Primary and Secondary Barriers to the Evaluation of Change

Evidence from Two Public Sector Organizations

DENISE SKINNER

Oxford Brookes University, UK

For more than two decades it has been widely acknowledged within the management literature that the evaluation of change rarely takes place within organizations, yet little coherent explanation has been offered as to why this should be so. This article reports on research undertaken in two UK public sector organizations, to explore the reality of evaluation in the context of change, with the intention of identifying the barriers to evaluation that existed and the factors that created them. The analysis of the empirical evidence led to the identification of two distinct types of barrier to evaluation, labelled primary (factors that act against an evaluation being undertaken) and secondary (factors that arise during an evaluation process). Primary barriers were rooted in contextual factors contained in the organization's history and culture; the background against which the change initiative took place. Secondary barriers relate directly to the choices and decisions involved in the evaluation process itself and therefore only become significant once a decision to undertake an evaluation has been made. In combination, they offer an explanation for the exclusion of the evaluation stage in many change initiatives.

KEYWORDS: barriers; change management; evaluation; organizational change; public sector

Introduction

In today's business environment, the ability to change continually and successfully is widely considered to be essential to organizational survival. This need for ongoing change is argued to require an organizational ability to learn (Lahteenmaki et al., 2001) on a continual basis in a co-ordinated and progressive way to achieve and maintain competitive advantage (Mabey and Salaman, 1995; Pettigrew and Whipp, 1991; Senge, 1990). Jenlink (1994) links continuous learning processes with enhanced capacities for change and notes the importance of past experience and the transfer of knowledge as factors in organizational learning.

Others who link learning and change (e.g. Drew and Smith, 1995; Dixon, 1994; Hendry, 1996) highlight the importance of knowledge acquisition processes and the need for organizations to share knowledge and learning as a key means of maintaining competitive advantage (e.g. Pedler et al., 1991; Senge, 1990).

An integral part of both effective change and effective learning is the reflection on experience – the evaluation of process and outcomes that enables informed progression to the next stage. Pedler et al. (1991) include the conscious structuring of evaluation as a characteristic of learning organizations and the role of evaluation in successful change initiatives is widely acknowledged within the change management literature (e.g. Hollinshead and Leat, 1995; Owen, 1993; Salaman and Butler, 1994). Evaluation, it is argued, provides information that can:

- improve management decision making (Calder, 1994; Love, 1991);
- create new insights and mutual understanding (Preskill and Torres, 1999; Weiss, 1990);
- lead to wider acceptance and commitment to change initiatives (Carnall, 1995; Kirkpatrick, 1985);
- develop innovation (Forss et al., 1994); and
- provide opportunity for reflection prior to undertaking further change (Patrickson et al., 1995).

Failure to reflect on past experience allows valuable knowledge to escape and condemns both individuals and organizations to repeat the, often unsuccessful, past (Garvin, 1993).

Herein lies the mystery. If many of those providing and promoting the prescriptions for organizational success recognize the need for evaluation in the context of the need for continual learning and change, why does this not translate into organizational practice? For there is evidence that, in the majority of organizational change initiatives in the UK, little is being done to assess the impact and degree of success of such initiatives before organizations progress to the next (Doyle et al., 2000; Torraco, 1997). The inevitable consequence of this is 'broken learning cycles' (Drew and Smith, 1995) in which mistakes are repeated and valuable knowledge lost (Doyle et al., 2000). No clear explanation has been provided for why this happens. Why has the norm in organizations been – and continues to be – that change efforts should be announced with great fanfare, yet, no announcement is made about the final evaluations (Reichers et al., 1997)?

This article reports the findings of a study that sought to understand why the evaluation of change initiatives remains a relatively rare occurrence in organizations. The research was undertaken to identify obstacles or barriers to evaluation that might occur, the factors that contributed to their creation and their effect on the evaluation process in relation to such initiatives. The article begins with a description of the case study organizations and the research that took place. This is followed by an explanation of the factors that emerged from the cross-case evidence and analysis of the findings in the context of the relevant literature. The analysis leads to the identification of two distinct types of barrier

to evaluation, labelled primary (factors that act against an evaluation being undertaken) and secondary (factors that arise during an evaluation process), which, in combination, offer an explanation for the exclusion of the evaluation stage in many change initiatives.

The Research

A qualitative, exploratory, multiple-case approach was used to observe the evaluation of change initiatives in two public sector organizations in order to identify the nature of the barriers that arose and the factors that contributed to their creation. Public sector organizations were chosen in the belief that they might reasonably be expected to have a raised awareness of process and measurement issues that would have relevance to evaluation. Since the early 1990s there have been significant, high-profile changes in UK public sector management primarily aiming to improve efficiency and effectiveness (Ghobadian and Ashworth, 1994; Kouzmin et al., 1999). As part of these changes, there has been considerable emphasis on the need for objective setting, monitoring, control and accountability (Driscoll and Morris, 2001; Hyndman and Eden, 2000) and pressure to introduce comprehensive performance measurement (Ghobadian and Ashworth, 1994) to improve the efficiency and effectiveness of individuals and organizations. Kouzmin et al. (1999) argue that this has shifted the emphasis to means rather than ends, and so might reasonably be expected to have increased awareness of process and issues relating to implementation.

In each case, data were collected using a variety of methods: interviews; focus groups; observations; internal and published documents. In each organization, an initial set of stakeholders in the evaluation was identified who would be interviewed and these tended to be the individuals most visibly involved, usually those actually managing the evaluation process. Decisions about who to interview were made using stakeholder analysis (Burgoyne, 1994) to ensure that all useful contributors were included and that 'completeness' would be achieved (Rubin and Rubin, 1995). Key stakeholders, those most actively involved, were interviewed on various occasions throughout the project, so providing evidence of changes that occurred over time in response to the unfolding evaluation process. Table 1 gives details of the data sources in each case study.

Interviews were semi-structured, and the topics covered by each interview reflected developments in both the evaluation process and the research. In each case the aim was to understand:

- the respondent's understanding and assessment of the evaluation process;
- how this related to previous experience and current expectation; and
- future expectations relating to the evaluation and its outcomes.

