

✓

②

WHO'S AFRAID OF EVALUATION? *

Barry MacDonald.

Centre for Applied Research in Education, University of East Anglia.

Evaluation and Assessment

It may be helpful to begin by offering a distinction between 'evaluation' and 'assessment' in education. The distinction is a fairly conventional one, not intended to be definitive or provocative, but simply to avoid confusion within the compass of this article. The purpose of assessment is to make statements about the recipients of an educational service, statements about their actual and potential accomplishments in relation to the opportunities for learning provided by that service. Assessment is the basis for decisions about what students will get in the way of further provision and for predictions of their future accomplishments. The purpose of evaluation, on the other hand, is not to make statements about the recipients, but to make statements about the educational service. Evaluation statements serve decisions about educational provision. It will be obvious that assessment can assist this purpose, and that some form of assessment is a likely component of any evaluation process. True enough, but the distinction by primary purpose is more significant than may at first sight appear to be the case.

The fact is that the instruments of assessment that we have laboriously cultivated and refined over the years, the examination procedures and the psychometric tests, are not much help to the educational evaluator. Such instruments are typically constructed so as to differentiate between the accomplishments of individual learners, that is, to achieve a distribution of scores. These instruments are known as norm-referenced, because they make it possible to rank the individual in relation to the performance of the group. This discrimination facilitates educational and vocational placement. But assessment for the purpose of evaluation calls for a different approach, because it is asking different questions. What did the students need to learn? Given the nature of the provision for learning, the circumstances of its implementation and the intentions of the

* This article was written for publication in Education 3-13 Vol 4 No. 2.

teachers, what could they reasonably have been expected to learn? What did they *all* learn and what did they all fail to learn? Instruments which set out to answer these kinds of questions require *criteria* rather than norms as their design basis. They are known as criterion-referenced, and they have an important function in the evaluation of educational programmes. The massive task of constructing such criterion-referenced tests which are reliable and valid is only now getting underway. Very few are readily available in this country at the present time, and this in part explains why during the last decade, evaluators have increasingly adopted and developed approaches to their task which place little emphasis on the measurement of student performance. I shall have more to say about this later: it is enough for the moment to have established that evaluation and assessment, as I have defined them, have different purposes and that, even though they share an interest in student learning, they require different instruments.

The Context of Evaluation - Values

But I don't want this introductory article to be primarily a technological review of the province of evaluation. The rising demand for evaluation in education and I think it wiser to talk of 'demand' rather than 'popularity', calls instead for a much broader consideration of the circumstances which surround this demand. Evaluation can so easily be an instrument of abuse that the context in which it's applied should always be carefully scrutinised, 'Who's Afraid of Evaluation?' the title of this article, is offered in a spirit of caution rather than contempt.

These are troubled times in education, are they not? The air is shrill with criticism, threats, challenges, war cries of one kind and another. Some of the rather sloganised exchanges that characterise the public debate have been with us for some time, and are familiar enough - formal versus informal teaching methods, comprehensivisation versus selection, streaming versus mixed ability, and so on. They may have sharpened in tone of late, but we know them well. In a sense, many of us feel that much of this debate lies outside education although it is about education. What actually happens within the schools is determined still by the values and preferences of those most intimately involved in educational provision - headmaster, teachers, and local administrators. This degree of autonomy in matters of curriculum provision, although a thorough analysis would no doubt qualify it in a number of ways, has been sufficiently substantial for us to feel that the affairs of the school are under professional control, and that this arrangement enjoys public confidence.

Can we still say that today? I think not, and I have in mind not only the effects of the economic disaster which has overtaken educational expenditure and which I shall come to shortly, but of other trends and events whose effects are rather more difficult to calculate. Events like the William Tyndale saga, which culminate in an inconclusive three-month-long 'evaluation', and which raised a number of issues that are much less familiar than the 'chestnuts' referred to earlier. Tyndale was an event which compelled the confrontation of long dormant issues in education in the light of contemporary trends. For instance, who is accountable for educational provision, to whom, and by what means? And, if a pluralist society entails a pluralist teaching profession, what are the implications for the school as an agency of cultural transmission? To put this last question more baldly, can schools still be relied upon to reinforce the existing social order? The 'problem school' in this sense is a new phenomenon, one which many people link with the inner urban crises of recent years, as well as with a general growth of political consciousness within the educational community. The extent to which Tyndale, and its successors, leads to an erosion of public confidence in the schools may well depend on the extent to which it is seen as an opportunity to change the existing distribution of power within the education system. The TES saw in the actions of the parents of Tyndale the emergence of a new tide of educational 'consumerism', The ILEA followed the inquiry by appointing a new group of inspectors to monitor school standards. The DES reinforced its recent call for a national monitoring system. There is at least incipiently, a sense of crisis in the air, and a tendency to respond by seeking strong measures to re-establish control. Whereas Rivinghill was seen by its critics as an aberration, Tyndale has been interpreted as an early warning signal.

