Directions in Qualitative Evaluation

ian shaw

It is regarded by many as not far short of bad taste to advance passionate claims based on the superiority of this or that methodology. The argument of most mainstream evaluation theorists is for a 'horses for courses' approach that aims to identify the strengths of different methods and discourage evaluators from over-claiming the relevance and application of any one approach to evaluation. I use this article to develop a few outline arguments in support of turning on their heads some conventional arguments about methodological choices for evaluation. I touch on four areas where qualitative methodology enables evaluators to re-cast central aspects of evaluation practice, viz causal understanding, methodological choice, the evaluation of professional practice, and the uses of evaluation.

Introduction

It is regarded by many as not far short of bad taste to advance passionate claims based on the superiority of this or that methodology. The argument of most mainstream evaluation theorists is for a 'horses for courses' approach that aims to identify the strengths of different methods and

Ian Shaw is Reader in Social Work at Cardiff University, School of Social Sciences, Cardiff (Wales). He is the author of Qualitative Evaluation (1999 Sage) and co-editor of the journal Qualitative Social Work (Sage).

discourage evaluators from over-claiming the relevance and application of any one approach to evaluation. For example, if the evaluator is operating with an evaluation-as-accountability perspective – eg measuring results or efficiency – then the randomised, controlled, clinical trial (RCT) provides the 'gold standard, the Rolls Royce of evaluation approaches' (Chelimsky, 1997: 101). Qualitative strategies, for example case study designs, should not be used to tackle questions and problems which are the province of quantitatively-oriented methods. However, if the perspective is

one of evaluation for greater knowledge and understanding of a given policy or programme, then qualitative approaches may be the method of choice.

I have some sympathy for this position. Even the best arguments for pulling the coach of evaluation methods behind the horses of some favoured paradigm are in the end self-defeating, and I do not intend to use scarce space rehearsing the reasons for this conclusion[1]. In that sense I agree with the sociologist David Silverman when he says 'there are no principled grounds to be either qualitative or quantitative in approach. It all depends on what you are trying to do' (Silverman, 1997b:14). Demonizing positions with which we disagree is a fool's errand. It entraps us in sentimentality and superstitious practices (Shaw, 1999). I like the story of Philip of Macedon who apparently employed a man with a stick, atop of which was a pig's bladder. The sole function this man fulfilled was to exercise the freedom to wake Philip at any time of the night, and beat him about the head with the bladder, as a reminder that he was mortal. We forget our methodological mortality at our peril. There are, as Lather remarks, 'no innocent positions' (Lather, 1991: 85), and we need to 'protect our work from our own passions' (Lather, 1986: 77).

And yet the horses-for-courses approach of the evaluation mainstream imposes a premature closure on questions that are far from closed, and thus reinforces the status quo, some issues remaining forever out of court for methodological challenge from other quarters. In practice this operates as a line of defence for traditional evaluation methodologies. Evaluating outcomes, understanding causal processes, providing confident generalizations, and so on, remain the territory of just some evaluators, and 'merely serve to sanctify one perspective at the expense of another' (Chelimsky, 1997: 108). Boundaries thus become fixed at some arbitrary point in time. Why should we agree to fix a division of labour that is the simple happenstance of the century's turn? This is likely to perpetuate unhelpful hierarchies. If Chelimsky is right when she concludes that RCTs are the Rolls Royce then maybe case studies are a Trabant or at best a rather quaintly English Morris Minor.

I want to use this space to develop a few picture-board outline arguments in support of turning on their heads some conventional arguments about methodological choices for evaluation. I will touch on four themes, viz causal understanding, methodological choice, the evaluation of professional practice, and the uses of evaluation.

My general position is that evaluation is characterized by a cluster of evaluative purposes. These enable plausible and productive responses to questions such as methodological choice, evaluation theorising, evaluation ethics, and advocacy evaluation. Evaluation is best understood as entailing the conduct of evaluative research rather than a discrete set of evaluation axioms or methodology separate from the wider research enterprise. In this regard, evaluation theory and methodology owe almost

---

1 have attempted this in some detail in Shaw (1999) Qualitative Evaluation London: Sage. Eg Chapter 3.

as much to work undertaken by writers and researchers who would not regard themselves as evaluation theorists, as it does to confessedly evaluation theorists.

The boundary fence between research and evaluation is not the only one that needs dismantling. A relevant and rigorous evaluation requires the development of inter-professional evaluation theorising and strategies, such that education, health, criminal justice, law, human services and so on, are mutually attentive.

