The Object of One's Passion:
Engagement and Community in Democratic Evaluation

Saville Kushner
Chair of Applied Research
Centre for Research in Education and Democracy
*University of the West of England, Bristol, UK*

"Maturity is the capacity to endure distance from the object of one's passions."
Peter Berger (1974: 26)

Indelible ideas

Some methodological ideas are ever-fragile; others endure. The former need persistent support and attention; to the latter we return as a default position. We return, not always because those methodological ideas serve us well, but because they speak to our need for stability and order – they offer us ways of seeing the world which seem to put events reassuringly within our control. Those fragile ideas – the ones to which we have only wavering commitment - are the ones through which we perceive a less than predictable world and which ask us, therefore, to take a risk, to suspend our belief in order. The question is knowing whether and when it is appropriate to confront or to contain our fears.

What stands the test of time in the world of programme evaluation is the *comparative model* – the belief that comparing outcomes with objectives will provide some measure of the wholesomeness or productivity of a programme. This has something of a default position about it, its influence is indelible – and it certainly speaks of order. It is most simply (admittedly, crudely) represented thus:
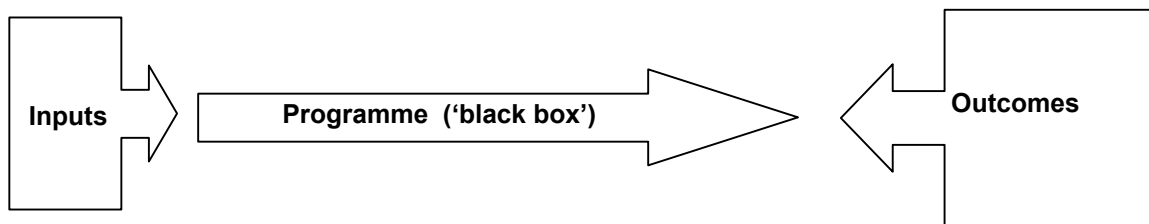
Inputs include programme objectives. The task of the evaluation is twofold: (a) to compare inputs with outputs to achieve a measure of accomplishment in value-added terms or to ensure compliance with stated aims; and (b) to use measures of outcome as an indication of the quality of programme process (hence, the outcomes arrow points back to the programme). The model is elegant in its simplicity, appealing for its rationality, reasonable in asking little more than that people do what they say they will do, and it is supremely efficient in its economical definition of what data counts. However, we have learned through serial evaluations of social innovation that, notwithstanding its power to measure programme productivity, the attendant shortcomings with this model are multiple. They include the following, all or some of which may apply in given situations:

- Social programmes are characterised by multiple and often contested objectives (since there are many stakeholder groups with different interests and values) No single set of objectives can, therefore, be fixed referents for the purposes of comparison;
- Outcomes may also be contested and, in any event, are not stable (ie. they erode and change over time and across contexts);

- The causal link between process and outcome is typically interrupted by so many intervening variables as to make it insecure – ie. we cannot derive from outcomes criteria to measure the quality of process;
- Key characteristics of perceived programme success may not be articulated in the vocabulary of outcomes and may not yield to measurement – while measurement instruments, themselves tend to favour those with technical knowledge;
- The model pre-empts the question of whether the initiative we are observing is a programme or otherwise – it persuades us to evaluate it as a programme (eg. with certain standard and consistent features) though it may measure up better as a movement or an idea or a policy;
- The approach is persuasive of conservatism in that it asks people to adopt a language of self-justification and to avert risk – eg. programme learning which leads away from initial objectives threatens failure against outcome criteria;
- Outcomes measurement delivers ready-made judgements which limit evaluation use by narrowing discretion in its utilisation;
- Outcomes are valued for their futuristic promise but are rooted in historical accountability.

This was the model most favoured by evaluators and governments in the post-war social reform periods in Britain and the USA, the period when evaluation was emerging as a discipline. Government administrations were seeking stable foundations for their social engineering projects. Employing social scientists to import conventional instruments to assess the merit of their programmes enshrined, very early in the life of the new discipline, this predilection for ordered states. A key problem, however, was that outcome measures may be useful for assessing the productivity of a programme but prove to be inappropriate for measuring its change potential. You cannot, it turned out, use the same methodology to hold programme people to account, on the one hand, and to encourage them to take risks, on the other. Those administrations wanted leverage, sought to understand the mechanisms of change so as to guide their social investments, sought, also, to encourage people to strive for new futures and to indulge in risks (Norris, 1990; House, 1993). What was needed was closer study and analysis of the programme, its processes and experiences – to get inside the 'black box' of change. Government wanted to be able to explain causality in a way applied social science largely failed to do. What they needed – and eventually sponsored – were bespoke methodologies designed for feeding judgement and for analysing change.

