COMING TO TERMS WITH RESEARCH

**An Introduction to the Language
for Research Degree Students**

CENTRE FOR APPLIED RESEARCH IN EDUCATION

The Centre for Applied Research in Education was set up in the University of East Anglia in 1970, and is now a constituent of the School of Education of the University, while maintaining its own identity.

It is primarily a research community of experienced educators, some of whom are predominantly engaged in sponsored or commissioned research projects, many of whom combine such work with teaching and supervision on the range of post-graduate degree courses supported by the Centre, and most of whom are, of course, students who come to the Centre to develop and refine their research experience within the framework of the degree programme. This research core of the community is supported by a fully integrated and experienced staff, some of whom are part of the Centre management team, others of whom are members of particular project teams, all of whom participate in the collaborative ethos of a busy, but friendly organisation.

The 'identity' of the Centre has elements of continuity and elements of change. It is made up of substantive concerns and methodological leanings. The phrase itself 'applied research in education' expresses the Centre's commitment to field-based investigations of educational practice and policy, and its responsiveness to problems of action. Since its inception the Centre has been continuously involved with programmes of educational change and development, particularly with the application, adaptation and invention of research methods to meet the needs of practical decision-making at all levels of the system.

Although CARE is probably best known for the pioneering work of its first Director, the late Lawrence Stenhouse, in initiating the 'teacher as researcher' movement in in-service education, the image of the Centre has also been shaped by Barry MacDonald's early work in developing a naturalistic methodology of case study in evaluation, John Elliott's influential leadership of the action-research tradition, and by a succession of research, development and evaluation studies in significant areas of curriculum policy. These are basic elements in the continuity of the Centre's identity - a tradition whose separate parts share an aspiration to place the power of research within the grasp of educational professionals, and the value of research within the judgement of the public they serve.

EVALUATION IN THE SOCIAL SERVICES

What evaluators do

Evaluators investigate the work of social service providers. Their task is to construct a body of evidence to support judgements, made by others, of the merits of existing practices and the policies from which they derive. Their focus is almost invariably highly specific - a community health care scheme, a training course for detectives, a mother tongue teaching programme in London for linguistic minorities.

They are usually commissioned by those who also have responsibility for the programmes under scrutiny, and the power to shape their future, and this puts a premium on the fairness, as well as the accuracy they bring to their representations. Their investigations are guided by the questions and expressed concerns of those whom the evaluation is intended to serve - initially the sponsors and perhaps other stakeholders involved in the commissioning process, subsequently others on whose cooperation the evaluation may well depend. As a general rule (frequently breached) evaluators should not seek answers to questions nobody is asking, though it is reasonable to draw attention to phenomena their user groups might logically be interested in. Clearly this service stance of evaluators with respect to their audiences commits them to methodological eclecticism if not paradigm surfing. They may be asked, in any one case, for a portrayal of the programme in action, for measures of learning outcomes, and for an assessment of union attitudes. The anatomy of decision making (a phrase coined by House) is vitally important for evaluation, and needs to be identified rather than assumed. Evaluation reports tend to be more comprehensive than other research reports in terms of the range of issues they address, activities they describe, and perspectives they present, and should be more tentative. Conclusiveness is not a virtue in evaluation. Evaluation reports are intended to inform the deliberation of decision choices, not to pre-empt that process. Evaluators are often invited, indeed pressed, to make recommendations for action. That is not their business, though the line can be a fine one. There is a difference, for example, between reporting that a certain activity, universally admired, is under-resourced, and recommending that more resources be allocated to it. The first is an evaluation function, the second a matter for those who are politically accountable for resource allocation.

You should note a conventional distinction in evaluation, between formative and summative orientations. Historically most evaluations were funded as adjuncts of innovative development programmes, and functioned as feedback loops to developers for the purpose of ongoing revision while the programme was still fluid and open to improvement. The emphasis then was this formative role for evaluation. Summative evaluation is oriented more towards external audiences seeking to judge the value of established forms of provision, though it may also be

applied (usually prematurely) to innovations. Stake offers a helpful analogy. "When the cook tastes the soup, that's formative evaluation. When the customer does, that's summative."