Evaluation is more – much more – than programme evaluation. Qualitative evaluation promises distinct but coherent perspectives on policy, programme and practice evaluation. Qualitative evaluation offers credible partial solutions to problems of causal analysis and outcome evaluation. It also enables us to avoid an unduly instrumental and rational approach to the uses of evaluation[2].

I don't want to be misunderstood as to the scope of the claims I am trying to support. It would be silly and self-defeating, for example, to leave the impression that qualitative evaluation should be the order of the day for all evaluative purposes. Rather, I hope to unsettle unquestioning faith in the evaluative benefits of some forms of mainly quantitative evaluation – those in particular that place undue confidence in

the possibilities of controlling and precisely measuring independent and dependent variables.

I am not convinced, either, that all forms and traditions of qualitative methodology lend themselves equally or even directly to evaluative purpose. This would simply replace one variety of uniformitarianism with another, through a tendency to treat qualitative methodology in an unduly homogenous way. As a corrective to this, I believe there is a need to develop the case for a dialectical mix of methods within qualitative research. This will need to proceed through the development of a set of critical features of knowledge for different qualitative methodologies. A helpful starting point for this is the paper by McKeganey and colleagues, in which they discuss the benefits and limitations of interviewing and observation methods as part of a study of professional decision-making when people may be offered a place in a home for the elderly (McKeganey et al., 1988). This initial analysis needs to be extended to the full range of qualitative strategies, and tied to the critical features of their associated knowledge claims (c.f. Greene and Caracelli, 1997: 12-13). I do not attempt this task here, but it is with such considerations in mind that I have deliberately avoided giving overall definitions of what constitute qualitative or quantitative studies, but have restricted myself to a number of illustrative examples, mainly from the fields of interpretive sociology, ethnography and case studies.

---

2 This paper has been written primarily to an evaluation audience. I have corresponding misgivings about the ways in which a majority of mainstream qualitative research academics are silent on the evaluation relevance of their field. For a recent example of the dog that does not bark, see Atkinson, Delamont, Coffey, Lofland and Lofland (2000).

Understanding causes

The conventional division of labour is that qualitative inquiry is useful for generating hypotheses/questions, and describing processes, while quantitative and more statistical designs are needed to analyze outcomes and verify hypotheses. Miles and Huberman summarize this view – from which they vigorously dissent - as:

qualitative studies are only good for exploratory forays, for developing hypotheses - and … strong explanations, including causal attributions, can be derived only through quantitative studies.
(Miles and Huberman, 1994: 147)

This is too limiting. Neither quantitative nor qualitative evaluation can solve questions of cause and effect in a straightforward way. Led by recent work on realist evaluation, changes have taken place in thinking regarding the nature of cause, and the corresponding models of causal hypotheses which flow from that thinking. The central idea is that there are underlying causal mechanisms which cannot be understood by surface workings and measurement. Hence, 'events themselves are not the ultimate focus of scientific analysis...Reality consists not only of what we can see but also of the underlying causal entities that are not always discernible' (House, 1991: 4). The underlying reality produces actual events, of which we have empirical experiences and sense impressions.

This is often described as a generative concept of causality.

When we explain an outcome generatively we are not coming up with variables or correlates that associate with one another; rather we are trying to explain how the association itself comes about. The generative mechanisms thus actually <u>constitute</u> the outcome. (Pawson and Tilley, 1997: 408) [italics in original]

The conventional concept of causation as regularities and associations is dismissed in favour of causal entities which have 'tendencies interacting with other tendencies in such a way that an observable event may or may not be produced' (House, 1991: 5). House quotes Manicas and Secord who say that, 'For the standard view of science, the world is a determined concatenation of contingent events; for the realist, it is a contingent concatenation of real structures. And this difference is monumental'. Hence, instead of merely documenting the sequence and association of events, the realist seeks to explain events.

While this view of cause does not necessarily require a qualitative methodology, it does clearly lend itself to such methods.

Qualitative studies are not designed to provide definitive answers to causal questions...(but) it can still be an appropriately qualified pursuit. (Lofland and Lofland, 1995: 136, 138)

Miles and Huberman are even less reserved. They describe the conventional view as 'mistaken' (Miles and Huberman, 1994: 147), and insist that qualitative evaluation research is well equipped to,

1. Identify causal mechanisms.
2. Deal with complex local networks.
3. Sort out the temporal dimension of events.
4. Cycle back and forth between different levels of variables and processes.
5. Provide a way of testing and deepening single case explanations through analytic induction.

