Evaluation: Business or Vocation?

JOE O'HARA
Dublin City University

GERRY McNAMARA
Dublin City University

Introduction

The following is offered as a response to the challenge to redefine the 'market' for evaluation as it currently exists in Europe, laid down by Frans Leeuw, Jacques Toulemonde and Andre Brouwers. This challenge was initiated by Leeuw and Toulemonde in the course of a lively debate which took place at the 1998 Rome Evaluation conference and was developed in conjunction with Andre Brouwers in their article 'Evaluation Activities in Europe: A Quick Scan of the Market' (Leeuw et al., 1998).

This response attempts to explore a number of the key issues which emerged in the course of the debate and to sketch an alternative, though it is hoped complementary, analysis of the characteristics of evaluation suitable for the potential market in the coming millennium. In order to do this, the authors argue that it is necessary not only to explore the nature of the market but also re-examine our understanding of the nature of evaluation as it is currently understood by the key stakeholders in the same market. Written from an educational perspective, this response looks at the increasing demands for evaluations to provide information which demonstrates the achievement of key, pre-defined objectives, whilst at the same time assisting in a process of organizational development and the democratization of decision making. These latter imperatives can often be at odds with the achievement of the same pre-ordained objectives. This then is the reality of the market for educational evaluators in Ireland, one which has a series of contradictions that require the development not only of a more rounded understanding of the needs and nature of the market, but also a revisiting of our understanding of the nature of the evaluation process that can most successfully meet its needs.

Post Paradigm War Evaluation

It is important to note when commenting on this debate that to those from an education background and relatively new to the world of evaluation, the tone and

substance of the debate on 'evaluation as a business' at the annual EES conference in Rome last October was a surprise. To be fair, this was only one of many sessions at a very interesting and wide-ranging conference and, given the theme, a certain hard-headed, realist view of the evaluation business might have been somewhat justifiable. Nonetheless in the context of the emergence of action research, qualitative studies and process, as opposed to product, objectives in educational research discourse, the objectivist, positivist approach to evaluation implicit and explicit in the session seemed out of kilter. Out of kilter, that is to say, not only with current practice in educational research but also with the ongoing debate that has featured in the journal *Evaluation* for some years past.

Schwandt (1997) argues that concepts of evaluation based on an essentially positivist, empiricist paradigm are untenable. He ascribes the development of this dominant evaluation paradigm to the work of Michael Scriven, Deborah Fournier and Nick Smith, and suggests that it is displayed in 'four formal steps':

- (1) establish criteria of merit (2) construct standards (3) measure performance and compare standards and (4) synthesize and integrate data on performance into a judgement of merit or worth. (Schwandt, 1997: 71)

Defined in these terms, evaluation, Schwandt argues, becomes a matter of:

... having the correct procedure for constructing descriptive, interpretative and/or evaluative statements, assertions or claims about various kinds of 'objects' that are evaluated. This approach, in turn is wedded to a model of strategic political action aimed at 'solving problems' in social programming. Administrators and policymakers seek to manage economic and social affairs 'rationally' in an apolitical, scientized manner such that social policy is more or less an exercise in social technology. (Schwandt, 1997: 74)

Schwandt goes on to argue for a 're-indexing' of evaluation theory which would link it to a 'practical hermeneutics' enabling a dialogue between the evaluator and those under evaluation.

To many engaged in educational research, certainly in Britain, Ireland and Scandinavia in recent years, the struggle which Schwandt sketches between the technical-rational approach to evaluation research design and a fundamentally post-positivist, interpretative paradigm is still contested territory. Subsequent articles in *Evaluation* referring to these issues have either sought to re-establish a realist agenda (Julnes et al., 1998; Pawson and Tilley, 1998), or to argue that these paradigm wars are giving way to 'dialogue and accommodation between paradigms by adopting the "epistemic triangle formed by systems thinking, dialectical thinking and constructivist thinking"' (Bhola, 1998: 330).

