Moving IS evaluation forward: learning themes and research issues

B. Farbey, a, *, F. Land, b, D. Targett, c

a University College London, London, UK
b London School of Economics and Political Science, London, UK
c Imperial College, London, UK

Received 19 July 1996; revised paper accepted for publication by Professor R.D. Galliers May 1999 following three further resubmissions

Abstract

“Value for IT money” has been and is an issue. There is concern that poor evaluation procedures mean it is difficult to select projects for investment, to control development and to measure business return after implementation. This concern has been matched by increased activity in researching IT evaluation. The purpose of this paper is to consider what is now needed to make further progress. We do this by drawing together diverse findings from three major research projects, conducted by the authors between 1989 and 1998, in the context of the literature. We suggest five “learning themes”, concepts which we believe would better inform future research and practice. These are: more extensive use of the theory of evaluation as a frame for classifying present prescriptions and discovering new ones; closer attention to multiple relationships among stakeholders; the development of more complex models of decision making; more emphasis on “whole life” studies of project dynamics and, finally, further exploration of the role of evaluation in management learning. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Information systems; Evaluation procedures; Learning organisations; IS research methodology

1. Introduction

Throughout the industrialised world investment in Information Technology (IT) and associated Information Systems (IS) is substantial. Some indication of the size of IT budgets comes from the annual surveys carried out by CSC and others. In 1997, for example, CSC reckoned average budgets in the companies they surveyed to be of the

* Corresponding author.

E-mail addresses: b.farbey@cs.ucl.ac.uk (B. Farbey), flandlse@aol.com (F. Land), d.targett@ic.ac.uk (D. Targett)

0963-8687/99/$ - see front matter © 1999 Elsevier Science B.V. All rights reserved.

PII: S0963-8687(99)00021-9
order of US$44 million (CSC, 1997). In the UK Spikes Cavell (1996) estimate that 58 of the top 150 companies spend more than 2% of turnover on IT and 96, more than 1%. Whether these high levels of spending produce a commensurate return is another matter, currently being debated in terms of a ‘productivity paradox’. At firm, industry and economy level there is contradictory evidence on whether IT expenditure has resulted in business value (for example Brynolffson, 1993). Hitt and Brynjolfson (1996) conclude that the confusion occurs because, although IT has led to increased productivity and consumer value, it has left business profitability unchanged. The statistical uncertainty adds substance to a continuing business worry about how the value of IT expenditure can be measured. There has been relentless pressure on IT departments and business generally to cut costs and account for money spent, yet few organisations appear to be satisfied with their evaluation procedures.

“Value for IT money” has been an issue for a number of years (Price Waterhouse, 1990–96) and is linked to the “productivity paradox” debate in two ways. The first is that poor evaluation of IT expenditure suggests that the statistics may be of doubtful quality, adding further confusion to the debate; the second is that poor evaluation practices have resulted in incorrect selection and management of projects, resulting in poor returns. In either case inadequate evaluation gives cause for concern.

Not surprisingly this concern has been matched by increased activity in researching IT evaluation. For example, there is now an annual European Conference on IT Evaluation (for example, Berghout and Remenyi, 1997). The authors have played a part in this increased activity having conducted three major research projects in the years 1989–1998. In this paper we ask what is now needed to move the topic forward, both theoretically and practically. Our aim is to draw together these diverse pieces of research in the context of the IS literature and from there move to suggest five “learning themes”, that is areas and emphases for research which we believe would better inform both theory and practice.

Section 2 of the paper provides the background, defining evaluation and reviewing recent developments in the literature. We then describe our three projects in Section 3. Sections 4–9 form the nucleus of the paper. They identify and discuss the five themes. We conclude (Section 10) by noting that, despite the attention already given to it, evaluation is still a “wicked” problem, and likely to remain so for a while yet.

2. The background to IT evaluation

The term “IT evaluation” is often used imprecisely. Sometimes it is referred to as an event taking place at the commencement of a project in order to decide whether the project should go ahead. At other times this decision point is called “appraisal”, reserving “evaluation” for a post-implementation review of achieved benefits. However, in both theory and practice the concept is much broader. Our working definition of evaluation for the rest of the paper is as follows. IT evaluation is:

A process, or group of parallel processes, which take place at different points in time or continuously, for searching and for making explicit, quantitatively or qualitatively, all the impacts of an IT project and the programme and strategy of which it is a part.
A contemporary approach to IT evaluation began to emerge in the 1980s. Journal papers and books in the late 1980s and early 1990s started to investigate in depth the earlier ideas and problems, proposing new perspectives and new solutions. The earlier writing included for example Emery (1976) and Land (1976). In the UK, the theme of evaluation was taken up in the late 1980s by, among others, Earl (1989), Hochstrasser (1990), Silk (1990), Symons (1990), Ward (1990), Symons and Walsham (1991), Willcocks (1992), Farbey et al. (1993) and Walsham, (1993a,b). There was similar interest elsewhere (for example Bjorn-Andersen and Davis, 1988; Remenyi et al., 1991, and Weill, 1992). In particular the newer writing addressed the question of the so-called Strategic Information Systems (SIS) which were perceived as different in kind from the older “transaction” systems and thus requiring new methods of evaluation (Parker et al., 1988; Earl, 1989).

