

How can we evaluate Educational research?

By

Harvey Goldstein

Institute of Education

h.goldstein@ioe.ac.uk

OFSTED recently has published a survey of educational research (Tooley and Darby, 1998) which appears against a background of political concern about the quality of such research and the relationship between research and policy. This debate can be traced to an article by Hargreaves (1996) for the Teacher Training Agency, with subsequent responses from several researchers (Gray et al., 1997) and a further review commissioned in 1998 by the DfEE.

In a preface to the OFSTED survey HMCI Chris Woodhead says 'much that is published is, on this analysis, at best no more than an irrelevance and a distraction'. As this remark implies, the survey is critical of the material it investigated and the authors invite others to carry the debate forward; the following remarks are intended to do just that.

While much of what I have to say will be critical of this survey, there are some important respects in which I believe that it does advance debate. First of all it sets out its aims clearly - namely to evaluate the quality of a sample of articles published in four British journals devoted to education. The judgemental criteria are set out in detail and the data used to form judgements are clearly identified. The report is also careful to state various caveats about the 'modesty' of its contribution, and particularly that evaluating people's reports of their research may not be a good guide to the quality of the research itself (P. 11). The authors also say that it is not intended that generalisations about educational research can be drawn from their work (P. 27) - a reservation which, to judge from the above quotation, has clearly escaped the notice of HMCI Woodhead. The authors' concern to provide reliable statistical data about quality is welcome, especially since the original article by Hargreaves and some of the subsequent debate is rather short on providing representative evidential support for assertions that are made.

It may be helpful if I reveal some of my personal views about research quality. From my experience of refereeing papers, grant applications and as a journal editor, I do perceive a widespread lack of understandings, including basic concepts of sampling and logical inference. On the other hand this is not confined to educational research. I see it in medicine and other disciplines, and of course it permeates general public and political debates. Indeed, policy makers themselves typically appear to have scant understanding of, or regard for, some basic canons of rational debate. Since current discussions of educational research have strong political connections this makes it difficult to arrive at a balanced view.

To return to the OFSTED survey, which concludes that only a third of the papers reviewed satisfy 'good practice'. A key issue is the decision to study just four British journals, using two evaluators. This raises several problems. First, much educational research is published in journals which are not solely concerned with education: such

research appears in mainstream psychology, sociology, statistical etc. journals. It may even be the case that the 'best' published research appears in such journals. Secondly, educational research from the UK gets published in international journals, but these were excluded from the survey. Again, it may be the case that the best educational research goes into such journals. Thirdly, a better design would have been to select more journals and fewer articles from each, although with the resources available for the exercise any sample would be very limited. To carry out a satisfactory review of published papers requires a much larger sample of journals and a more representative set of evaluators: as careful as they might be, no two researchers can hope to be expert enough adequately to review the variety of educational publications.

The authors (P. 77) are critical of the peer review process which led to the publications they evaluate. In effect they are claiming that in the majority of cases their own judgements are superior to those of the referees used by these journals. This is indeed a strong claim and it does seem reasonable to ask the wider world to suspend judgement until the research community itself (of which the authors are a part) has had some time to evaluate this claim.

The authors state that the journals reviewed 'represent an important strand of academic educational research' (P. 27). As already pointed out, this statement is contestable because it begs the question of the proper way to make judgements of the corpus of educational research. Such research is produced by educational researchers. If we wish to study its quality then we should be sampling the researchers, research teams and institutions carrying out that research. The sampling frame would include individual academics, research teams, survey organisations, institutions such as OFSTED etc. From such sources we can study the activities of the various participants, what kinds of research they engage in, how much it costs, who funds it, how it is communicated, critical reactions to it, as well as an evaluation of its qualities. Such an approach has the advantage that it allows comparisons, e.g. between research funded by DfEE, ESRC and others. It would be able, in principle, to capture the totality of educational research in a way that the Tooley/Darby survey is unable to do. Certainly, if we wish to have a sensible debate about educational research funding we require information of this kind: we cannot make reliable inferences about the effects of funding from a survey such as the present one, despite the authors' attempts to do so (P.78).

The comparative issue is also important for educational research as a whole. Whatever conclusions we may come to about educational research, it needs to be evaluated against the situation in other fields of study - medicine, economics or whatever. This suggests a much wider perspective than currently envisaged.

A further key issue when debating research output is to examine how far its extent and quality is influenced by external factors. The authors briefly address this in terms of the universities Research Assessment Exercise (RAE) (P. 78-79) and suggest how the RAE may have had deleterious effects on research quality. A more complete analysis would study other aspects such as the requirements of funding bodies, the attitudes of the media and overt and covert political pressures. To evaluate quality without attempting to understand the relationship between quality and such other factors seems to miss the point. Thus, for example, Hargreaves (1996) fails to acknowledge how influential such

factors may be and by so doing implicitly lays the 'blame' for what he sees as poor quality research entirely at the doors of the researchers themselves.

In this brief review of the OFSTED survey I have made no attempt to enter into a dispute with Tooley and Darby over the judgements they make about individual articles. The authors of those articles, if they feel misrepresented, will be perfectly capable of responding themselves. Indeed, it would be a constructive initiative if OFSTED were to organise a symposium to which the authors of the articles were invited to make a contribution. My aim, rather, has been to reflect upon the usefulness of this kind of survey and to offer some suggestions about how a more valid study of educational research might be constructed.

For all its flaws and the severe caveats surrounding its conclusions, the present survey has raised some useful issues. It will be interesting to see how far the DfEE study of educational research can provide a more complete picture, and especially how far it is prepared to follow the suggestions outlined above.

This paper is available on the Web at www.ioe.ac.uk/hgoldstn/

References

- Gray, J. and others (1997). Educational research and evidence based practice: the debate continues. *Research Intelligence*, February 1997.
- Hargreaves, D. H. (1996). *Teaching as a research based profession: possibilities and prospects*. London, Teacher Training Agency.
- Tooley, J. and Darby, D. (1998). *Educational research - a critique*. London, Office for Standards in Education.