Causal accounts will be local and 'now-oriented' (Lofland and Lofland, 1995: 141). Miles and Huberman develop analytic methods which address causal attribution in both single and multiple case explanations. For example, they advocate the use of field research to map the 'local causal networks' which informants carry in their heads and to make connections with the evaluator's own emerging causal map of the setting. Such maps start from 'causal fragments' which lead on to linked building of logical chains of evidence. Such causal networks

are not probabilistic, but specific and determinate, grounded in understanding of events over time in the concrete local context - and tied to a good conceptualisation of each variable. (Miles and Huberman, 1994: 159)[3]

Much of this reasoning was anticipated by Cronbach's arguments regarding causal

models. Rejecting the idea of causation as events that can be predicted with a high degree of probability, Cronbach developed twin arguments. First, he argued that causes are contingent on local interactions of clusters of events. More than one cluster may be sufficient, but no one cluster is necessary. Second, he accepted that there are usually missing events or conditions that affect the outcome of a given programme, but about which we know little. He was the first evaluation theorist to produce a plausible explanation of contextual factors in evaluation. Hence he concludes that 'after the experimenter with his artificial constraint leaves the scene, the operating programme is sure to be adapted to local conditions' (Cronbach et al., 1980: 217). Furthermore, 'a programme evaluation is so dependent on its context that replication is only a figure of speech' (p. 222).

Qualitative evaluation cannot resolve the problems of causal conclusions any more than quantitative evaluation, but it can assess causality 'as it actually plays out in a particular setting' (Miles and Huberman, 1994: 10).

Lofland and Lofland make the important observation that causal answers are by and large based on passivist conceptions of human nature. Qualitative inquiry has often steered away from causal accounts, not because the methodology is weak in

---

3   An interesting connection can be drawn between Miles and Huberman's assessment and current work, mainly in American evaluation writing, on logic models. A qualitative methodology stance on logic models suggests the potential value of inductive 'logic models'. I see the existing USA approach as a way of com-

bining a consensual stakeholder input with outcome analysis. The drawback of this approach is that it under-emphasizes the likely persistence of variant informal logic models. Inductive, informal logic models offer a use of 'logic' in a sense not far from ideas of frames of meaning as used in Anthony Giddens and others.

that area but because of a commitment to an activist conception of human nature. The Loflands argue that an activist conception will lead to a focus on questions that address both structures and strategies. This will involve 'deciphering and depicting exactly what sort of situation the participants are facing' (Lofland and Lofland, 1995: 146), and understanding the 'incessantly fabricated' strategies people construct to deal with the situation.

Take for example, Silverman's work on HIV counselling. He is right to conclude that 'it is usually unnecessary to allow our research topics to be defined in terms of...the »causes« of »bad« counselling or the »consequences« of »bad« counselling' (Silverman, 1997a: 34), insofar as such topics reflect the conceptions of social problems as recognized by professional or community groups. Nonetheless, this does not require the abandonment of causal inquiry in qualitative evaluation. Inquiry into the ways in which professionals incessantly fabricate service forms and structures does promise a better way to understand causes.

By way of illustration, Shaw and colleagues describe a case study evaluation

Figure 1.

Stakeholder models of a Rural Activity centre

	Training for Work	Personal/social growth	Education for life
Aims	Credible work skills for independent/ sheltered work	Personal and social growth	Alternative occupation to enhance the quality of life
Target group	Demonstrable ability to benefit; younger	Wide range of age and ability	Wide ability range; younger
Programme	Time limited stay; skill learning; assessment and review; contracts; move-one facility; integration into work	Open stay period; small project; small-group activities; counselling; liaison with carers and social work agencies	Loosely held time limits; the best learning context; interest-led contracts; community based activities and outside links; craft work and homemaking skills
Staffing	Education and special needs employment skills; plus volunteers	Social and group work qualifications; plus expert consultants	Education and social work qualifications; plus volunteers
Outcome	Regular throughput; work placements; normalization of work patterns; skill learning	No clear distinction between programme and outcome	Wide range of social skills; integration into community networks; change of attitudes on the part of outside community members

Ian Shaw: Directions in Qualitative Evaluation

of a rural activity centre for people with learning disabilities. They observed and interviewed project participants, parents, carers, management group members, key workers and other professionals. Project records were analyzed.

When describing and explaining the workings of the centre, the people who were interviewed appeared to draw on one or more of three different models of the scheme. These were a 'training for work' model, a 'personal and social growth' model, and an 'education for life' model. These operated in part as causal maps which entailed an array of model-specific positions on the aims of the project, optimal target groups, desirable programme patterns, staffing requirements, future development strategies, and likely or desirable project outcomes[4].