The data collected from multiple sources were analysed within each case using the NUD.IST qualitative data analysis software before emerging themes were tested against both the literature and the data from the other case. The approach used to analyse the data is captured in Boyatzis' (1998) thematic analysis wherein the raw data are used to generate themes or patterns that, at a minimum, describe

Table 1. Data Sources

<i>Cast study One – HighEd</i>	<i>Case study Two – GovtAg</i>
<ul style="list-style-type: none"> • Project Director (4 interviews) • Director of Equal Opportunities Unit • 2 × Project Team Members (1 interviewed on 3 occasions, the other once) • 2 × Trainers • Head of Training • 3 × consultants (2 interviews each) • Manager and recruiter • 2 × recruitiees <p>Total: 21 interviews plus observation of three meetings of project team with consultants and two focus groups run as part of the evaluation.</p> <p>Secondary data sources: committee papers, policy documents, letters and internal communications.</p>	<ul style="list-style-type: none"> • Senior Development Consultant, champion of the project (3 interviews in addition to ongoing discussions as part of the evaluation project) • The first CE, original champion of empowerment • Chair of the Management Development Group, Area Director • 3 × Area Directors • Director of Personnell (retired 1995) • Head of Personnel Branch <p>Total: 10 interviews plus, during the evaluation project</p> <p>Interviews with:</p> <ul style="list-style-type: none"> • 2 senior managers • 9 middle managers • 6 Focus groups with junior management each group from a different location • 6 Focus groups with clerical staff grades each group from a different location <p>Secondary data sources: annual reports, policy statements, internal reports and memoranda.</p>

and organize observations and, at a maximum, interpret aspects of the phenomenon. Thematic analysis begins with 'sensing themes' (Boyatzis, 1998: 11) within the data, progressing through the development of codes to the interpretation of the information and themes in order to contribute to the development of knowledge. All the primary data were transcribed and entered into NUD.IST in their entirety for coding. However, much of the secondary documentation did not exist in word-processed form and the volume was too great to permit complete transcription. Secondary documentation was therefore coded manually and cross-referenced as appropriate within NUD.IST. Paragraphs rather than single words or sentences were used as the basic unit of analysis as this ensured that the context of words was preserved and was clear when the information attached to categories was revisited at later stages. In addition to comparison within and between cases, the themes and patterns that were being identified were also compared with the literature. This allowed categories to be refined and possible relationships to be explored based on what had, or had not, already been found. This provided a means of reflecting on the emerging findings to encourage deeper insight, which, in turn, served to reinforce the credibility of the findings (Eisenhardt, 1989).

The Case Studies

HighEd HighEd is a public sector service organization in the higher education sector. The main site, where this research took place, employed approximately 3000 staff and had responsibility for the development of policy, regulations and the creation of materials and services to be offered to consumers. The structure was a complex combination of matrix and hierarchy with a strong dependence on committees for the formation and monitoring of policy. The initiative being evaluated formed a strategic part of the organization's commitment to equality of opportunity and was intended to ensure that recruitment and selection practices were fair. 'Open Recruitment' had a direct impact on recruiters who were required to undertake training in 'fair' selection practices, but also, due to the nature of the initiative, it indirectly affected the whole organization. Implementation of this initiative had taken place over a five-year period in which no evaluation had been undertaken. At the end of this period, the absence of evaluation activity within the organization had become an issue due to the decision to achieve the Investors in People (IiP) accreditation. The IiP standard required that the organization demonstrate that it evaluated its development of people within the context of business goals and targets at the individual, unit and institutional level. 'Open Recruitment' was chosen as the vehicle for developing and demonstrating an evaluation model that would fulfil the IiP requirements. In the belief that in-house skills were insufficient, external consultants were used to assist and support the in-house team responsible for undertaking the evaluation. The evaluation project took place over a six-month period during which time data for this project were collected.

GovtAg This was one of a number of ex-Civil Service Departments in the UK to be given Agency status under the Next Steps programme in 1991. At the time of the research, GovtAg had around 65,000 staff based in a national network of offices that provided a variety of informational and support services to their client groups. Having gained Agency status, the organization introduced initiatives intended to deliver the desired behavioural and attitudinal changes perceived as necessary to create a supportive, enabling culture that would improve service to external and internal customers. Eight key areas were identified, one of which was the empowerment of local management through maximum devolution of responsibility. The project that forms the basis for this case study was a means of assessing the value of empowerment to this organization five years after its introduction when the political emphasis had changed and the Chief Executive (CE) – who had personally championed empowerment – was no longer part of the organization. At this point, the future of empowerment in GovtAg was unclear and uncertain. The Department to which GovtAg was responsible was aiming to reduce costs; it was estimated that an annual 20 percent downsizing would be needed to keep within the public expenditure ceilings, and there was increasing ministerial interest in private sector involvement as a means of achieving this. The political context against which this evaluation project took place encouraged the use of external consultants because an external assessment served the purpose of

public accountability by being seen to judge the initiative impartially. Provided the outcomes were as the sponsors hoped, the evaluation could also be promoted as unbiased 'expert' support for what had been achieved – and continued to be by the new regime – in the name of empowerment. The evaluation took place over a six-week period.

The Prompt for Evaluation

The distinction between primary and secondary barriers was drawn following the discovery that, in each case, an external prompt had occurred that had led to these particular evaluations being undertaken. The research project began with the intention of exploring the barriers to evaluation based on the expectation that such barriers occurred during the evaluation process. However, in each of these cases, a change in circumstances following the implementation of the initiative had been the prompt for the evaluation to be undertaken. This both indicated the existence of barriers prior to the evaluation and also offered an opportunity to explore the nature of both types of barrier. In HighEd, the prompt had been the requirements of IiP, while, in GovtAg, it was the change in political climate. There were clearly barriers in place prior to this prompt (primary barriers) that had prevented evaluation taking place until a change in circumstances had occurred that created a perceived need for evaluation that was stronger than the barriers. At that point, factors relating to the process of evaluation (secondary barriers) became an issue, those factors that make it difficult for 'good' (thorough, unbiased, relevant) evaluations to take place and that determine the extent to which useful findings are produced.