And What They Don't Do - Some Distinctions

Bear in mind that evaluation investigates the enactment of ideas, beliefs and values (i.e. policy) in the forms of social action to which they give rise, so that they may be reviewed in the light of their effects. It's focus is the quality of provision, and it should not be confused with assessment or appraisal, which are concerned with the performance of people. Assessment is about the achievements, needs and potential of students, whilst appraisal serves a parallel function with respect to employees. Assessment and appraisal data may constitute part of the evidence assembled by an evaluation, but only in so far as they cast light on the merits of the provision from which they arise. In some cases such data may tell us how effective a programme is in achieving its objectives, but nothing about the costs, the side effects, the unanticipated or unintended outcomes, or the worth of the objectives themselves. Nor should you confuse evaluation with monitoring or audit, both of which are system maintenance procedures. Monitoring is a surveillance instrument that provides regular, standardised information on the level of functioning of a given system. It may be carried out by direct inspection or indirectly by performance indicators or by a combination of the two. Audit, on the other hand, is a regular or occasional check on the adequacy and integrity of internal accounting procedures.

You may find these distinctions useful, if somewhat arbitrary, as you attempt to come to terms with a literature which has a tendency to call all of them 'evaluation'.

Evaluation and Government

The boom years for the growth of evaluation as an academic pursuit date from 1965, when the federal government of the USA invested a billion dollars in compensatory programmes of education for poor children, and required all such programmes to be independently evaluated. The idea was that evaluators would ride shotgun on the federal coach, to make sure that the money got through to its intended beneficiaries. The idea spread to other social investments and to other countries, surviving as it did policy shifts from investment to disinvestment in the boom/bust economics of the West. And it was reconceptualised as a form of social reality testing for governments increasingly intent upon social welfare engineering.

Evaluation has been, and continues to be, largely concerned with what governments and their agencies do with our money for our supposed benefit. Who gets what, and at whose expense? We are talking about winners and losers in the arena of social action, and we are not just talking

about money, or even what money buys. In a pluralist democracy we have to add values and beliefs to the mix. In such circumstances judgements of worth are inherently problematic and unlikely to be consensual. Whose judgement is to count, whose process of judgement formation is to be informed, whose questions get answered? These are political issues, mediated by the inescapable need, in the case of most government funded programmes, to inform judgements at many levels of responsibility for action.

But what's the problem, some may ask? We elect governments to make those judgements, to take those decisions. If we don't like the results we vote them out. Fair enough, if you believe that that singular power is sufficient, and sufficiently informed by the media. If, on the other hand, you believe that democracy is a continuous debate about the pros and cons of government actions, and that such debate is underinformed by the fitful attention span of the media, misinformed by government itself, and in need of a more credible information service than either provides, then the possibilities opened up by evaluation-oriented research are worth exploring. You might want to argue that the public has a right, as well as a need, to know what the Government is up to, and how it is doing. Some concepts of evaluation (MacDonald's 'democratic' evaluation, Stake's 'responsive' evaluation, House's notion of 'impartiality', and Cronbach's notion of the evaluator as 'public scientist') reflect in different ways resistance to the idea of evaluation as a technical service to social system managers, and support for a more communal, or at least 'stakeholder' resource.

So much for theory, most of it generated in more benign economic environments than we now enjoy, most of it by academics for whom public knowledge is a stock in trade, most of it by educationists at a time when educational investment was a top priority. As House has noted, all the major theories of evaluation were ideologically 'liberal', and for a time they were broadly acceptable to the public bureaucracies which saw in programme evaluation, conducted by independent social scientists, not only a source of feedback on the effectiveness of their actions, but a means of legitimating such actions in the eyes of the public.