Choosing Methods

Qualitative evaluation methodology is not only well equipped to address the local outworking of cause and effect, but also enriches the choice of methods relevant to evaluative purposes. The inter-relationship of qualitative and quantitative methods is

---

4 Rethinking approaches to understanding causal processes leads naturally to rethinking the ways in which it is possible to generalize from one program and its evaluation to another program. Qualitative researchers and evaluators have developed partial answers to this question through ideas about vicarious experience (Stake and Trumbull. 1982), transfer (Eisner, 1991), 'thick description' (Geertz, 1973) and analytic generalisation.

not only, nor even primarily, about choice of methods. It is about the questions in Figure 2 and is also inextricably relevant to issues of the politics and purposes of social work research, values, participatory forms of research, interdisciplinary research, and the uses of research.

Figure 2.
Qualitative and Quantitative Methodology

Single cases or comparison.
Cause and meaning.
Context as against distance.
Homogeneity and heterogeneity.
Validity and the criteria of quality in social work research.
The relationship of researcher and researched.
Measurement.

Qualitative inquiry may shed light on programme outcomes in ways that are less susceptible to quantitative methodology. Miller, for example, discusses ways that institutional texts constructed to explain past decisions inevitably gloss over the openness and complexity of the decision-making process (Miller, 1997). He gives the mundane example of evaluation research on a bowel-training programme in a nursing home. The evaluation consisted of counting when and how patients had bowel movements. The programme was judged to have a successful outcome if patients used a toilet or bedpan and ineffective for those who continued soiling beds. One patient had soiled her bed. However, observation methods enabled the researcher to view a nursing aide contesting the definition of this as 'failure' on the grounds that the patient knew what she was doing and had

soiled her bed as a protest act against staff favouring another patient. This illustrates how observing the context of text construction illuminates mundane, everyday life. This would not have found a way into the formal outcome record. Text production in institutions is 'micro-politically organized', and this includes textual outcome records.

A further illustration of the relevance of qualitative methodology for outcomes evaluation can be traced through the surprising impact of Denzin's interpretive interactionism (Denzin, 1989 and 2002; Mohr, 1997). Mohr, for example, extends Denzin's argument to the evaluation of clinical outcomes in health research. She argues that the method leads us to inspect the relationships between personal difficulties, experiences, policies, interventions, and institutions. 'Interpretive interactionism permits intensive scrutiny of the ramifications and outcomes of various interventions' (1997: 284). It can:

1. Sort out different ways problems are defined.
2. Show how patients experience care. What it is about interventions they find helpful or not, and in what circumstances.
3. Identify 'secondary causes' eg contexts, culture, and the meanings patients bring.

'Strategic points for intervention can be identified by contrasting and comparing patients' thick descriptions, and these can be used to change, to improve, or to negotiate and renegotiate interventions' (p.284). It is valuable when 'an outcome may not be readily apparent, and…the intervention is something that only the patient and not the professionals can define' (p.285).

Constructive, if cautious, dialogue regarding the relative merits and characteristics of quantitative and qualitative methodologies has emerged more recently. The social work literature provides a useful example. From the quantitative side of the case, Reid in the USA and Sinclair in Britain have developed mediating positions. Reid seeks to 'redefine the nature of the mainstream so that qualitative methodology is a part of it not apart from it'. He regards quantitative research as strong when dealing with linkages, control, precision, and larger data sets, while qualitative research is able to depict system workings, contextual factors, and elusive phenomena, and provide thorough description. 'Neither method is superior to the other, but each provides the researcher with different tools of inquiry' that can be set against a single set of standards (Reid, 1994: 477).

Sinclair adds to Reid's conclusion, in his discussion of randomized control trails (RCTs), when he says that qualitative methods are in many ways 'more adapted to the complexity of the practitioner's world than the blockbuster RCT'.

Qualitative research draws attention to features of a situation that others may have missed but which once seen have major implications for practice. It counteracts a tendency to treat the powerless as creatures with something less than normal human feelings. It contributes to an ethically defensible selection of outcome measures. And, in combination with simple statistical description, it can lead to an informed and incisive

evaluation of programmes in social services. (Sinclair, 2000: 8)

He turns common assumptions on their head when he concludes that,

Quantitative social work research does face peculiarly acute difficulties arising from the intangible nature of its variables, the fluid, probabilistic way in which these variables are connected, and the degree to which outcome criteria are subject to dispute. (pp. 9-10)

Evaluating professional practice

Qualitative methodology also provides a strong purchase on the evaluation of direct service delivery. The main point I wish to make is that thinking and practice in the evaluation field have been too much influenced by ideas of evaluation as being equivalent to programme evaluation (indeed, often as equivalent to programme outcome evaluation). I believe we need to distinguish more strongly in our thinking between evaluation of policies, programmes, projects within programmes, and direct practice and service delivery. I am especially interested in the last of these. Rather than see good practice as being subject to evaluation – whether internal or external evaluation is immaterial – I believe good practice should in and of itself entail evaluation for and with service users. Influences on my position include:

- Reflective practice.
- Qualitative methodology.

