

Improving assessment: a response to the BERA Policy Task Group's report on assessment

HARVEY GOLDSTEIN

Institute of Education, London

THE CONTEXT

The BERA Policy Task Group report has been produced when central government is increasing its control over education policy, weakening local accountability and systematically disregarding established canons of rational debate and decision-making in the area of educational policy.

It is appropriate that the body representing the educational research community makes its views known, since government policies threaten to undermine the ability of that community properly to pursue its work. The report identifies the failures of current government policies, and counters them with its own set of recommendations.

While I agree with many of the recommendations, I do have certain reservations, as follows.

A RESURRECTED ASSESSMENT OF PERFORMANCE UNIT (APU)

In principle I agree that a survey using light sampling, carried out regularly would provide useful information about education. The report, however, advocates using a 'bank' of items for constructing these assessments and claims that feedback from a new APU 'would enable schools to interpret the achievements of their own pupils in terms of national standards'.

First of all, an 'open' bank of assessment items, nationally calibrated, which schools could use to judge their pupils and which could form the basis for APU surveys, simply will not work. For reasons discussed elsewhere in detail (Goldstein and Wood, 1989), it is not possible to have an open choice item bank which also conveys unbiased information on national performance levels, whether for feedback to teachers or for incorporation into a national survey. Second, the APU as it existed was a highly inadequate vehicle for providing understanding of national performance standards because it collected very little information on the students, their teachers and their schools. Nor did it collect longitudinal information. Only if it could be properly contextualized using such information would a resurrected APU be worth having.

RECORDS OF ACHIEVEMENT

The report makes a useful distinction between formative assessment and other kinds of assessment, but then proceeds to link the former to summative assessment by proposing that it be 'the basis of an on-going running record of progress'. It certainly *can* be that, but the problem is that knowing this to be the case will alter the manner in which the subjects of the assessments, that is the pupils, approach the formative component. Knowing that what you say and do will count towards a public statement about how you perform will surely influence how you say it and do it. In particular it is likely to militate against complete honesty and admissions of weakness, and in this way undermine the foundations of formative assessment. At the very least the report should have aired this issue and addressed the argument that formative assessment may need to be completely decoupled from summative assessment.

There are plenty of problems attached to records of achievement, many of which have been discussed in the recent national evaluations. Ensuring comparability of teacher judgements, obtaining a high enough status for them and making them of importance for the academically able as well as the academically weak are serious problems which need to be solved if records of achievement are to succeed. It seems to me that the report fails properly to address such problems.

SCHOOL EFFECTIVENESS

The report does not really question the usefulness of schools publishing exam and assessment results, even in 'value added' form. Yet the research in this area above all urges caution. Even the most sophisticated 'value added' results have large measures of uncertainty attached to them, and because they are inevitably out of date, are of little use for parent choice (Goldstein, 1992). We have no

evidence that the publication of assessment results for schools yields specific improvements, and for the report to argue in favour of the publication of school level reports, even highly contextualized ones, strikes me as naive and dangerous. Why, one may ask, *should* the results be published? Are there not more direct ways of encouraging schools to debate their policies and activities?

POLICY CONSIDERATIONS

The educational policy of the present administration is based upon some very clear objectives, revolving themselves around the need to centralize control and to put in place a particular kind of market place ideology. The methods which have been used to achieve this are now well known. In so far as assessment policy is concerned, this has been a principal vehicle for achieving these aims, though this could well change if the manipulation of assessment strategies is seen to have failed to achieve its political purposes. Yet the report has nothing to say about the political context. One consequence is that assessment is not viewed as part of educational policy-making so that, as a critique of current policy, the report lacks incisiveness.

Assessment does not consist merely of a neutral set of procedures for 'promoting learning' but has to be contextualized within the general aims of educational policy. Our attitudes to such ventures as the APU or the use of records of achievement or the public accountability of schools through test scores, must depend on how we think the products of these activities will be used. After all, it makes an enormous difference if school comparisons are to be used to close down those which are at the bottom of a list rather than to give them extra resources.

CONCLUSION

There are serious theoretical as well as practical problems surrounding all the issues dealt with in this report, and it is a pity that the report misses the opportunity to insist upon the need to pilot and evaluate its recommendations. As a research organization BERA is well placed to offer advice on this and I would strongly suggest that the Task Group now considers how to set up such a programme of implementation and evaluation research.

REFERENCES

- Goldstein, H. (1992) 'Statistical information and the measurement of education outcomes' (editorial). *J. Roy. Statist. Soc. A*(155): 313–15.
Goldstein, H. and Wood, R. (1989) 'Five decades of item response modelling'. *Brit. J. Math. Statist. Psychol.* 42: 139–67.