The Context of Evaluation - Efficiency

If cultural fragmentation is one context which gives rise to calls for closer monitoring of the school system, inept management of resources is another. During the seventies, we have seen, both at the national level and the local level a significant change in the model of resource management applied to educational expenditure. Using techniques and procedures evolved primarily in industry and defence, the planners in Elizabeth House have designed, and are seeking ways to apply, a system of management which involves a shift from input to output budgeting, a system which calls for evaluation of the effects of alternative resource allocations in relation to objectives and costs. To put into operation such a system requires the co-operation of the planners at local level, and since 1971, there has been considerable exploration of the integration problems of local and national planning.¹ The reorganisation of local government in 1974 gave an immediate, and

perhaps crucial, boost to the new model of management, introducing as it did the 'corporate management' structure. Corporate management involves the formation of cross-departmental executive committees, horizontal movement of staff at senior management level, and a strong emphasis on policy evaluation of a kind that is required if the aspirations of the Department planners are to be fulfilled. Now, at the local as well as the national level, we have a commitment to goal-setting and the evaluation of goal attainment. The new power structure of local authorities ensures that Education, the 'big spender', will come under pressure to justify its resources in terms of demonstrable outcomes. Game, set and match to the men from the Ministry.

The New Rhetoric

Add to the foregoing contexts the economic recession and we begin to understand why the mood of optimism which imbued public discussion about curriculum development in the sixties, and the rhetoric of constructive support and co-operation that went with it, are sadly frayed. Now we hear, suddenly it seems, a new vocabulary of relatively unfamiliar terms, a vocabulary more militant in tone than we are accustomed to, more strident in advocacy than many of us feel comfortable with. Top of the list is the word 'accountability', a concept we used to associate with financial audits, but which is now applied to the processes and outcomes of schooling, and to the obligations of those who have responsibility for those processes and outcomes. Many other 'new' words we are being asked to learn seem, like accountability, to belong to the world of production organisations, words like 'cost benefit' 'efficiency', 'management by objectives', 'programme budgeting', 'control systems' and the like. The tendency of language like this is to suggest that the production of educated people is much like the production of anything else, a technological problem of specification and manufacture. And, like the production of anything else, it can best be organised hierarchically, each part or sub-system dovetailing into a master plan which encompasses and orchestrates the whole organisation.

In the case of education, specifying the product is the really tricky bit of the process. First of all, everyone in the organisation must agree what the product is to be. To derive an analogy from car production, it's no use if some of the workers think it's a Mini while others think it's a Land Rover. Objectives must be pre-determined. Secondly there must be agreement about what the product will look like, how it will be identified. In education, this means that the objectives must be defined in terms of learner behaviour. And this is where another word enters the contemporary vocabulary, the word 'standards'. Not a new word of

course, 'standards' is the traditional war cry of the educational 'right'. 'Standards' are what grammar schools stand for, what you get if you concentrate on the three Rs, what you lose when your hair touches the back of your neck or when you enter a comprehensive school. 'Standards' is the central construct of a familiar polemic at the political level of education policy disputes. But of late the concept has come more and more to be employed in its operational sense, as criteria of acceptable performance. When the Department of Education's newly fledged Assessment of Performance Unit talks of standards, it means levels of pupil achievement which can be used as the basis of a national monitoring system.² In other words, specified behavioural objectives.

While it would be going too far to suggest that what is taking place is a transformation in the way that schooling is viewed as resource investment, it would be wrong and dangerous, to dismiss this trend as mere verbal fashion unlikely to impinge in any significant way upon the activities of the school. The changing rhetoric of educational management reflects in a very direct way the transformation of the economic circumstances of schooling that has been wrought in the past two years.

It would be difficult to resist the proposition that the British educational system, and particularly the school system, is about to launch itself into a new phase, perhaps a decade of unprecedented concern with evaluation. Unprecedented that is, in modern times: some of the rhetoric of the platform speakers bears a chilling resemblance to the remark of Robert Lowe, Vice-President of the Education Department, when, in 1861, he recommended to the House of Commons the issue of a 'revised code' which introduced the notorious and ill-fated system of 'payment by results' - 'If the new system is not cheap, it shall be efficient, and if it is not efficient, it shall be cheap'. In one sense, current rhetoric goes one step further. The context of Lowe's promise was one of rapidly expanding expenditure on schooling, which payment by results failed to stem. The context today, with the announcement by the Chancellor, in February, 1976, of plans to cut public expenditure by £3,000 million, including £619 million off education, is one of a declining resource future. The new system therefore, has no options. It shall be efficient, and it shall be cheap.