- Empirical evidence on evaluation as a dimension of professional practice.
- Advocacy evaluation and user-led research.
- Action research.

The model I have tried to develop, and have tested out in some measure with practitioners, is premised on a composite image of good practice as requiring evaluative evidence, evaluative learning, and evaluative justice. It draws, therefore on what in the UK is increasingly being called 'knowledge-based practice'; on the learning organization and reflective learning literature; and on advocacy models of evaluation. The main practical approach I have developed has been to seek to 'translate' and 'colonize' methods, especially, but not exclusively, qualitative ones (eg Shaw, 1996, 1997). Examples include life histories, simulations, focus groups, narrative methods, cultural reviews, inductive local logic models, and peer interviewing.

Empirical work has shed some light on how practitioners seek to make sense of and resolve evaluative issues in their day-to-day work. Elks and Kirkhart urge that 'an alternative research model is needed, one that is exploratory rather than confirmatory, building a model of evaluation from the practitioner's own accounts rather than superimposing an ideal model and testing for conformity' (Elks and Kirkhart, 1993: 555). They interviewed seventeen social workers asking them how they evaluated what they were doing, and how they knew whether they were doing a good or bad job. Practitioners acknowledged difficulty in knowing if they were

effective. They also perceived an incompatibility between the roles of evaluator and practitioner. The researchers suggest that practitioners hold an implicit model of practice evaluation which they describe as a 'pragmatic-professional model'. This included a reliance on intuition and experience, an internalised notion of an ideal practitioner, a dependence on feedback from colleagues, friends and family, and a model of an ideal client which always included growth and change.

Humphrey and Pease conducted a corresponding study in which they interviewed British Probation Officers regarding their perceptions of probation effectiveness. Probation officers tended to 'de-couple' the process of supervision from the out-turns of probation. Thus one person said, clients 'frequently get into more trouble but I don't think that in any way is a reflection on whether or not I have been effective'. Indeed, there was widespread belief that an element of luck operated in being effective. One might do 'brilliant' work but if the circumstances are against you they will still re-offend. Thus, 'if luck is seen to determine outcome, probation supervision becomes merely a matter of keeping an offender in the community for luck to strike' (Humphrey and Pease, 1992: 40).

Subsequent work by Shaw and Shaw suggests that social workers appear to have two contrasting models of evaluation in their heads - a formal 'evaluation proper' and self-evaluation. Formal evaluation is experienced as largely alien to the realities of social work and in almost complete contrast to social workers' evaluative 'maps' of their actual day-to-day evaluating (Shaw and Shaw, 1997b). Evaluation strategies were constructed from a 'game plan', the success of which was viewed - as in Humphrey and Pease' research - as partly contingent on the untoward operation of 'sheer luck'. Social workers judged their practice according to whether their work produced emotional rewards; the case was 'moving'; intervention won steady, incremental change; practice was accomplished without inflicting harm through the operations of the welfare system, and confirming evidence was available from fellow professionals. These practitioners were preoccupied with causes and reasons for outcomes of their work, held strongly worked views about the complexity and ambiguity of social work evidence, and were aware of the constant interplay of knowing and feeling in practice (Shaw and Shaw, 1997a). The significance of emotions echoes Erikson's remark of the clinician that

The evidence is not »all in« if he does not succeed in using his own emotional responses during a clinical encounter as an evidential source and as a guide to action. (Erikson, 1959: 93)

The emphasis of these previous paragraphs has been on evaluation as endemic and taken-for-granted within professional practice, and for the considerable gains derived from rendering it visible as a step to embedding it at the core of good practice. This embedding will be further promoted if we avoid over-simple distinctions between insider and outsider evaluation. A drawback of much practitioner research

stems from a tendency to regard 'practice' as distinct from theory, and hence to regard 'being theoretical' as something that happens in the mind and the 'practical' as having a derivative, 'applied' relationship to these guiding ideas. Conversely to this, good practitioner evaluation ought to give attention to exploring different kinds of tacit knowledge. The significance of personal contact and practical knowledge sharing between practitioners will be brought out, and sources of trust and mistrust between social workers made clear (Collins, 2000). This is a big agenda, and one that social work and other occupations have only begun to tackle. One consequence is that we cannot hang on to a narrow distinction between practitioner research as being 'insider' evaluation and 'academic' research as being 'outsider' evaluation. In a recent analysis of qualitative social work research, several of the contributors reflect on these issues. Hall and White, for example, record how they held both insider and outsider roles in relation to their research participants. Hall 'arrived' as an outsider but became in different ways a partial insider (Hall, 2001). White started as an insider, yet found herself undergoing a fruitful, if potentially hazardous, process of de-familiarisation through which she became in some degree a marginal 'inside »out« member (White, 2001). In the same volume, Scourfield focuses his reflections on the research and practice relationship through his consideration of what it was like to interview expert professional social work interviewers (Scourfield, 2001).