However, whether this post paradigm war *rapprochement* is a temporary cessation or a permanent peace is a moot point. Kushner (1997), in a fierce attack on the most recent work of Scriven, argues that:

Scriven's concerns with robust theoretical underpinning for evaluation practice (Scriven, 1996a) and with the demanding set of competencies that define the eponymous 'evaluator' (Scriven, 1996b) sit uneasily with the reality that much evaluation is currently commissioned by anti-intellectual firms of management consultants who trade in formulaic solutions; that they and sponsors often hire solitary and vulnerable

individuals to conduct the investigation; that much of evaluation methodology is held to be 'commercially sensitive' and, therefore, secret; and that the intellectual upstarts of constructivism and postmodernism have scored some undoubted successes in engendering mistrust of conventional approaches to theory and role. (Kushner, 1997: 364)

Kushner goes on to suggest that Scriven (and implicitly much of current evaluation practice) systematically downplays 'process, portrayal and formative evaluation in favour of comparative, summative and product evaluation until the consequence that in the end formative collapses back into summative "product" evaluation' (Kushner, 1997: 367).

Evaluative Process or Product?

This then brings us to the heart of the matter. As educationalists, we are faced with policies – for example, teacher appraisal, league tables of successful and unsuccessful schools and school improvement through what is sold as cooperative school planning (but often amounts to little more than what Hargreaves (1994) calls contrived collegiality). Some of these we may regard as potentially useful if given a chance to develop (school planning), or very divisive and dangerous (most of the school appraisal schemes so far developed). Nevertheless, as educational researchers and evaluators, we are being asked to assist in the implementation of these policies, for example by facilitating school planning processes, and to evaluate the outcomes in order to disseminate 'best practice' to other schools.

This set of expectations creates all sorts of dilemmas. Can we be involved in the implementation and evaluation of policies that we mistrust or actively oppose? Can such involvement be from the standpoint of subverting these policies or at least limiting the damage they do? Should one be working to design a 'product'-centred, time-driven evaluation methodology designed to meet the demands of the commissioning agency while knowing full well that the process, for example collaborative, practitioner-led whole school planning, may be damaged or destroyed by such an approach?

To date our answers to these questions have been largely in the affirmative for a number of complex reasons. As already suggested some processes, such as whole school planning, have the potential for enhancing collaborative decision making and democratic involvement if the process can be both separated and protected from product demands and the tyranny of short-term objectives. In the case of other policies, such as school and teacher appraisal, one can only argue that it is probably better that educators with a lively appreciation of the dangers inherent in much of what is happening in these areas be involved in research, implementation and evaluation, rather than leave the field to the type of evaluation described by Kushner above.

The decision to become involved in educational evaluation, for whatever reason, invariably leads one to a set of key questions regarding underpinning philosophy and research stance. It will be clear to the reader by now that our view of what constitutes a defensible approach to educational research lies squarely in the realm of the naturalistic/interpretative paradigm. This does not mean that it

is impossible to evaluate programmes or policies in the sense of value or worth (how else could the foolhardiness of judging schools largely on examination results be demonstrated?). Rather it suggests that such judgements can only be made in the context of the multiple realities that constitute each individual case. Thus, evaluation in this context must involve the individual schools, the teachers, pupils and other stakeholders so that the factors which influence, constrain or liberate teachers, and the highly complex issues of power, control and politics of schools, can be taken into consideration. On this account of evaluation we soon move away from narrow objectives, short timeframes and certain outcomes which can be easily replicated in other contexts. Instead our investigation becomes exploratory in nature, and involves itself to a significant degree in the frames of reference of those involved. A focus on values is essential at the heart of collective empowerment processes, and thus evaluation is central and in this context will, of necessity, involve cyclical research processes allowing dialogue to change the nature and direction of the evaluation.

Of course, tendering to evaluate, for example, a school improvement project for a Department of Education or other governmental agency using the approach to evaluation outlined above is tantamount to ensuring that somebody else gets the job. It was this dilemma which featured in the article which brought us to Rome in the first place and first involved us in the debate in which we are now engaged. That article posed the question as to whether there are models or a mode of evaluation that can do justice (not damage) to improvement processes in education and yet simultaneously meet product/outcome measurement demands from funders? The article described the evaluation of school improvement/whole school planning projects in the Irish context and the attempts by the evaluators to resolve the dilemma detailed above. The approach developed attempted to marry the Habermasian notion, as interpreted by Carr and Kemmis (1986), of human knowledge constituent interests at three levels: technical, practical and critical emancipation, with three stages of the evaluation process producing different types of narratives for different audiences or end users of the evaluation.