There were a number of bespoke evaluation strategies proposed for getting inside that 'black box' – principle among them Stake's (1967, 1975) *Responsive Evaluation* in the USA; Parlett & Hamilton's (1972) *Illuminative Evaluation* in Britain; and Barry MacDonald's (1976) *Democratic Evaluation*, also in Britain – and Carol Weiss's (1983) *Stakeholder Evaluation*, again in the USA. Each of these pulled back from an emphasis on outcomes and from attempting to explain causality, each was designed to open up the black box and to confront its complexity – in relation to, respectively, programme experience, process, politics and multiple aims. The aim of each was to expose contingencies - relationships between people and between people and events which allow us to theorise about change in complex ways. Each of them made the world a less predictable place, but each promised closer engagement in the management of change. Understanding in place of measurement was the key.

Once inside the black box the evaluator immediately confronted issued in the control over data and meaning – the methods had to be more interactive and iterative, requiring the evaluator to both build and sustain relationships in order to gain and maintain access. What was needed was close and sustained observation of interactions, probing interviewing and collaborative theorising. Perhaps the most radical promise of these approaches was the capacity to generate 'practical theories', theories which were collaborative attempts to understand and explain action – all of these approaches valued the respondent not just as 'informant', but as judge. Suddenly the evaluation resource was multiplied by the number of respondents who fell within the sampling frame of evaluation interviewing. Formal theory (Economics, Psychology, Ethnography) switched from being a determinant to a resource.

Black-box methodologies were designed for engagement. Here were the sources of contestation over judgement criteria, value pluralism, conflicts over interests, varying perceptions and parallax views. The authority of the evaluation – its social and political warrant – was easily brought into question as the sheer proximity of the evaluation raised questions of ownership of data, control over the evaluation and its agenda, the ethics of evaluation conduct, the politics of use of the evaluation data and, centrally, over the intersection between the evaluation and those power structures implicit in programme arrangements.

These have been called 'transactional' approaches, and in transaction evaluation is made transparent. The disturbances this brings to ordered situations is what makes political support for such approaches fickle and render the methodologies fragile, Where the comparative model could, to some extent, stand on the political sidelines administering its interrogative instruments, process-based approaches had to develop a political, 'transactional' ethic. In a post-war culture of liberalism that ethic was always likely to be democratic.

The comparative model is, notwithstanding its flaws, enduring – while democratic and process-focused approaches to enquiry remain fragile. The recent (global) return to reformist administrations have reinvigorated the model as governments have sought (as they did in the earlier period) to measure and then build upon the productivity of their pilot programmes. Pawson & Tilley (1998) have recently sought to restate the promise of evaluation to generate causal explanations in the laboratory-like conditions of a programme, but they, too, have been overwhelmed by the issues given above in a period of intensive and multiple innovations in Britain – all bearing down on vulnerable, excluded or under-performing groups, none coordinated, all claiming attributed effects from the same database. Nonetheless, a range of modern factors have militated against the sponsorship of those alternative approaches: government centralism, the suffocating effect of centralised target-setting and performance management, punitive accountability systems and coercive innovations, the confusion of quality control with creative potential – all of these favour an evaluation focus on outcomes. So the question is whether government is forced, once again, to confront the contradictions and shortcomings of the outcomes approach. We must assume so, for the problem of managing change has not gone away, and informed management serves the interest of governments who want more than the quick-fix of a successful programme. What goes round, comes round.

Democratic evaluation

So it is appropriate to revisit Democratic Evaluation as proposed by Barry MacDonald in the 1970s. He observed three categories of enquiry, each with its unique set of implications for the political positioning of evaluative enquiry.