As the planned and managed societies of the post-war era evolved, however, the circumstances that had to some extent favoured the open society and the 'liberal' evaluator were transformed. Without going into detail, the combination of economic decline/instability and rising social welfare demands made for regimes that had to promise investment to get elected but cut it to stay there. As the gap between political rhetoric and social reality grew, so did the need to control the production and dissemination of knowledge about the effects of government action. Evaluation thrived as an indispensable tool of expenditure management, but became increasingly incorporated as an in-house function of administrations or farmed out to private consultancy firms. The evaluation community diversified, peopled now as much by civil servants, management consultants and accountants as by social scientists. Academics still

compete for evaluation contracts - on domain knowledge, on research skills, on social and professional acceptability and credibility, but find themselves bound by stipulative clauses that limit many of the freedoms they used to take for granted. (see Norris and Pettigrew) Mind you, that is not to say that they use those freedoms to good effect, or that they weren't always mindful of the dangers of offending a sponsor. Eyes down and palms up is not a hitherto unknown posture in the research community.

Fortunately this gloomy picture is not the whole picture, even in those countries like Britain where it most clearly applies. Bad news isn't always suppressed or distorted, evaluation reports get leaked to newspapers by unidentified miscreants or lost in pubs, validation procedures and consultation exercises erode control of the data, and oral dissemination is particularly difficult to monitor. Self-censorship is possibly the greater danger, particularly as universities become more and more dependent on research earnings. In any case, government is not the only sponsor (at least in more Western countries). In this country there are the Research Councils and the Charitable Foundations, both of which, though fearful of offending Government, still fly the flag of academic freedom whilst disassociating themselves from the views they sponsor.

It's not the whole picture in another sense. What we have been talking about so far is what we might call 'big' evaluation that facilitates a critique of macro policy for which government is directly accountable. But most of the evaluation that goes on is negotiated, funded and carried out at lower and more localised levels of the social system, where the political agendas are more varied and the focus of the evaluation designed to reflect responsibilities and aspirations at those levels and locations. Nor should we forget the considerable number of research theses devoted to evaluations of social service delivery. These are public documents, their conception and construction not (yet) subject to political approval. Few reach commercial publication; they constitute an undertapped resource that may become more important to exploit should the noose on public knowledge tighten further.

Evaluating the New

Modern governments are committed to social engineering, continuous intervention in the way things are, despite a disappointing record in the field of change initiatives. Social modernisation (variously interpreted by ideological leanings) is both difficult to achieve and electorally sensitive. Despite almost forty years of varied experience in different countries, we are still unsure about how best to combine central control of financial resources with the conditions of grassroots human productivity. That is not to say that there hasn't been a learning curve, on the part of those who finance innovation, those who carry it out, and those whose task is to evaluate new ventures.

What follows is a profile of contemporary programme evaluation in the form of five statements which attempt to summarise how far evaluators have travelled in terms of the perspective they bring to bear upon the new. The profile is by no means uncontentious.

- We no longer treat programmes as if they were poorly designed experiments from which, with a bit of care we can extract some useful (ie, decontextualised) generalisations. Rather we see them as new elements (new in kind or in quantity) introduced into a social situation in order to promote certain values at the expense of others.

- We no longer treat programmes as disembodied ideas enacted through the interplay of role sets. Rather we see them as resources in the marketplace of individual and institutional aspiration at a point in biographical time.

- We no longer treat programmes as if they will happen but rather as game plans conceived on the basis of unsafe projections and subject to continuous modification in the light of experience.

- We no longer treat programmes as if they had an unquestionable right to exist. Rather we see them as options exercised by those in delegated authority, using resources for which they are publicly accountable.

- We no longer see ourselves either as technicians or as judges but rather as reporters of action, interpreters of meaning and brokers of information.

This process of redefining the objects of evaluation in terms of dynamic social action has been slow to emerge, as has the parallel review of evaluation as a particular form of research which both changes and is changed by the situation with which it interacts. It seems fair to say, however, that there is now widespread recognition of the need for evaluators to adopt a responsive, rather than a preordinate, approach to the design of their enquiry and to think more in terms of a process-oriented rather than a product-oriented view of the contribution they can make, particularly to Programme development (Stake, 1975). What do these changes entail, what difference do they make to the work of the evaluator? Let us look briefly at an 'ideal' model of contemporary evaluation.