Primary Barriers

Contextual Factors

Within the literature, the contextual factors relating to evaluation tend to be ignored (Henderson and McAdam, 1998); within the case studies, however, they proved to be the source of real and significant barriers to an evaluation being undertaken. Historically, there was little evidence of effective evaluation within the organizations, particularly in the context of 'soft' change initiatives. In GovtAg, Area Director A stated: 'we don't tend to evaluate "touchy-feely"', and Area Director B explained: 'there are no established measuring devices for the less tangible criteria'. Furthermore, a Project Director of HighEd stated:

I don't think that managers are encouraged to think in evaluative terms and I think little of the literature suggests any kind of systematic approach. . . . I mean, I've worked in a number of soft change areas and I think it's particularly hard to pin down evaluation techniques that are useful for those areas. (Project Director, HighEd)

No one felt that evaluation was something that was done well in their organization. In addition, in each case, there was evidence that individuals perceived themselves as operating in an organization in which the use of previous evaluation findings had been negative and divisive.

Skinner: Primary and Secondary Barriers to the Evaluation of Change

How can we evaluate what we've been doing while still retaining a sense of being a team, because some of the evaluation is bound to go back in and say, 'well, I think you were wrong there, or that was wrong there, or we should have done it better, but also I'm suggesting that you could have done it better'. (Project Director, HighEd)

In organizations where individuals believe that they 'work in a blame culture, increasingly so' (District Manager A, GovtAg) then 'evaluation opens up a can of worms for everybody' (Project Team Member, HighEd). Clark et al. (1998) observe that much of the popular management theory, and the change initiatives it inspires, supports an ideology of management with managers cast as heroes or heroines who make success possible (Clark and Salaman, 1998). Managers become closely associated with the success of particular projects because managers are 'not passive bystanders when it comes to the importation of new ideas' (Tyson, 1999: 45), selecting, reinterpreting and giving relative emphasis to ideas according to their own agendas. Inevitably against this background an evaluation that might suggest anything less than total success is likely to be in conflict with personal self-interest, not least because, as Fox (1989: 192) notes, 'evaluative judgements must be made openly about people'.

The 'Grand Strategy' Approach

With respect to the initiatives under consideration here, there were no incentives to challenge the perceptions that had formed as a result of previous experience. Tichy (1983) argues that a senior management focus on 'grand strategy' means that evaluation is unlikely to feature in implementation plans. Previous research found that senior managers are more likely to be heavily involved in the conception and initiation stages of an initiative than in its implementation (Skinner and Mabey, 1997). Patrickson et al. (1995) note that there may also be a reluctance to allocate time and resources to a process that focuses on the past rather than the future: 'I do think people are very resistant to go digging around – you know, what's past is past; we're here now, let's just go forward' (Project Director, HighEd). In each of these cases, senior management had focused on the creation of the initiative rather than the specifics of its implementation. The consequences included the failure to define success criteria or to designate responsibility for monitoring progress: 'we didn't have a framework of accountability' (former CE, GovtAg). The result was that neither the information, nor the resources, necessary for an evaluation were readily available.

In keeping with the focus on 'grand strategy' there had been little to indicate to those involved in these initiatives that evaluation was an activity required or valued by senior management. In HighEd it was suggested that there was little senior management interest in evaluation in general either because 'they are always rushing on to something else' (Director of Equal Opportunities, HighEd) or because 'they don't need to evaluate, they think that they are the senior management team; therefore by their very nature what they do is good' (Recruit A, HighEd). This perception continued even after these particular evaluations had taken place.

Unless I make an effort before my time in this place is done, I don't think any of my

Evaluation 10(2)

line managers will be saying, 'Well, I'd like your evaluation report on how this has gone', so it's, you know, if you like, if you aren't being pressed by something in the hierarchy. (Project Director, HighEd)

In GovtAg the atypical nature of the empowerment initiative, 'the general vague introduction' (former CE, GovtAg), resulted in an incremental and emergent implementation rather than the systematic project plan, including evaluation, which would have been normal for a large project in this organization.

Nobody actually voted any money at all for it, what happened was that [the former CE] and his influence started to work and people started to set up training, development events, team building with empowerment as the theme. (Senior Development Consultant, GovtAg)

The result of this was:

... if you haven't got a Board pronouncement of it and there hasn't been something gone out in writing and it hasn't found its way onto their agendas then there is no stimulus to actually monitor progress in a structured way or evaluate it in a structured way. (Senior Development Consultant, GovtAg)

Managers at all levels of organizations struggle constantly with both the pressure to succeed and the pressure of time (Swanson, 1997). As a consequence, they are unlikely to undertake activities that they consider are not seen as priorities in the minds of their superiors and for which they have not been given specific responsibility. The perception in both cases was that for five years there had been 'no stimulus to evaluate in a structured way' (District Manager B, GovtAg) because 'if your line manager isn't interested in it, it's not going to be in your objectives, and if there's nothing formal about it...' (Project Team Member, HighEd) it is unlikely to happen.

The Assumed Value of the Initiative

In these organizations, the initiative in question directly related to the culture, values and public image perceived as desirable and which senior management wished to promote. 'Equal Opportunity awareness is very much part of the organization's culture' (Manager, HighEd), and 'in a large organization such as GovtAg, empowered people are essential if we are to respond successfully ... it's perfect for that organization' (former CE, GovtAg). In each case, senior management perceived the initiative as a means of addressing issues, both potential and actual, which related to the desired state. As suggested by Brunsson and Olsen (1998), these reforms were perceived to have an inherent value of their own and assumptions were being made, albeit to some extent subconsciously and in very general terms, that there would be benefits that would inevitably result from their introduction. There was no evidence of a detailed or considered assessment of either the organizational need or the appropriateness of these particular strategies in these contexts before they were introduced. This unquestioned belief that there would be an inevitable benefit to the organization from introducing such initiatives also serves to reduce the perceived need to evaluate formally – those responsible for its initiation already 'know' that its effect will

be positive: 'empowerment is the way forward for most organizations' (Senior Development Consultant, GovtAg); 'on "Open Recruitment" it is a brave person who challenges a senior member of staff' (Project Director, HighEd).

Informal Evaluation

Easterby-Smith (1994) observes that the evaluation process is a complex one that cannot be divorced from issues of power, politics, value judgements and human interests, and so it proved in the case studies. The research began with the intention of focusing on the formal evaluation process that took place. It quickly became clear, however, that informal evaluation was a continual, ongoing activity that occurred at all levels, and was a key factor in the creation of barriers. At the most senior levels, the adoption of the grand strategy approach and the unquestioned assumptions about the inherent goodness of the initiatives already noted demonstrate the power of informal, often personal, evaluation, but this was not restricted to those in the highest levels of management.