- Autocratic evaluation
- Bureaucratic Evaluation
- Democratic Evaluation

The first of these is conducted within and for a community of academic peers. It is, for example, conducted by a researcher in pursuit of an independent agenda and reported, principally, in academic journals. The controls on the enquiry are derived from academic standards and the impulse for the evaluation comes from the scholar. In the second, Bureaucratic Evaluation, the evaluation is conducted on behalf of the administrative system – a management service for civil servants, who are a privileged audience. Control remains within the bureaucracy who set the agenda and whose values dominate the process.

The third category is clearly the one MacDonald was advocating and is the one he has pursued and elaborated since. Here, the evaluation serves multiple audiences and does not escape the democratic obligation of open reporting to the citizenry. This approach emphasises rights and obligations – for example, balancing the right of the general public to know about the programme being observed with the obligation of the evaluator to protect participants' confidentiality.

Here is how MacDonald (1976) characterised Democratic Evaluation (left column) with (right column) its associated reference point in the argument above. I suggest you first read the left column and then go back to pick up its referents in the right column:

"Democratic Evaluation is an information service to the community about the characteristics of an educational programme. Sponsorship of the evaluation does not in itself confer a special claim on this service. The democratic evaluator recognises value	*Democratic intent* *Control over the evaluation not exclusive to powerful groups (ie. we are 'inside' the black box)*

pluralism and seeks to represent the range of interests in [their] issue formulation. The basic value is an informed citizenry, and the evaluator acts as broker in exchanges of information between groups who want knowledge of each other. [The] techniques of data gathering and presentation must be accessible to non-specialist audiences. [The] main activity is the collection of definitions of, and reactions to, the programme. [The democratic evaluator] offers confidentiality to informants and gives them control over the use of the information they provide. The report is non-recommendatory, and the evaluator has no concept of information misuse. The evaluator engages in periodic negotiation of [...] relationships with sponsors and programme participants. The criterion of success is the range of audiences served. The report aspires to 'best seller' status. The key concepts of Democratic Evaluation are confidentiality, negotiation *and* accessibility. *The key justificatory concept is* the right to know.*"*	*Addresses plural aims and the instability of objectives – agendas constructed rather than pre-set* *Evaluation is part of the learning milieu* *Democracy of instrumentation* *Accumulates and then feeds judgements but does not displace them* *Gives informants control over their own data* *Enhances rather than narrows discretion in the use of evaluation* *Recognises programme learning and the shifting identity of the programme* *Emphasises external validity*

A key value underpinning both democratic evaluation and its methodological sister, case study, came from, *"the commitment to the belief that the subject being studied can impose its own authority on the sense that is made of it by the investigator"* (Walker, 1980: 224). Case study was synonymous with Democratic Evaluation in that it was conceived as a way of designing ethical sites within which power inequalities could be neutralised in favour of the elicitation and valuing of plural voices and in favour of broader publishing opportunities than are normally available to the evaluator. Democratic Evaluation operates through field-based, interactive, negotiated, situation-specific methods – which may be quantitative or qualitative.

With Democratic Evaluation MacDonald proposed, the first of evaluation authors to do so, a political model for evaluation (though Weiss & Rein, 1969, had identified political issues deriving from the fact that programmes typically had a multiplicity of aims – notwithstanding a single set of objectives). MacDonald saw programmes as sites for political action and described a political theory of action for evaluation in order to properly regulate its intervention.

Once in the 'black box' the evaluator was to discover political process deriving from the various contestations over value and criteria which are implicit in the list above. Democratic Evaluation saw programmes as microcosms of political society where, in well-understood ways, the greater your power the greater your potential for living in a self-determining and meaningful world, the more immediate the opportunity for self-expression and the greater your strength in the competition for resources. Evaluation, now fully immersed in the competition over voice and values had either to arbitrate among them – ie. favour one over others (for which evaluation has no warrant); or find a methodological strategy for valuing all and for making transparent the various contests. Democratic Evaluation chooses the latter. Its self-denial in respect of recommendations, its procedural disinterest in respect of outcomes creates an ethical space within which the evaluator can generate the trust and the safety which encourages people to engage in open information exchange. The democratic evaluator seeks not to change the world – that is the job of the programme – but to understand it. Democratic Evaluation and case study represent the fusion of political, ethical and methodological values into a single strategy.