Thus baldly stated, the demand represents a transformation of the social and political circumstances of the school, one which has materialised with bewildering rapidity. The demand is backed by fiscal control and made legitimate by public concern about 'standards', it cannot with impunity be ignored. Already there are some ominous signs. The ailing National

Foundation for Educational Research is showing all the signs of revival as orders for new tests come in from the Department, itself gratified by Bullock's call for national monitoring of reading standards. The Assessment of Performance Unit has announced the membership of its Consultative Committee which will eventually yield, subject by subject (including morals and aesthetics) the lists of objectives which will serve as the bases of evaluation. A change of role is planned for the HMI's who want the local advisors to take over much of their individual school inspection function, and it is interesting to note a change of nomenclature at the local level, where some former advisers are now called evaluators. And here come the volunteers. I note in a recent issue of the *Times Educational Supplement* that the principal of a College of further Education has welcomed accountability in the form of an external audit. Who's afraid of evaluation?

Accountability and Evaluation

Where does all this leave the school, and the evaluators, whoever they may be? The context which I have attempted to outline suggests that future evaluation efforts in the schools are likely to arise out of, and feed back into, a centralised system of performance monitoring based on pre-set targets. We should note several salient features of this system. In the first place, viewed as an accountability system, it is clear that some are more accountable than others. The Department of Education and Science for instance, thought part of what Margaret Thatcher called the least accountable bureaucracy in Western Europe,³ and though it has led the crusade for accountability, is not offering to render its own operations and accomplishments more open to scrutiny. Of course the management model does in theory render the Department more accountable to its political masters but if one believes, with Professor Vaizey, that such accountability has 'vanished, even as a pretence',⁴ and if one adds to this Tyrell Burgess's view that the effect of corporate management is to make 'it harder than ever for the representatives to control the officials',⁵ then we face the possibility of an accountability system which only bites at the level of the schools. In the second place, the model of evaluation which is being advocated is extremely rudimentary and simplistic, one which no professional evaluator working in this country today would try to defend. Evaluation as the measurement of pupil attainment of pre-specified objectives is an excellent instrument for the purpose of central control of education, but a poor instrument of quality assurance, and an even poorer instrument for increasing understanding of the problems and potentials of educational provision. It is a bureaucratic concept of evaluation for a bureaucratic concept of accountability. Let the buyer beware. Ernest House has observed, in relation to the enactment of comparable accountability procedures

in America, 'I believe such schemes are simplistic, unworkable, contrary to empirical findings, and ultimately immoral. They are likely to lead to suspicion, acrimony, inflexibility, cheating, and finally control -which I believe is their purpose.'⁶

However we do not need to postulate malign intentions in order to predict malign effects and indeed would be unjustified in doing so. Few would deny that schools should be accountable, more so than they have been in the past, and most of us would support the need for more systematic study of alternatives in educational provision - their values, their processes, their circumstances and their effects. But there are many ways to be accountable, many audiences with legitimate claims for information and different notions of what constitutes educational excellence. Evaluators increasingly acknowledge these propositions in their designs, which have moved away from simple output measurement in an effort to portray for decision makers and others the complexity of the relationships between actions, circumstance, and consequence in educational affairs. It is this complexity that is in danger of being ignored if we subscribe to a form of evaluation that places overmuch reliance on a single productivity index, an index which, as I pointed out at the beginning of this chapter, is underpinned by an undeveloped technology.

One final point. To judge schools by what we can measure in pupils is to judge them by our instruments, not by our heads or by our varying values. Ledger-book evaluation may lead us to assign false priorities, undue emphasis in our provision. These recent words of Robert Stake are worth bearing in mind; 'The world encourages a great variety of competences across persons and tolerates a great incompetence in persons. Poor achievement is not despicable (though failure to provide opportunity for better achievement is). Poor achievement is often the scapegoat focus of attention when a person is rejected really because of social class, race, personality, or appearance. The incompetences of the handsome rich are greatly tolerated. A successful life is possible for any person with any combination of talents. With a heavy accent on low competence in certain academic areas, the schools can help deny a person the ordinary opportunities of a successful life.'⁷

And what of those who hold that education is characterised by indeterminacy of ends, anyway? Who's afraid of evaluation? I am, are you?

References

1. See *Management in the Education Service: Challenge and Response* an SEO Occasional Paper, printed by Carter, Halls & Co. Ltd., Coventry, 1974.
2. See Brian Kay, 'Monitoring Pupils' Performance in *Trends in Education*, July, 1975.
3. In a televised interview with Robin Day during the Common Market Debate.
4. John Vaizey, 'Expansion and Contraction', in *Times Educational Supplement* 28th June, 1974.
5. In *The Guardian*, 7th May, 1974.
6. Ernest House, *The Price of Productivity: Who Pays*, University of Illinois (mimeo), 1973.
7. Robert Stake, *Measuring Achievement in the schools*, University of Illinois (Mimeo) February 1976.