In GovtAg, the sponsor had already assessed the success of empowerment on a personal level and, if it had been left to him, would not have evaluated at all because the initiative's value was 'manifestly observable' (Senior Development Consultant, GovtAg), a statement that assumes that everyone shares the same perspective. In the context of busy and demanding organizational life, the role of most managers frequently requires that they have to make decisions/take action on the basis of their personal interpretation of limited information. On this basis, a predisposition to accept their own personal assessment as an accurate evaluation of outcomes and impact is understandable, particularly with the absence of any external incentives/requirements to evaluate further. Clark et al.'s (1998) ideology of management seems likely to reinforce the belief of managers in the value of their own judgements and dispose them towards pre-reflective reasoning (King and Kitchener, 1994). As a result, they are unlikely to question their own interpretations or to recognize the limits of their own understanding. Aligned with this, Easterby-Smith (1994) notes that managers, particularly at senior levels, have a preference for information received via their own informal information channels, and that this information tends to be far more influential than that produced via more formal channels.

The Impact of Primary Barriers

Clearly then, within each of the case studies, there were significant deterrents to the undertaking of a formal evaluation as perceived by those with the necessary power and influence to publicly assess what had been achieved. In their reasoned action model, Ajzen and Fishbein (1980) suggest that it is reasonable to feel social pressure not to perform a behaviour if those with whom you wish to comply think you should not perform it. In both organizations, the past experience of many was of evaluation that was not well executed and that resulted in personal criticism and the attribution of blame. Each of the managers responsible for the evaluations could articulate the 'accepted' reasons why evaluation was necessary and how it could benefit the organization. However, on a personal level, they were clearly aware of the negative connotations of formal evaluation

for managers and staff in these organizations. In that sense, there were risks involved in choosing to evaluate, not only in terms of receiving personal criticism but also of the activity being unpopular with peers and others who might feel vulnerable. On this basis, it seems unlikely that the attitude of individuals towards the behaviour, the undertaking of evaluation, would be very positive in the context of their own organization.

Ajzen and Fishbein (1980) also argue that an individual's behaviour follows logically and systematically from whatever information is available. On this basis, neither the person's attitude towards the behaviour nor the norm in relation to evaluation were likely to result in it being perceived as a priority activity. Combined with a belief that 'we can't measure people's feelings and views' (Senior Development Consultant, GovtAg), and that evaluation is 'perverse, interesting and difficult' (Project Director, HighEd), there would seem little incentive to pursue it unless the circumstances changed and raised its degree of importance at an individual or organizational level. A prompt for evaluation was required to create a perceived need for formal evaluation. Without these prompts, it is questionable whether a formal evaluation of the two initiatives reported here would ever have taken place. In each case, the prompt was sufficient to change the balance in the reasoned action equation in favour of evaluation being undertaken and, in each case, once the decision to evaluate had been taken, other issues and obstacles (secondary barriers) became relevant.

Secondary Barriers

Once the need for a formal evaluation had been agreed, it became necessary for decisions to be made in relation to the process itself and, at this point, the secondary barriers emerged. A variety of issues relating to choices of approach, evaluator, participation and use of findings had to be addressed at each stage of the evaluation. The choices to be made would affect the extent and nature of the evaluation and the ultimate effectiveness of the process, but the decisions made were in themselves constrained by the existence of barriers and, in turn, they had the potential to create further barriers. These are discussed below and examples of the supporting evidence are given for each section.

The Purpose of Evaluation

The initiatives and the evaluations did not have clearly considered and stated objectives that had been shared with those participating. The absence of success criteria for each initiative removed the need to assign responsibility or resource for assessment, with the result that, in HighEd and GovtAg, the data necessary to undertake an evaluation were not collected, a situation that was irretrievable and made the subsequent evaluation difficult. However, necessity required both organizations to articulate objectives for the evaluations in order to brief the consultants. These objectives had their limitations however, and were not widely shared.

In HighEd, the situation was complicated by the needs of IiP being superimposed on the evaluation of the initiative, and this increased the level of difficulty involved in defining a clear set of objectives. From the beginning, this translated

into a lack of focus and vagueness of purpose that caused uncertainty and tension, not least because the scope and extent of the project kept changing. In HighEd even those directly involved commented that management 'weren't clear what they wanted to do' (Consultant A), 'they don't seriously seem to know whether they want to do this or not. They don't know whether they want to do it for real or just to get into the portfolio' (Consultant B).

The original objectives identified by GovtAg did not reflect the reality of what was either possible or necessary within the constraints and needs of the project and had to be revised. A lack of clarity still remained, however, which was demonstrated by the additional questions identified by the sponsor once the evaluation had been completed and the 'gaps' in the agreed brief were exposed. Similar to HighEd, those participating in the evaluation in GovtAg were unaware of any detailed objectives for the evaluation, even at senior levels, but many had a broad idea and had made assumptions based on their expectations and experience.

I've no idea why they've decided to look at it now, it's perhaps that senior management have noticed it's changing, or that they are reviewing the impact, . . . or the worst case scenario perhaps there have been crappy decisions lately so we'd best see what empowerment is doing. (Area Director A)

In each case, the assessment was intended to serve a purpose in addition to understanding the value and outcomes of the initiative itself; it was about demonstrating or proving something that was important to those involved. Contrary to warnings given within the literature (e.g. Easterby-Smith, 1994) concerning the dangers of hidden agendas that distort the purpose of the evaluation, the hidden agendas held by those responsible for the evaluation served as a driving force rather than as a barrier. In both HighEd and GovtAg, the evaluation was summative, clearly reactive (Stufflebeam, 1972) and, in terms of the initiative itself, fits Easterby-Smith's (1994) definition of proving (the measurement of worth and impact). However, this was merely contributing to an overarching purpose, more akin to Nevo's (1986) socio-political function, where the primary purpose of the evaluation was to be used for a communicative purpose intended to create awareness or motivate desired behaviour. This was quite explicit in HighEd, as pursuit of IiP status and the need to meet the requirements of the standard were visible activities within the organization, and it was quite clear that the evaluation of 'Open Recruitment' took place in that context. In GovtAg, the objective of demonstrating what the initiative had achieved in order to influence senior management attitudes, thereby protecting the initiative, was not a formally stated objective; nevertheless, the desirability of pursuing this purpose was quite widely perceived and shared.