Personalising Evaluation

Stake's Responsive Evaluation was designed to focus on the experience of the programme and to use this as a source for theorising about the programme. He chose the term 'responsive' to suggest an inversion of the conventional approach in which the evaluator emits a stimulus (question) to which the programme responds. Rather, he argued for the programme as the original stimulus to which the evaluation responds (methodologically). This was but another route to case study and reflects the same values and intentions of Democratic Evaluation – in fact, Stake and MacDonald were key members of the transatlantic group of evaluation theorists who brokered the emergence of case study.

One way of elaborating the methodological and political challenge of responsive and democratic evaluation is through the personalisation of evaluation, an approach which combines methodological individualism with democratic purpose (Kushner, 2000). Here is that same inversion. The intention is to invert the relationship between programme and person in such a way as to capture a more authentic view of the significance of a programme and its impact. This inversion can be represented simply in diagrammatic form:
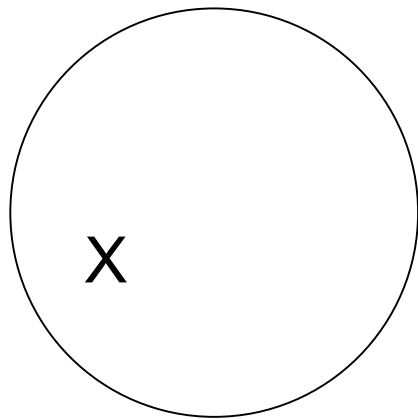
Conventionally we portray programmes as context and locate people within them – the circle is the programme; the X the person (or group) or event. Those persons are read, typically, as programme-people – they are 'teachers', 'patients', 'officials', 'pupils', for example, 'read' in terms of programme meaning, assigned programme-related roles. People tend not to be portrayed as 'lovers', 'mothers', 'sisters', 'lonely', 'fulfilled' or whatever. There are a number of effects, here, the principal one being that we lose any sense of scale in attributing significance to a programme. In fact, we are vulnerable to over-emphasising significance through a relentless focus on programme – to the point, often, where we create the impression that a programme may be the most significant event in certain people's lives (empirically unlikely other than for those who conceive of and are accountable for a programme and even there experience suggests that programme experience is a surrogate for other ambitions). We may sometimes be forgiven for forgetting that people go home after programme events. Nor does the conventional approach recognise this possibility, for example…..
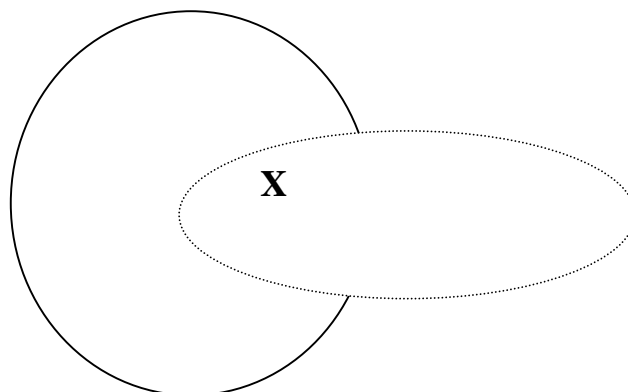
….that the person or event being observed may be more meaningfully located elsewhere, in another programme, say, and only incidentally or accidentally lie within this frame of reference. Think, for example, of a young person identified as 'at risk' or 'vulnerable'. It is likely that this young person will be touched by many reform initiatives – youth justice, health, education, social work, employment, family-centred. In a world of 'joined-up' or 'whole-of-government' approaches to professional reform we risk being incoherent by focusing on programme exclusively – since coherence is only achieved by the youth who seeks to make sense of the multiple innovations bearing down upon them.

The aim of Personalising Evaluation is to portray life and work as context within which to 'read' the significance of a programme – the circle becomes the work and life of the individual (or group) and the X stands for programme. This admits the democratic intent - not only to hold people to account for realising the aims of policy, but also to hold policy to account for its capacity to support the aspirations and needs of professionals and citizens. It reflects the aims of Democratic Evaluation and case study to draw closer to respondents so as to enter into collaborative theorising relationships with them – to properly articulate their voices (be they ministers, teachers or pupils) and to invite them to locate the programme in the context of their hopes, fears, expectations and abilities. As people talk to the evaluator about their life and work the conversation will be generating criteria against which, in due course, to assess the significance and meaning of the programme and to make judgements about its merits. Such conversations allow us to see not just what aspirations and efforts people bring to a programme, but also the source of those dispositions in experience, beliefs and values. People can define themselves independent of the programme, as they were before the programme came along and will be when it has passed.