In what ways do we expect evaluation to be useful?

Finally, a focus on qualitative methodology leads fairly directly to the wider question of how evaluation might be useful for policies, programmes, projects and professional practice. At the broadest level, evaluation would be judged useful if it demonstrably contributed to one or more of the following.

- Better policies, services and practice.
- Strengthened the moral purpose of professional practice.
- Promoted methodological rigour, scope, depth and innovation.
- Strengthened the sense of a profession's intellectual nature and location.

We tend to make simple distinctions between research that has a direct, applied purpose and research that is basic, and hence where direct use questions are less relevant. This does not work (just as the related sharp distinction between 'research' and 'evaluation' does not work either).

Conventional quantitative research on outcomes is linked to a confidence in the instrumental utility of research. The problem with this is that it does not square with evidence on how evaluation is actually used, and it misunderstands the nature of the policy making process. It is based on a rationalistic model. The rationalist model of policy making sees it as a series of discrete events, where each issue to be decided is clearly defined, and decisions are taken by a specific set of actors who choose between well-defined alternatives, after weighing

the evidence about the likely outcome of each (Finch, 1986: 149-150).

An important early figure for this question is Carol Weiss, whose work in the 1970s explored how political considerations intrude on evaluation. She addressed three related issues. First, she delineated the political context in which evaluation is located. Although she has been primarily concerned with evaluation and policy research at the federal level, her empirical work with policy and programme staff resonates throughout evaluation theory and practice. This underlines the importance of being clear about the audiences for any inquiry.

Second, she exposed the limitations of conventional instrumental views of the political use of information, through her conceptualisation of use as enlightenment. With her colleagues she interviewed 155 senior officials in federal, state and local mental health agencies. Officials and staff used research to provide information about service needs, evidence about what works, and to keep up with the field. However, it was also used as a ritualistic overlay, to legitimize positions, and to provide personal assurance that the position held was the correct one. At a broader conceptual level, it helped officials to make sense of the world. For all these purposes, 'It was one source among many, and not usually powerful enough to drive the decision process' (Weiss, 1980: 390). As for direct utilization of research, 'Instrumental use seems in fact to be rare, particularly when the issues are complex, the consequences are uncertain, and a multitude of actors are engaged in the decision-making process. (p. 397).

Research use was also reflected in officials' views of the decision-making process. Decisions were perceived to be fragmented both vertically and horizontally within organizations, and to be the result of a series of gradual and amorphous steps. Therefore, 'a salient reason why they do not report the use of research for specific decisions is that many of them do not believe that they make decisions' (p. 398). Hence the title of one of her papers - 'Knowledge creep and decision accretion'. This provided the basis for her conclusion that enlightenment rather than instrumental action represents the characteristic route for research use.

The enlightenment model 'offers far more space to qualitative research, through its emphasis on understanding and conceptualisation, rather than upon providing objective facts' (Finch, 1986: 154).

Third, Weiss imbued models of use with a realistic view of the public interest. More than anything she has struggled towards a realistic theory of use. Others subsequently have developed such realistic views. For example, Chelimsky suggests that evaluation may have a deterrence function. 'In other words, the mere presence of the function, and the likelihood of a persuasive evaluation, can prevent or stop a host of undesirable government practices' (Chelimsky, 1997: 105).

Several cautious remarks are in order. First, we should not become over-pre-occupied with models of research use. Chelimsky believes 'it is often the case that….evaluations are undertaken without any hope of use'. Expected non-use is characteristic of some of the best evaluations, including 'those that question widespread popular

Ian Shaw: Directions in Qualitative Evaluation

155

beliefs in a time of ideology, or threaten powerful, entrenched interests at any time' (p. 105). Thus, 'there are some very good reasons why evaluations may be expert, and also unused' (p. 105). Chelimsky's comments are both sane and plausible.

To justify all evaluations by any single kind of use is a constraining rather than an enabling idea because it pushes evaluators towards excessive preoccupation with the acceptability of their findings to users, and risks turning evaluations into banal reiterations of the status quo. (Chelimsky, 1997: 106)

Second, the enlightenment model should not be adopted as universally appropriate. For example, practitioner research is likely to proceed on a more immediate instrumental view (see, for example, Figure 3).