Managerial Focus

The managerialist bias in the practitioner literature (Clark et al., 1998) perpetuates the notion of the management perspective as pivotal in organizations. The underlying assumption is that all the parties involved subscribe to this view and accept the processes of authority and their consequences. This bias was also evident in the evaluations. In each, the primary motivation was political (Patton, 1997), the intention of each sponsor being to use the evaluation findings to

influence management thinking, decision making and behaviour at higher levels of the organization. This created a barrier in the sense that the evaluation process became focused towards that purpose without any recognition in HighEd or GovtAg of the impact that this might have in terms of the evaluation and the wider organization. In HighEd it had simply 'not been thought through how we will communicate the results of the evaluation project, to whom or how' (Project Team Member, HighEd).

Similarly, Huzzard and Ostergren (2002) note unitarist assumptions in the organizational learning literature with its emphasis on shared vision, values and the tendency to suggest that learning requires a consensus of interpretation. Pettigrew et al. (1992) suggest that it may be important to influence the conditions that determine how situations are interpreted in order to enable the organization and its members to be able to deal with new situations and create lasting change. This highlights the political nature of knowledge and the importance of power and influence in the utilization of information, cynically reflected by the following opinions:

... they will sit on it for a long time, I think it will be pushed around on bits of paper for a very, very long time, and eventually we may end up with something and that something will be very watered down. (Middle Manager A, GovtAg)

... the senior management team will send it out to senior managers who will decide whether it goes down, it will depend what's in it. If they like it they will put it in a glossy brochure, very glossy, if they don't like it they will put it in their filing cabinet in the sky won't they? (Junior Manager's Focus Group E, GovtAg)

In each of these cases, sponsors did not accord equal status to each group of stakeholders, or indeed recognize all the potential stakeholders, through their inclusion. The need to share findings on a wider scale was an afterthought; there was little recognition in either organization of the potential, or need, for shared learning or interpretation, particularly outside management circles: 'I'm not sure it's in the least appropriate to feed this back' (Project Director, HighEd). Yet, Bruce (1998) argues that evaluation, through helping to ensure that learning takes place, provides closure for a project and serves to motivate people to be willing to participate actively in the future. The involvement of others as respondents in the evaluation had also created a desire for reciprocity that had the potential to lead to dissatisfaction and to create a barrier for the future if unfulfilled.

It's important that I get some feedback from being involved in this process because my time is in demand. I want to see the evaluation report; I want to see if my experience is different from other people. (Recruit B, HighEd)

... communication is a two-way process; if there is to be communication I need to get feedback; if I don't get feedback, it is of no use to anyone. (Clerical Staff's Focus Group A, GovtAg)

Without feedback, the process had no perceived added value for the individual and, as Fulop and Rifkin (1997) suggest, a lack of reciprocity or mutual self-disclosure can result in fear and a sense of threat that can, in turn, lead to concealment.

Once the wider interest in the findings had been acknowledged the instinctive, almost paternalistic, reaction of the managers was that the findings would need to be interpreted in some way before being shared. This would help people understand and put things into perspective, not least because people may not have 'the skills or knowledge to think about things in a reasonable way' (Project Director, HighEd). There was evidence that organizational defensive routines, as defined by Argyris (1986), were at work wherein the manager's automatic reaction was to ensure that, within the wider organization, the presentation of the analysis did not embarrass, expose or threaten management as a collective group. It was certainly evident from the interviews that staff from the lower grades in both HighEd and GovtAg expected there to be some selectivity and interpretation if, and when, any findings were shared with them, particularly if the findings were critical of activities: 'They never hear anything negative. Nobody wants to say we got it wrong' (Recruitee A, HighEd).

Lahteenmaki et al. (2001) note that, despite the close relationship between the change management and learning organization literatures, those writing about change management tend to ignore the mechanisms through which learning translates into organizational performance. Evaluation is no different: the change management literature is strangely quiet about the practicalities of its inclusion in the change process. Nor does its absence appear to be an issue that many of those writing and researching in this aspect of the management literature have recognized as problematic. In 1975, Davis and Salasin observed that the lack of dialogue between those specializing in evaluation and those in organizational change was puzzling, a situation that had not changed in 1993 when Wiggins observed that there was still a paucity of thinking about how they might fit together or when, in 1998, referring to management and business, Preskill wrote that people in this field have rarely looked to the field of evaluation for insights and knowledge about evaluation.

Conflicting Research Perspectives

Organizational bias and some uncertainty about assessing 'soft' initiatives led to a conflict of research perspectives that differed from a crisis of verification (Legge, 1984). Rather than difficulties caused by the incompatibility of academic requirements and the need for a pragmatic approach to research design, there was evidence of an incompatibility between the methodological bias of the organization and the pragmatic needs of the evaluation research. Each of the evaluation processes considered here used qualitative approaches to data collection yet, in common with the organizations described in previously reported research (Skinner et al., 2000), there was a definite bias towards quantitative measures and 'hard facts'.

In GovtAg, this reflected not only the organization's past history but also the prevailing political climate at the time of the evaluation: 'the Agency criteria are visible numbers and targets met within budgets' (Area Director B, GovtAg); 'targets are more prominent now, it's league tables . . . staff also say we only seem to talk about targets' (District Manager B). The initial intent in GovtAg had been to gather data for the evaluation through quantitative means, a choice largely

made on the basis that the organization was perceived to give more weight to quantitative results, rather than any sense that quantitative measures were the most appropriate. However, the absence of objectives for the initiative and lack of any formative data made this difficult within the constraints of timescales and budgets. The adoption of a qualitative approach was, however, largely a result of the sponsor being receptive to the consultants' views, based on their knowledge and expertise.

The situation in HighEd was rather different as it was 'not a very systematic organization' (Project Director, HighEd) that was 'not target- and measurement-oriented' (Manager, HighEd). Indeed, there was evidence that people were target- and measurement-averse! Due to the nature of the organization, there was also a greater awareness of the range of research methods that could be used in this type of project. Yet even here, there was evidence of a quantitative orientation when any measurement had previously been undertaken, for example 'the primary criteria for development or training being deemed a success is throughput, bums on seats' (Project Team Member, HighEd) and any performance indicators that were specified tended to express increases or improvements in purely numerical terms. However, this should not be overstated. While the IiP standard required a certain amount of quantitative data in terms of throughput and cost, the IiP team's inclination from the beginning was to incorporate qualitative methods, such as interviews and focus groups, to gather information. The desirability of this approach was largely based on pragmatic concerns, as the absence of historical data meant any assessment of improvement or change over time could only be made on the basis of people's experiences and perceptions.