There is no intention, here, to subvert our evaluation contracts, to avoid our responsibilities as evaluators for feeding judgements as to the merit or worth of a programme. To the contrary, the aim is to draw closer to a proper and adequate accounting for programme outcomes as they are rooted in values, struggles, beliefs and disbeliefs, allegiance and betrayal. Programmes of innovation are never blind leaps into uncertain futures. They are always rooted in personal and group histories, expressions of long-held ambitions, momentary alliances of people at all levels whose association represents a confluence of biographies. Even where it is a central policy which spawns an innovatory programme, the people it attracts to realise its aims are those for whom the programme promises a suitable vehicle. All programmes have values as engines.

The personalisation of evaluation is aimed at articulating those voices and portraying those events which commonly elude programme evaluation. This is not confined to the views and the work of the powerless. Too often the powerful are also marginalised in evaluation, portrayed in ways which prevent us from theorising about how their values, too, are mediated through their practices; how their aspirations may be frustrated by institutional constraints. All participants in a programme have the right of access to evaluation; all stakeholders have a right to know how the programme functions through actors at all levels.

Democratic engagement and social capital

Evaluators are drawing closer to community, making their services available there. Perhaps this is a symptom of the growing distance between western governments and their electorates, a reaction by communities and professions against centrally-imposed, coercive programmes – an attempt, perhaps, to recapture the change initiative for the grass-roots. As the notion of social capital comes to describe more what we want, so we are tempted to engineer more of it through community projects, building community self-determination. For whatever reason, there seems to be a proliferation of, often small-scale, studies evaluating community development projects. Here, the evaluator has to make no effort to discover desirable agendas. These are often values-driven projects which grow directly out of need and are neither mediated through the logic of government nor their ethical purposes overwhelmed by distantly-conceived targets. They are often more or less pure well-springs of humanism.

Here is where evaluation is likely to enter into its next manifestation – in relation to social capital. This may be the next move for Democratic Evaluation. Where evaluation has hitherto been steeped in accountability discourses at one remove from the citizenry, this community engagement may see a shift to more direct service to the citizen. A key question is whether evaluation enhances or diminishes social capital – ie. whether it enhances or diminishes self-determination in community. I will close by focusing

on two dangers in evaluation engagement one, too distant (distal) an engagement; the other, to close an engagement.

Distal engagement: Social capital is made up of collective resources which contribute to life quality in a community and which enhances self-determination. As governments realise this as a component in their struggle to raise the social indicators, so improving social capital becomes enmeshed in social engineering strategies. Then it needs measuring for its productive value and, once having been measured, it is susceptible to target-setting and it loses its roots in community endeavour. Health evaluators have already been asked to start this process by, for example, developing 'quality of life' indicators which guide medical practitioners in rationing treatments. Here, evaluators seek engagement in order to generate the data they have been asked to collect to support distant agendas and to enhance distant, not proximal, control.

Too close engagement: It is not hard for evaluators – especially those who have recently come from harsher evaluation contexts – to be coopted into the value system of these community projects and into the very assumptions and strategies they are supposed to be evaluating. In these contexts engagement is often demanded by the community – demanded as a token of faith and as an expression of trustworthiness and solidarity. It is, indeed, sometimes sought by evaluators as a strategy for ensuring that their evaluations are more socially meaningful, are part of the sought after social capital – and why not? In a world where there seems to be a diminishing wealth of creative, humanist and responsible solutions, is it not the obligation of the evaluator to promote those they find? So here, evaluators seek engagement to intensify their defence of local solutions.

Where, on the implied spectrum, does the democratic evaluator lie?

Here is a series of statements which outline the meaning and implications of democratic engagement. You will see that mostly they confer legitimacy on what might, under other social conditions, be thought to be disreputable activities and relationships. Evaluators share something with priests and prostitutes – that part of their social warrant requires them to fall in and out of love rapidly and strategically – but with no disrespect to their clients or dishonour to themselves. In no particular order:

