researchers and users of research—Scottish style, in: J. RUDDUCK & D. MCINTYRE (Eds) *Challenges for Educational Research* (London, Paul Chapman).
- BYNNER, J. (1999) New routes to employment, integration and exclusion, in W. HEINZ (Ed.) *From Education to Work: cross-national perspectives* (Cambridge, Cambridge University Press).
- COFFIELD, F. (1999a) Past failures, present difficulties and possible futures for research, policy and practice, in: F. COFFIELD (Ed.) *Speaking truth to power: research and policy on lifelong learning* (Bristol, Policy Press).
- COFFIELD, F. (1999b) Breaking the consensus: lifelong learning as social control, *British Educational Research Journal* 25, pp. 479–99.
- CROSLAND, A. (1966) Comprehensive education (speech to the 1966 North of England Education Conference), in: *Socialism Now and Other Essays* (London, Cape, 1974).
- DAUGHERTY, R. (1995) *National Curriculum and its Assessment: a review of policy 1987–1994* (London, Falmer Press).

- DEPARTMENT FOR EDUCATION AND EMPLOYMENT (1998) *Proposed Action Plan in Response to the Report 'Excellence in Research on Schools'* (London, HMSO).
- ECCLESTONE, K. (1998) Euston Road and the ivory towers: the impact of the GNVQ research industry, *Journal of Education Policy* 13, pp. 679–97.
- EDWARDS, T. (1987) Language codes and classroom practice, *Oxford Review of Education* 13, pp. 237–47.
- EDWARDS, T. (1997) Educating leaders and training followers, in: T. EDWARDS, C. FITZ-GIBBON, F. HARDMAN, R. HAYWOOD & N. MEAGHER, *Separate but Equal? A levels and GNVQs*. (London, Routledge).
- EDWARDS, T. & FURLONG, J. (1977) Time to go inside? *Times Educational Supplement* 3 June.
- EDWARDS, T., FITZ, J. & WHITTY, G. (1989) *The State and Private Education: an evaluation of the Assisted Places Scheme* (Lewes, Falmer Press).
- EDWARDS, T. & WESTGATE, D. (1994) *Investigating Classroom Talk* (2nd edition) (London, Falmer Press).
- FITZ-GIBBON, C. (1996), *Monitoring Education: Indicators, Quality and Effectiveness*, (London, Cassell).
- GIBSON, A. & ASHANI, S. (1998) School performance, school effectiveness and the 1997 White Paper, *Oxford Review of Education* 24, pp. 195–210.
- GILLBORN, D. & GIPPS, C. (1996) *Recent Research on the Achievements of Ethnic Minority Pupils* (London, HMSO).
- GLENNERSTER, H. (1999) *Tackling Poverty at its Roots: Education* (London School of Economics, ESRC Centre for the Analysis of Social Exclusion).
- HALSEY, A., LAUDER, H., BROWN, P. & WELLS, A. (Eds) (1997) *Education: Culture, Economy, Society* (Oxford, Oxford University Press).
- HAMMERSLEY, M. (1998) The profession of a methodological purist, in: G. WALFORD (Ed.) *Doing Research about Education* (London, Falmer Press).
- HARGREAVES, D. (1998a) The production, mediation and use of professional knowledge among teachers and doctors: a comparative analysis (Paris, high level seminar, OECD Centre for Educational Research and Innovation).
- HARGREAVES, D. (1998b) Improving research to enhance teaching, *Professional Development Today*, April, pp. 7–55.
- HARGREAVES, D. (1999a) Revitalising educational research: lessons from the past and proposals for the future, *Cambridge Journal of Education* 29, pp. 239–249.
- HARGREAVES, D. (1999b) The knowledge-creating school, *British Journal of Educational Studies* 47, pp. 122–144.
- HARLEN, W. & MALCOLM, H. (1999) *Setting and Streaming* (Edinburgh, Scottish Council for Research in Education).
- HILLAGE, R., PEARSON, R. ANDERSON, A. & TAMKIN, P. (1998) *Excellence in Research on Schools* [Institute of Employment Studies, Southampton University] (London, HMSO).
- KOGAN, M. (1999) The impact of research on policy, in F. COFFIELD (Ed.) *Speaking Truth to Power: Research and Policy on Lifelong Learning* (Bristol, Policy Press).
- LYND, R. (1939) *Knowledge for What?* (Princeton, Princeton University Press).
- MILLETT, A. (1997) Speech to the Teacher Training Agency Research Conference (December).
- MORTIMORE, P. & WHITTY, G. (1997) *Can School Improvement Overcome the Effects of Disadvantage?* (London, Institute of Education).
- MOUZELIS, N. (1995) *Sociological Theory: What Went Wrong?* (London, Routledge).

- POWER, S., WHITTY, G., EDWARDS, T. & WIGFALL, V. (1999) Destined for success? Educational biographies of academically able pupils, *Research Papers in Education* 14, pp. 321–339.
- REYNOLDS, D. (1998) Teacher effectiveness: better teachers, better schools, *Research Intelligence* 66, pp. 26–29.
- REYNOLDS, D. (1999) School effectiveness, school improvement and contemporary education policies, in: J. DEMAINE (Ed.) *Education Policy and Contemporary Politics* (London, Macmillan).
- RUTTER, M., MAGHAN, B., MORTIMORE, P. & OUSTON, J. (1979) *Fifteen Thousand Hours* (London, Open Books).
- SEDDON, T. (1997) Markets and the English: rethinking educational restructuring as institutional design, *British Journal of Sociology of Education* 18, pp. 165–186.
- SOCIAL EXCLUSION UNIT (1999) *Bridging the Gap: New Opportunities for 16–18 Year Olds not in Education, Employment or Training* (London, Stationery Office).
- WHITTY, G. & EDWARDS, T. (1998) School choice policies in England and the United States: an exploration of their origins and significance, *Comparative Education* 34, pp. 21–27.
- WHITTY, G., AGGLETON, P. GAMARNIKOW, E. & TYRER, P. (1998) Education and health inequalities, *Journal of Education Policy* 13, pp. 641–52.
- WOODHEAD, C. (1998) 'Academia gone to seed', *New Statesman* 20 March.

Correspondence: Professor Tony Edwards, 2 Hollin Hill Terrace, Riding Mill, Northumberland NE44 6HR.