The Choice of Evaluator

On the surface, pragmatic considerations such as lack of resources or expertise were responsible for the choice of evaluator. In HighEd, the perception of the necessary sequence of events set against the time and resources available led to the conclusion that an input of additional resource was necessary. The belief was that the requisite expertise was not available in-house, although in reality, there was evidence to suggest that the expertise was there, but that there was little effort or inclination to explore this possibility:

... at the time, people were feeling run ragged, aware of time pressures. We believed that we didn't have the skills to do it ourselves, but I'm sure we did. In hindsight, perhaps we convinced ourselves that we didn't have the skills. (Project Team Member, HighEd)

In GovtAg, although issues relating to the availability of internal resources, expertise and the timing were relevant, the primary driver was the perceived need for a detached, unbiased, external assessment of the initiative.

In order for people, in particular, the kind of people that are in our organization, to buy into that kind of belief, they want what they would describe as some sort of objective evidence. (Senior Development Consultant, GovtAg)

or 'an evaluation done in-house would be given a positive gloss' (Area Director C, GovtAg).

Skinner: Primary and Secondary Barriers to the Evaluation of Change

Within the cases there was no evidence of the evaluator conferring an inappropriate credibility on the evaluation in the form described by Legge (1984). Management within HighEd and GovtAg felt that there was some benefit in an external, detached view, particularly if it was linked to an understanding of what was happening in other organizations, and the consultants chosen were felt, by management at least, to bring significant credibility with them. However, this was not a perception shared by all. Staff in both organizations were sceptical about the extent to which those who were unconnected with the organization could possibly understand the reality and add any value. In HighEd there were concerns about 'artsy-fartsy consultants doing things that people couldn't get into' (Project Director, HighEd), a feeling that consultants 'only heard what they wanted to hear' (Recruit B, HighEd) and a belief that 'it's better to work through our own solutions' (Project Team Member, HighEd). Reservations were also expressed within GovtAg:

I can't see the point of somebody coming in and speaking to us and seeing what they have found and then passing it back. Why can't the management themselves do that, why can't they bring someone down to do it? (Clerical Staff's Focus Group D, GovtAg)

I've never been keen on outside consultants with empowerment . . . I think we're more than capable within a large organization to have those skills . . . and I think it helps when you have people come to talk to you who have some knowledge about the area of work you're working with. (Junior Manager's Focus Group C, GovtAg)

Individuals could do it on their own, just by keeping their eyes open and talking to people and then they could make their own mind up. It doesn't need to be a piece of central research. If people were of a mind to make their own mind up, or felt they were able to make their own mind up, and then act on it without being backed up by either sanction from above, central research, etc. (Clerical Staff's Focus Group E, GovtAg)

In the case of HighEd, the Project Team found that there were problems caused by the consultants being external to the organizational culture, not least the need to 'translate' documents for presentation to make them acceptable to committees and senior managers: 'one of the problems about using an external consultant to write a report for a committee is, of course, that they don't understand the politics of committees' (Project Director, HighEd). Experience gained during the project led the Project Director to the conclusion that a member of the Project Team needed to accompany the consultants to enhance their credibility. In GovtAg, the use of external consultants detracted from the credibility of the project for some non-management staff who perceived their use as a means by which management could abdicate their responsibilities: 'They haven't got to act on it, they just ask you to write a report . . . just a fact-finding thing and then that's the end of it' (Clerical Staff's Focus Group A, GovtAg). Thus, contrary to Legge's (1984) prediction of evaluators lending unwarranted credibility to an evaluation, in these two cases, the presence of the chosen evaluator had the potential to undermine the credibility of the project.

The issue of evaluator independence also has a bearing on the credibility of outcomes. Within the literature, the desirability of evaluator independence is debated in terms of fairness, impartiality and accuracy and the degree to which the

bias of dominant stakeholder groups may affect the findings (e.g. see House, 1977; Simons, 1984). In HighEd, the choice of evaluator and the nature of the relationship between the evaluator, the project and the organization, created some perception of evaluator bias. The issues of bias related to the evidence that the consultants hoped for further work (something that was recognized by the IiP team and led to some cynicism) and thus were keen to provide whatever the client required.

They [consultants] would like a large-scale evaluation project going at HighEd to pay their mortgages and further their academic careers. A large-scale activity was what they wanted. (Project Director, HighEd)

Conclusion

In conclusion, were these evaluations deemed a success by those involved? In the case of HighEd, against the formal outcomes of the project as stated in the original brief, the evaluation project might be judged a success: an evaluation model, information for the IiP portfolio, the consultants' report on 'Open Recruitment' were all produced, and recommendations for improvement and development were agreed by the relevant committee. However, the reality was that little had changed with respect to evaluation and the organization had failed to gain IiP accreditation. Ultimately, the model was not successfully presented within the organization and it was not adopted for use due to its perceived complexity because, as the Head of Training and Development explained, 'people haven't really got the model in their heads; it's not workable; it's unwieldy, cumbersome and unworkable'. If, as Patton (1997) suggests, the true measurement of a good evaluation lies in its utilization, then this evaluation failed on all counts; little was learnt about the impact of the 'Open Recruitment' initiative, the model was not adopted as a framework for evaluation within the organization, and attitudes towards evaluation did not appear to have changed.

The evaluation of the GovtAg project reported that significant headway had been made in changing staff behaviour and attitudes, and that empowerment had been instrumental in what had been achieved. In the political and economic climate in which this organization was operating, the conclusion reported was that empowering staff was probably the only realistic option and that further work was necessary to reinforce an empowering management style. The importance of trust existing between the organization and its staff, and the difficulties of maintaining that trust, were also highlighted. The evaluation was greeted by management within the organization as a well-researched piece of work demonstrating that the investment had produced results. However, following the presentation, the main use of the findings was to provide supporting evidence in bids made by the evaluation sponsor for further development initiatives. In a sense, this could be interpreted as intentional misuse of the evaluation findings (Alkin, 1990; Patton, 1997) as selective use was being made of certain aspects to support bids for other initiatives, and thus there is a danger of distortion. However, insofar as the original aim had been for the evaluation to demonstrate the positive impact of such an initiative, use in this way is compatible with the original intent and could be argued to be contributing to the improvement of subsequent initiatives.