Conclusion

The success or failure of the above approach is an issue for the article in question – and in any case, is, in the key sense of its impact on the collaborative planning process, still unclear. As of now, it appears all too likely that the evaluative process in this case will, in time, be seen as a negative and inhibiting factor in the emergence of a collaborative and democratic culture in the schools in question, despite being used as a model to encourage such improvement elsewhere (a very neat example of the evaluation dilemma as we see it).

The point is that in the education and social science milieu in which we operate, these issues and questions are contested in a way which hardly seemed to impinge at all on the debate on ‘Evaluation as a Business’ and by extension were not in any way factored in at the time in the definition of the ‘market’ for evaluation.

It was heartening, therefore, to read the attempt by Leeuw, Toulemonde and Brouwers to broaden the debate surrounding the nature of the market in their

article, although we would argue that there is still a way to go if we are to encompass the reality of the evaluation environment that we have attempted to detail above. Their exploration of the 'why' of evaluation activities, while in our view correctly asking what stage we are at in terms of our definition of the market, fails to acknowledge the necessity of extending the concept of 'evaluation' to the type of cyclical, iterative, organizational development processes that have become increasingly common in our experience. We would argue that the developmental evaluations which have emerged as a result of the growing popularity of action research and action learning in the educational and business communities are one of the key growth areas in the evaluation market in Ireland. This approach to evaluation, although initially small-scale and project-based, has started to appear at the regional and indeed national level, and has fundamentally changed the understanding of the nature of evaluation as understood by the 'patrons of evaluation'. Perhaps our real point of departure from Leeuw, Toulemonde and Brouwers comes with our understanding of the nature of the evaluative process. While they can say that 'many studies tend to be labelled evaluations nowadays which would have been called action-research or organisational learning in the past' (Leeuw et al., 1998: 7), we would argue that this broadening of the accepted definition of evaluation will be a key feature of evaluation theory and practice in coming years.

Perhaps it is this difference in scale that lies at the heart of our understanding of the potential of the new market. By re-focusing our attention on the organic evaluations, which are emerging at a small and medium scale, we can perhaps help in the maturation process mentioned by Leeuw, Toulemonde and Brouwers. A process which can demonstrate its worth to all stakeholders at every level of the commissioning process is one that can ensure a continuance and indeed a growth in interest in the evaluation market as a whole, the ultimate purpose of all our debates and discussions.

References

- Bhola, H. S. (1998) 'Constructivist Capacity Building and Systematic Evaluation of Adult Basic Education and Training in South Africa: One Discourse, Two Scripts', *Evaluation* 4(3): 309–50.
- Carr, W. and S. Kemmis (1986) *Becoming Critical: Education, Knowledge and Action Research*. London: The Falmer Press.
- Hargreaves, A. (1994) *Changing Teachers, Changing Times: Teachers' Work and Culture in the Postmodern Age*. London: Cassell.
- Julnes, G., M. Mark and G. T. Henry (1998) 'Promoting Realism in Evaluation: Realistic Evaluation and the Broader Context', *Evaluation* 4(4): 483–504.
- Kushner, S. (1997) 'A Critical Review of Some Recent Writings of Michael Scriven', *Evaluation* 3(3): 363–74.
- Leeuw, F. L., J. Toulemonde and A. Brouwers (1998) 'Evaluation Activities in Europe: A Quick Scan of the Market in 1998', article presented at the European Evaluation Society Annual Conference, Rome, October.
- Pawson, R. and N. Tilley (1998) 'Caring Communities, Paradigm Polemics, Design Debates', *Evaluation* 4(1): 73–90.

Evaluation 5(4)

- Schwandt, T. A. (1997) 'Evaluation as Practical Hermeneutics', *Evaluation* 3(1): 69–84.
Scriven, M. (1996a) 'The Theory behind Practical Evaluation', *Evaluation* 2(4): 393–404.
Scriven, M. (1996b) 'Types of Evaluation and Types of Evaluator', *Evaluation Practice* 17(2): 151–61.

JOE O'HARA is a lecturer at the School of Education Studies, Dublin City University. His areas of research include investigations into the role of evaluation as a catalyst for organizational change and the changing role of the evaluator in collaborative organizational settings. He has worked as an evaluator on a number of national programmes and is currently acting as a UNDP project evaluator.

GERRY McNAMARA is Head of the School of Education Studies, Dublin City University. His major areas of research interest include practitioner-led school-based research, organizational change and improvement and curriculum development and evaluation. He has published extensively both nationally and internationally and has acted as programme evaluator on a number of transnational programmes.