Third, the adoption of enlightenment assumptions about research use can easily translate into a defensive posture, arising from the fact that they can readily be used to support an incrementalist approach to social change. The step from an empirical recognition that policy and practice change often proceed through incremental enlightenment to a tacit assumption that this is how social change ought to proceed may be logically insupportable, but it is deceptively easy.

Fourth, in counterpoint to the previous point, we should avoid being unduly sanguine about the ability of research to change practice. Hammersley has criticized some professionally-driven research approaches on the grounds that they are based on too narrow a concept of research relevance and an overly optimistic faith in the ability of research to influence policy and practice. He suggests two grounds for concluding 'there are good reasons to believe that research cannot routinely solve teachers' problems' (Hammersley, 1993: 430). 'There is no scientific method that guarantees results' (p. 430) and teacher circumstances are diverse and unlikely to be amenable to action in any routine sense. Rather, 'sound practice cannot amount to the straightforward application of theo-

Figure 3.

Evaluation models and information use

Model of use	Evaluation focus	Evaluation base	Discipline links	Time span for action
Enlightenment	Policy evaluation Programme development Advocacy evaluation	Higher education	Stronger discipline links Commitment to theorising	Longer term
Instrumental	Project evaluation; programme feasibility studies; practitioner evaluation	Agency sponsored; 'Insider' evaluation; self evaluation	Limited social science theorising	Immediate 'applications'

retical knowledge, but is an activity that necessarily involves judgement and draws on experience as much as on…scientific knowledge' (p. 430).

Fifth, there is often an important connection between research use and ethics. Consider the ethical and political dilemmas about how research material is used. These issues are sometimes sharper in qualitative research. This arises partly from the greater closeness and consequent trust that may develop between evaluator and participant. In quantitative research the greater distancing may make these issues less agonising. The risk of betrayal is also increased because of the typical use of smaller samples, the consequent difficulties of protecting the confidentiality of individuals, and the emphasis on the details of how people live their lives. Finch describes from her playgroups research her 'sense that I could potentially betray my informants as a group, not as individuals' (Finch, 1986: 207). 'Where qualitative research is targeted upon social policy issues, there is the special dilemma that findings could be used to worsen the situation of the target population in some way' (Finch, 1985: 117).

But none of these caveats reduces the overall value of careful elaboration of what we mean when we speak of research being useful. For instance, everyday discussions of evidence-based practice frequently proceed on a misconception at this point. Practice is not and cannot be 'based' on evidence in the straightforward and unproblematic way envisaged by many of its advocates.

I have argued in this paper that the field of evaluation has much to gain by holding back from a premature consensus approach to evaluation methodology. I have touched on four areas where qualitative methodology enables evaluators to re-cast central aspects of evaluation practice, viz causal understanding, methodological choice, the evaluation of professional practice, and the uses of evaluation. I could as easily have developed the same argument for other domains – quality standards in evaluation, evaluation ethics and governance, evaluation synthesis, the relation of evaluation to mainstream social science, interdisciplinarity, and so on. In all these areas I am convinced by Paul Feyerabend's provocative aside that 'it is not the puzzle-solving activity that is responsible for the growth of knowledge, but the active interplay of various tenaciously held views' (quoted by Trend, 1979:84).

References

Atkinson, Paul, Delamont, Sara, Coffey, Amanda, Lofland, John and Lofland, Lyn (2000) Handbook of Ethnography London: Sage Publications.

Chelimsky, Elinor (1997) 'Thoughts for a new Evaluation Society', in Evaluation, 3 (1): 97-118.

Collins, Harry (2000) Tacit Knowledge, Trust and the Q of Sapphire Cardiff University School

of Social Sciences. Working Paper. <http://www.cf.ac.uk/socsi/>

Cronbach, Lee, Ambron, S., Dornbusch, S., Hess, R., Hornik, R., Phillips, D., Walker, D., and Weiner, S. (1980) Toward Reform of Program Evaluation. San Francisco: Jossey-Bass.

Denzin, Norman (1989) Interpretive Interactionism, Englewood Cliffs, NJ: Prentice Hall.

Denzin, Norman (2002) 'Social work in the seventh moment' in Qualitative Social Work: Research and Practice 1 (1): 25-38.

Eisner, Elliot (1991) The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice. New York: Macmillan.

Elks, Marion and Kirkhart, K. (1993) 'Evaluating effectiveness from the practitioner's perspective' in Social Work, 38 (5): 554-563.

Erikson, Erik (1959) 'The nature of clinical inference' in Lerner, D. (ed) Evidence and Inference. Illinois: Free Press.