Skinner: Primary and Secondary Barriers to the Evaluation of Change

If, as the literature claims, the effective management of change and achievement of well-developed organizational learning strategies are the key to future competitive success for organizations, evaluation, both informal and formal, has a crucial contribution to make. By differentiating between primary and secondary barriers and their causes, this research has highlighted the complexity of the barriers that exist within organizations and demonstrated that each type of barrier has its roots in different factors (summarized in Figure 1). In combination, they provide significant obstacles and disincentives to the undertaking of evaluation in the context of a change initiative.

The barriers that were identified also highlight the highly political nature of evaluation (Clarke and Dawson, 1999) within organizations. The reality is that the power and resources necessary to instigate formal evaluation processes, and to make use of the findings, rest with the dominant stakeholder group. Garvin (1993) and Jenlink (1994) highlight the importance of an organization's ability to learn from its own experience and history and from the best practices and the experiences of others. Ironically, in the case studies, these skills, as exercised by individuals, led to widespread perception among the dominant group that evaluation was a difficult, unwelcome and time-consuming undertaking. The political realities of organizational life neither required nor encouraged individual managers to pursue evaluation proactively in these contexts. There were uncertainties and concerns grounded in the previous history of staff and organizations where evaluation had been experienced as a negative process, a vehicle for criticism rather than learning, and something held in low esteem by those in the most senior positions. The absence of detailed guidance, advice and examples within the management literature simply served to reinforce these perceptions, which were then further reinforced by some of the secondary barriers encountered.

In each case a change in context occurred that shifted the political balance in favour of evaluation, but towards an evaluation undertaken to meet a specific political agenda. This determined the form of the evaluation and the intended use of the findings and, within this limited focus, each evaluation could be deemed to

Figure 1. Barriers to Evaluation

have served its purpose. However, the dominance of the management agenda led to a focus on process and outcomes that served that agenda and failed to recognize or acknowledge the needs of other stakeholder groups who were contributing to the information on which the evaluation was based. The cases demonstrated that the sharing of information was important from the perspective of other stakeholders but that previous experience and history led to an expectation that the dominant management group would filter, restrict and suppress findings. This resulted, albeit unintentionally on the part of the managers, in reducing future participation and levels of trust while missing opportunities to share learning and increase readiness for further change. This was perhaps the greatest failing and the one that should be of greatest concern to the public sector organizations if, as Kouzes et al. (1999) argue, the values of public sector organizations are rooted in an individualistic, localized culture that reinforces short-term events and group heroes rather than collective learning.

References

- Ajzen, I. and M. Fishbein (1980) *Understanding Attitudes and Predicting Behaviour*. Englewood Cliffs, NJ: Prentice Hall.
- Alkin, M. C. (1990) *Debates on Evaluation*. Newbury Park, CA: Sage.
- Argyris, C. (1986) 'Reinforcing Organizational Defensive Routines: An Unintended Human Resources Activity', *Human Resource Management* 25(4): 541–55.
- Boyatzis, R. (1998) *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks, CA: Sage.
- Bruce, A. (1998) 'Aiming for Change? Stay on Target', *Professional Manager* 7(5): 24–5.
- Brunsson, N. and J. P. Olsen (1998) 'Reform as Routine', in C. Mabey, G. Salaman and J. Storey (eds) *Strategic Human Resource Management*, pp. 297–309. London: Sage.
- Burgoyne, J. (1994) 'Stakeholder Analysis', in C. Cassell and C. Symon (eds) *Qualitative Methods in Organisational Research*. London: Sage.
- Calder, J. (1994) *Programme Evaluation and Quality*. London: Kogan Page.
- Carnall, C. (1995) *Managing Change in Organisations*. London: Prentice Hall.
- Clark, T. and G. Salaman (1998) 'Telling Tales: Management Guru's Narratives and the Construction of Managerial Identity', *Journal of Management Studies* 35(2): 137–61.
- Clark, T., C. Mabey and D. Skinner (1998) 'Experiencing HRM: The Importance of the Inside Story', in C. Mabey, D. Skinner and T. Clark (eds) *Experiencing Human Resource Management*, pp. 1–13. London: Sage.
- Clarke, A. and R. Dawson (1999) *Evaluation Research: An Introduction to Principles, Methods and Practices*. London: Sage.
- Davis, H. and S. Salasin (1975) 'The Utilization of Evaluation', in E. Streuning and M. Guttentag (eds) *Handbook of Evaluation Research*, pp. 621–68. Beverly Hills, CA: Sage.
- Dixon, N. (1994) *The Organizational Learning Cycle: How We Can Learn Collectively*. Maidenhead: McGraw-Hill.
- Doyle, M., T. Claydon and D. Buchanan (2000) 'Mixed Results, Lousy Process: The Management Experience of Organizational Change', *British Journal of Management* 11(S1): S59–S80.
- Drew, S. and P. Smith (1995) 'The Learning Organization: "Change Proofing" and Strategy', *The Learning Organization* 2(1): 4–14.
- Driscoll, A. and J. Morris (2001) 'Stepping Out: Rhetorical Devices and Culture Change Management in the UK Civil Service', *Public Administration* 79(4): 803–24.