- People deserve to be properly understood and this will often demand the kind of intimate knowledge which comes from close relationships. This enables – sometimes requires – the evaluator to work through field-based friendships – but always in pursuit of data. *Engagement involves strategic intimacy*;
- The evaluator must at least appear to be the friend of all participants – ie. show no favour, accept as reasonable and legitimate all claims and dispositions. *Engagement should be impartial and homogeneous*;
- Evaluation friendships are legitimately intensive but ephemeral ('fleeting intimacy' said Jennifer Greene to me) – they are, as it were, inauthentic-but-real, binding for as long as they last, but limited by purpose. There is inevitable betrayal (the evaluator will write up, and perhaps publish, the products of these friendships), but betrayal cannot involve harm. *Engagement must be respectful*;
- We should not evaluate projects whose survival as useful experiments we would not support. This does not mean buying into their values or their strategies, though it does mean we have no warrant to do anything which might compromise their existence or esteem. *Engagement should be circumspect*;
- We are obliged to defend projects we evaluate to the point – and no further – where they can withstand public scrutiny and critique. *Engagement should be safe but productive of public insight*;
- The evaluator must not allow portrayals of communities and their people, nor their claims which feature in the evaluation, to be undermined by public beliefs that the evaluation was 'in the pocket' of the evaluated community and so lacking in credibility. *Engagement must not erode the undeniability of the evaluation*;
- The democratic evaluator has all the time to be balancing the participant's right to privacy with the public right to know about the project or programme. Engagement is a strategic process for achieving that balance. *Engagement is founded upon negotiation*;
- The evaluator must honour the strivings of their respondents, but not seek to embrace those strivings in the evaluator's own value system. *Engagement is respectful of respondent autonomy*.
- The evaluator must formally appreciate the aspirations and efforts of programme people but hold back from endorsement of their objectives or outcomes. *Engagement should be mediated through independence*.

The principal challenge, here, is how to achieve Peter Berger's (existential) state of 'distance from the object of one's passions'. The accident of what may be our personal passion for the educational or social worth of a programme is irrelevant to the need to make available that programme's experience for public learning. Indeed, where the evaluator expresses a preference or declares a value or an interest in relation to the programme's work, this invites dismissal of their report on grounds of partiality. Any recommendation made by the evaluator, for example, will inevitably favour one interest group over others. This, in a sense, characterises the maturation of evaluation as a discrete field of knowledge and action (House, 1993) – when concepts of impartiality and political independence emerged, when formal disinterest became a role definition. It was Lawrence Stenhouse (1980: 260) who, as a subject of evaluation, put it in characteristically succinct terms: *"instead of discriminating between alternative courses of action [democratic evaluation] seeks to make actors more discriminating"*. The democratic evaluator cares too much for his or her respondents to take their side.

Acknowledgments

References

Berger, P (1974) *Pyramids of Sacrifice: Political Ethics and Social Change*, London: Allen lane

House, E (1993) *Professional Evaluation*, London: Sage

Kushner, S (2000) *Personalising Evaluation*, London: Sage

MacDonald, B. (1976) 'Evaluation and the control of evaluation' in D.A. Tawney (Ed.) *Curriculum Evaluation Today: Trends and Implications*, Schools Council Research Studies, London: Macmillan

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

Parlett, M & Hamilton, D (1977) 'Evaluation as illumination: a new approach to the study of innovatory programmes' in D. Hamilton, D. Jenkins, C. King, B. MacDonald, M. Parlett (Eds) *Beyond the Numbers Game: a Reader in Educational Evaluation,* London: Macmillan, pp. 6 - 22

Pawson, R & Tilley, N (1998) *Realistic Evaluation*, London: Sage

Stake, R.E (1967) 'The countenance of educational evaluation' in *Teachers College Record*, 68:7, pp. 523-540

Stake, R.E (Ed.) (1975) *Evaluating the Arts in Education: a Responsive Approach*, Columbus, Ohio: Charles E, Merrill

Stenhouse, L (Ed.) (1980) *Curriculum Research and Development*, London: Heinemann

Walker, R (1980) 'Making sense and losing meaning' in H. Simons (Ed.) *Towards a Science of the Singular*, CARE Occasional Publications No. 10, Norwich: University of East Anglia, pp. 222-235

Weiss, C (1983) 'The stakeholder approach to evaluation: origins and promise' in A.S. Bryk (Ed.) *New Directions for Programme Evaluation*, No. 17, London: Jossey Bass

Weiss, C & Rein, M (1969) 'The evaluation of broad-aims programmes: a cautionary case and a moral' in *Annals of the American Academy of Political and Social Science*, 385, pp. 133-142