We negotiate with sponsors, participants and audiences, seeking and accommodation of interests to guide our questions and observations. We seek access to all the levels of determination relevant to programme origin and action, so as to get a grasp of how and why the programme came about and what the consequences of changing, continuing or abandoning it

might be. We stay close to the programme from start to finish, responding to changes of context or of key personnel that may have implications for the direction, focus or timing of evaluative feedback. We see ours as an educative role, widening the sphere of deliberation both in terms of the numbers participating and in terms of the range and depth of the information that is taken into account in a graduated process of reshaping beliefs. We are case-oriented, relatively non-comparative, respecting actor frames of reference and values in representing their work to others. We are sensitive to the human consequences of reporting social action and the threat implicit in the evaluative act. We operate from no particular discipline base, produce no grand theory, pursue no personal theory. We work within the language of those we seek to influence. we offer methodological competence in the construction of new knowledge rather than substantive expertise. We depend upon educative interaction to achieve impact rather than upon authority. We do not conspire, we do not collude and we insist upon our obligation to represent the views of those who hold neither power nor office. In the latter respect, we are political brokers.

Reporting Evaluation

Evaluators study public life in private places. They record what particular people do and think at particular times in particular places. There's a story in that, as Stake (1967) pointed out a long time ago. Evaluation data offers the possibility of organisation in narrative form. Think about the kinds of questions evaluators have to deal with. WHO did WHAT to WHOM, WHERE, WHEN, WHY, HOW, and how did it all work out? We might call that CIPP (Stufflebeam,1971) - Context, Input, Process, Outcome. It's just another way of looking at it, and another way of telling. Given the known difficulties of getting evaluation reports read, let alone used, evaluators need to develop more accessible and engaging ways of reporting, Along the way they may lose some of the stamp of science, but gain some of the power of art.

References

House, E.R. (1980) *Evaluating with Validity*, London, Sage Publications

Stake, R.E. (1975) *Evaluating the Arts in Education - a Responsive Approach*, Colombus Ohio, Merrill Publishers

MacDonald, B. (1974) 'Evaluation and the Control of Education' in MacDonald, B. and Walker, R. (eds) *Innovation, Evaluation, Research and the Problem of Control*, Norwich, CARE Publications

Stufflebeam, D.J. et al (1971) *Educational Evaluation and Decision making*, Itasca, Illinois, F.E. Peacock

Cronbach, L.J. et.al (1981) *Towards Reform of Program Evaluation*, San Francisco, Jossey-Bass

Norris, N. and Pettigrew, M. (1993) 'Expand and Contract - the conditions of government sponsored social research' Norwich, CARE Mimeo

Bibliographic Guide

Key Papers

Stake, R.E. (1967) 'The Countenance of Educational Evaluation', *Teachers College Record*, 68, 7, pp 523-540

House, E.R. (1972) 'The Conscience of Educational Evaluation', *Teachers College Record*, 73, 3, pp 405-14

MacDonald, B. (above)

Simons, H. (1971) 'Innovation and the case study of schools' *Cambridge Journal of Education*, 3, pp 118-23

MacDonald, B. and Parlett, M. (1973) 'Rethinking Evaluation: Notes from the Cambridge Conference', *Cambridge Journal of Education*, 3, pp 74-82

Cronbach, L.J. (1975) 'Beyond the Two Disciplines of Scientific Psychology', *American Psychologist*, 30, 2, pp 116-27

Selected Books

Best of the vast USA literature - the House and Cronbach books already mentioned, plus

Patton, M.Q. (1978) *Utilization Focussed Evaluation*, Beverley Hills, Sage

Stake, R.E. (1986) *Quieting Reform*, University of Illinois Press

House, E.R. (1993) *Professional Evaluation*, London, Sage Publications

Best of British:

Hamilton, D. et al (1977) *Beyond the Numbers Game*, London, Macmillan Education

Simons, H. (1987) *Getting to Know Schools in a Democracy*, London, Falmer Press

Norris, N. (1990) *Understanding Educational Evaluation*, London, Kogan Page Ltd.

Notes:

(1) Most of the literature cited above is concerned with external evaluation. The exception is Helen Simons, the best guide to internal evaluation, especially of schools.

(2) All of House's contributions end up as chapters in his books. Not so Stake, whose (generally) brief contributions are well worth hunting down.

(3) CARE Publications include *SAFARI 1* and *SAFARI 2*, particularly useful on methodology, and a number of substantive evaluation studies.