Skinner: Primary and Secondary Barriers to the Evaluation of Change

- Easterby-Smith, M. (1994) *Evaluating Management Development, Training and Education*. Aldershot: Gower.
- Eisenhardt, K. (1989) 'Building Theories from Case Study Research', *Academy of Management Review* 14(4): 532–50.
- Forss, K., B. Cracknell and K. Samset (1994) 'Can Evaluation Help an Organization to Learn?', *Evaluation Review* 18(5): 574–91.
- Fox, S. (1989) 'The Politics of Evaluating Management Development', *Management Education and Development* 20(3): 191–207.
- Fulop, L. and W. Rifkin (1997) 'Representing Fear in Learning Organisations', *Management Learning* 28(1): 45–63.
- Garvin, D. (1993) 'Building a Learning Organization', *Harvard Business Review* 71(4): 78–91.
- Ghobadian, A. and J. Ashworth (1994) 'Performance Measurement in Local Government – Concepts in Practice', *International Journal of Operations and Product Management* 14(5): 35–51.
- Henderson, J. and R. McAdam (1998) 'A More Subjective Approach to Business Improvement and Organizational Change', *Total Quality Management* 9(4/5): S116–S120.
- Hendry, C. (1996) 'Understanding and Creating Whole Organisational Change Through Learning Theory', *Human Relations* 49(5): 621–41.
- Hollinshead, G. and M. Leat (1995) *Human Resource Management: An International and Comparative Perspective*. London: Pitman.
- House, E. (1977) 'The Logic of Evaluative Argument', *CSE Monograph Series in Evaluation* No 7. Los Angeles: University Of California, Los Angeles, Centre for Study of Evaluation.
- Huzzard, T. and K. Ostergren (2002) 'When Norms Collide: Learning Under Organizational Hypocrisy', *British Journal of Management* 13(S2): S47–S59.
- Hyndman, N. and R. Eden (2000) 'A Study of the Coordination of Mission, Objectives and Targets in UK Executive Agencies', *Management Accounting Research* 11(2): 175–91.
- Jenlink, P. (1994) 'Using Evaluation to Understand the Learning Architecture of an Organization', *Evaluation and Program Planning* 17(3): 315–25.
- King, P. M. and K. S. Kitchener (1994) *Developing Reflective Judgement*. San Francisco, CA: Jossey-Bass.
- Kirkpatrick, D. L. (1985) *How to Manage Change Effectively*. San Francisco, CA: Jossey-Bass.
- Kouzmin, A., E. Loffler, H. Klages and N. Korac-Kakabadse (1999) 'Benchmarking and Performance Measures in Public Sectors: Towards Learning for Agency Effectiveness', *International Journal of Public Sector Management* 12(2): 121–44.
- Lahteenmaki, S., J. Toivonen and M. Mattila (2001) 'Critical Aspects of Organizational Learning Research and Proposals for its Measurement', *British Journal of Management* 12(2): 113–29.
- Legge, K. (1984) *Evaluating Planned Organisational Change*. London: Academic Press.
- Love, A. (1991) *Internal Evaluation*. Newbury Park, CA: Sage.
- Mabey, C. and G. Salaman (1995) *Strategic Human Resource Management*. Oxford: Blackwell.
- Nevo, D. (1986) 'Conceptualisation of Educational Evaluation', in E. House (ed.) *New Directions in Educational Evaluation*, pp. 15–29. Falmer Press: London.
- Owen, A. (1993) 'How to Implement Strategy', in C. Mabey and B. Mayon-White (eds) *Managing Change*, pp. 143–50. London: Paul Chapman.
- Patrickson, M., V. Bamber and G. Bamber (1995) *Organisational Change Strategies*. Melbourne: Longman.
- Patton, M. Q. (1997) *Utilization-focused Evaluation: The New Century Text*. Thousand Oaks, CA: Sage.

- Pedler, M., J. Burgoyne and T. Boydell (1991) *The Learning Company*. Maidenhead: McGraw-Hill.
- Pettigrew, A. and R. Whipp (1991) *Managing Change for Competitive Success*. Oxford: Blackwell.
- Pettigrew, A., E. Ferlie and L. McKee (1992) *Shaping Strategic Change: Making Change in Large Organizations*. London: Sage.
- Preskill, H. (1998) contribution posted on the EVALTALK Internet discussion list, 06.05.98.
- Preskill, H. and R. Torres (1999) 'Evaluations of Practice', in M. Easterby-Smith, J. Burgoyne and L. Araujo (eds) *Organizational Learning and the Learning Organization*, pp. 92–114. London: Sage.
- Reichers, A. E., J. P. Wanous and J. T. Austin (1997) 'Understanding and Managing Cynicism about Organisational Change', *Academy of Management Executive* 11(1): 48–59.
- Rubin, H. and I. Rubin (1995) *Qualitative Interviewing: The Art of Hearing Data*. Thousand Oaks, CA: Sage.
- Salaman, G. and J. Butler (1994) 'Why Managers Won't Learn', in C. Mabey and P. Iles (eds) *Managing Learning*, pp. 34–42. London: Routledge.
- Senge, P. (1990) *The Fifth Discipline: The Age and Practice of the Learning Organization*. London: Century Business.
- Simons, H. (1984) 'Negotiating Conditions for Independent Evaluators', in C. Adelman (ed.) *The Politics and Ethics of Evaluation*, pp. 56–68. Beckenham: Croom Helm.
- Skinner, D. and C. Mabey (1997) 'Managers' Perceptions of Strategic HR Change', *Personnel Review* 26(6): 467–84.
- Skinner, D., C. Tagg and J. Holloway (2000) 'Managers and Qualitative Research: The Pros and Cons', *Management Learning* 31(2): 163–80.
- Stufflebeam, D. (1972) 'Should or Can Evaluation Be Goal Free?', *Evaluation Comment: The Journal Of Educational Evaluation* 3(4): 7–9.
- Swanson, R. A. (1997) 'HRD Research: Don't Go to Work Without It', in R. Swanson and E. Holton (eds) *Human Resource Development Research Handbook*, pp. 3–20. San Francisco, CA: Berrett-Koehler.
- Tichy, N. (1983) *Managing Strategic Change*. New York: Wiley.
- Torraco, R. (1997) 'Theory Building and Research Methods', in R. Swanson and E. Holton (eds) *Human Resource Development Research Handbook*, pp. 114–37. San Francisco, CA: Berrett-Koehler.
- Tyson, S. (1999) 'How HR Knowledge Contributes to Organisational Performance', *Human Resource Management Journal* 9(3): 42–52.
- Weiss, C. (1990) 'Evaluation for Decisions: Is Anybody Out There, Does Anybody Care?', in M. Alkin (ed.) *Debates On Evaluation*. Newbury Park, CA: Sage.
- Wingens, M. (1993) 'Towards a General Utilization Theory: A Systems Theory Reformulation of the Two-communities Metaphor', *Knowledge: Creation, Diffusion, Utilization* 12(1): 27–42.

DENISE SKINNER is a Reader in Strategic Human Resource Management, and was previously a manager in the public sector. Her research interests include the management of change, the end-user experience of HRM initiatives, trust in the employment relationship and diversity issues. Please address correspondence to: Oxford Brookes University Business School, Wheatley Campus, Oxford OX33 1HX, UK. [email: denise.skinner@brookes.ac.uk]