Finch, Janet (1986) Research and Policy: the Uses of Qualitative Methods in Social and Educational Research. London: Falmer Press.

Geertz, Clifford (1973) The Interpretation of Cultures. New York: Basic Books.

Greene, Jennifer and Caracelli, Valerie (1997) Advances in Mixed Method Evaluation: the Challenge and benefits of Integrating Diverse Paradigms. New Directions For Evaluation, No 74, San Francisco: Jossey-Bass.

Hall, Tom (2001) 'Caught not taught: Ethnographic research at a young people's accommodation project' in Shaw, I and Gould, N. Qualitative Research in Social Work London: Sage Publications.

Hammersley, Martyn, (1993) 'On the teacher as researcher', in Educational Action Research, 1 (3): 425-445.

House, Ernest (1991) 'Realism in research', in Educational Researcher, 20 (6) 2-9.

Humphrey, Chris and Pease, Ken (1992) 'Effectiveness measurement in the Probation Service: a view from the troops' in Howard Journal, 31 (2): 31-52.

Lather, Patti (1986) 'Issues of validity in openly ideological research', in Interchange, 17 (4): 63-84.

Lather, Patti (1991) Getting Smart: Feminist Research and Pedagogy with/in the Postmodern. New York: Routledge.

Lofland, John and Lofland, Lynn (1995) Analysing Social Settings. Belmont: Wadsworth.

McKeganey, Neil, MacPherson, I. And Hunter, D. (1988) 'How »they« decide: exploring professional decision making' in Research, Policy and Planning, 6 (1): 15-19.

Miles, Matthew and Huberman, A, (1994) Qualitative Data Analysis: an Expanded Sourcebook. Thousand Oaks: Sage.

Miller, Gale (1997) 'Contextualizing Texts: Studying Organisational Texts' in Miller G and Dingwall R (eds) Context and Method in Qualitative Research London: Sage.

Mohr, Wanda K. (1997) 'Interpretive Interactionism: Denzin's Potential Contribution to Intervention and Outcomes Research' in Qualitative Health Research 7 (2): 270-286.

Pawson, Ray and Tilley, Nick (1997b) 'An Introduction to scientific realist evaluation', in Chelimsky, E. and Shadish, W. (eds) Evaluation for the 21st Century. Thousand Oaks: Sage.

Reid, William J. (1994) 'Reframing the epistemological debate', in Sherman, E. and Reid, W. (eds), Qualitative Research in Social Work. New York: Columbia University Press.

Scourfield, Jonathan (2001) 'Interviewing interviewers and knowing about knowledge' in Shaw, I and Gould, N. Qualitative Research in Social Work London: Sage Publications.

Shaw, Ian (1999) 'Seeing the trees for the wood: the politics of evaluating in practice', in Broad, B. (ed), The Politics of Research and Evaluation. Birmingham (England): Venture Press.

Shaw, Ian (1996) Evaluating in Practice, Aldershot (England): Ashgate.

Shaw, Ian (1997) Be Your Own Evaluator: a Guide to Reflective and Enabling Evaluating. Wrexham (Wales): Prospects Publishing.

Shaw, Ian and Shaw, Alison (1997a) 'Keeping social work honest: evaluating as profession and practice', in British Journal of Social Work, 27 (6): 847-869.

Shaw, Ian and Shaw, Alison (1997b) 'Game plans,

buzzes and sheer luck: doing well in social work', in Social Work Research, 21 (2): 69-79.

Silverman, David (1997a) Discourses of Counselling: HIV Counselling as Social Interaction. London: Sage.

Silverman, David (1997b) 'The logics of qualitative research' in Miller, G. and Dingwall, R (eds) Context and Method in Qualitative Research. London: Sage.

Sinclair, Ian (2000) 'Methods and measurement in evaluative social work', paper from ESRC seminar series 'Theorising social work research'. National Institute for Social Work web site http://www.nisw.org.uk/tswr/sinclair.html

Stake, Robert and Trumbull, D. (1982) 'Naturalistic generalizations' in Review Journal of Philosophy and Social Science, 7, 1-12.

Trend, M.G. (1979) 'On the reconcilation of qualitative and quantitative analyses, in Cook, T and Reichardt, C. Qulitative and Quantitative Methods in Evaluation Research. Beverly Hills: Sage.

Weiss, Carol (1980) 'Knowledge creep and decision accretion', in Knowledge, Creation, Diffusion, Utilisation, 1 (3): 381-404.

White, Sue (2001) 'Auto-ethnography as reflexive inquiry: the research act as self surveillance' in Shaw, I and Gould, N. Qualitative Research in Social Work London: Sage Publications.