There were a great number of methods on offer. These ranged from the general, for example, Return on Investment (ROI), (Brealey and Myers, 1991) to the IT specific, for example, Return on Management (Strassmann, 1985). They included methods relying on quantitative assessment of costs, benefits and risks such as Information Economics (Parker et al., 1988). There were also “softer” methods for identifying and assessing benefits, for example Multi-Criteria, Multi-Objective methods (Chandler, 1982) as well as those based on modelling and experiment, for example systems dynamics models (Wolstenholme et al., 1993).

Meanwhile organisational research showed that in practice little attention was paid to the formal evaluation of IT investment. Instead many substantial investments were defended as “acts of faith” or “got to do” or simply “strategic”. Where formal evaluation did take place it tended to be an accounting based technique, undertaken very early in the project life cycle or at the point where the system was handled on from the developers to the users (Kumar, 1990; Willcocks, 1992). Post implementation audit was rare. Appraisal and evaluation were viewed negatively as hoops which had to be jumped, an organisational imposition which took up valuable working time. They were not perceived as positive activities leading to clearer understanding, better decisions, tighter control, happy users and greater benefits.

In examining this gap between theory and practice, academic writing highlighted specific concerns (Symons, 1990; Ward, 1990; Willcocks, 1992; Galliers, 1995; among others):

- the treatment of costs, risks and benefits was partial and inaccurate;
- the difficulty of dealing with intangible benefits, and their frequent omission from business cases was leading to neglect of important aspects of the investments;
- the emphasis on tangible costs, benefits and risks was leading to the bending of rules;
- there had been a failure to align IT strategies with organisations’ strategic directions;
- the focus of evaluation was narrow, concentrating on the system in itself, rather than the intervention as a whole of which the (new) system was just a part;
- evaluation methods were aimed at project level while other levels, for example the IT portfolio, were neglected.

Academic prescriptions followed the diagnosis, calling for:

- a clearer examination of the strategic value of systems: Earl (1989);
- a broad view of what was to be evaluated, including IT portfolios: Ward (1990);
a more holistic approach: Wolstenholme et al. (1993); Galliers (1995);
recognition of the role of evaluation in organisational learning: Symons (1990);
greater appreciation of the wider purposes served by evaluation, including its political
and persuasive effects: Farbey et al. (1993);
more attention to accounting systems in theory and practice: Willcocks (1992);
further exploration of softer methods for determining costs, benefits and risks: Symons
and Walsham (1991); Walsham (1993a); Keen (1995), among others;
greater emphasis on post-implementation audits: Farbey et al. (1993).

Taken together these prescriptions represented if not exactly a new consensus, then at
least a fresh approach: broader, more insistent on the social nature of evaluation, more
situated and contingent than before. The objectives of evaluation had widened and
mechanistic approaches based on limited accounting and other hard data were seen to
be inadequate.

3. Three projects

Our own conclusions are based on three major projects carried out over the years 1989–
1998. The types of organisation involved are summarised in Table 1.


The main focus of the research was to study how the projects had been evaluated. The
systems concerned were management-related and in the area of office automation. They
were located in a wide range of private and public sector organisations and had just been
implemented or were about to be implemented. Evidence about the projects, from original
conception to the latest position, was collected from documentation and from interviews.
with people involved in the development and running of the systems. In particular the research sought to identify and interview the “champions”, the people who had a major influence in getting the project accepted. As far as evaluation was concerned *ad hoc procedures* were used in the majority of systems investigated. Even when a formal prescribed method was used it was frequently customised to suit the exigencies of a particular situation. For a full description of the overall findings of the research, see Farbey et al. (1993).

3.2. Project Two: 1993–1997: longitudinal research on 12 projects in 9 organisations

Project One suffered from two major drawbacks: it had investigated only projects that had gone ahead and it was based on data collection methods which relied on retrospection. This indicated a need for longitudinal research which would have three important advantages. First, it would be able to investigate project dynamics and observe how decision processes evolved. Second, projects which were not completed according to plan could be researched. Third, a method for matching project to evaluation method was developed empirically in Project One. It worked effectively but we wished to validate it and develop a theoretical basis. These concerns led to the formation of Project Two. The methodological approach was to develop an in-depth, interpretative study, following the progress of planned investments over a sustained period. Case studies were the primary research instrument, based on action research, collaborative enquiry, interviews, documentation and literature. The projects ranged across the systems life cycle from strategy formation through development and implementation to benefits management, maintenance and decommissioning. Each represented a significant capital investment.

To collect data we attended company meetings, wrote reports, participated in workshops, maintained telephone contact and conducted interviews. The interviewees were key personnel including project sponsors, project managers, senior IS managers and users. Some of the data was provided by our collaborators directly, in the spirit of collaborative enquiry. A systematic framework was employed to keep the accounts coherent both within and between cases. In addition the research team kept a series of informal research notes to record ideas and comments which might otherwise have been lost in the long timescale of action research. Some results of this project have been published in Farbey et al. (1998, 1999).


Our third research project stemmed from interest in the early results of the second. Our remit was to investigate the appraisal and evaluation of IT investments in the public sector and consider the extent to which benefits had been achieved from them. We were then to recommend improvements to the appraisal and evaluation processes. The six cases were chosen because they were significant projects likely to yield insights beyond the cases studied. All were either fully implemented or close to implementation; they ranged from clear success to probable failure. Our research was conducted in the context of the whole process by which the systems are developed, managed and implemented: appraisal and evaluation were threads running through that process with substantial influence at all parts of it. For each project we studied relevant documents and interviewed key members of
staff, including both providers and users of the systems as well as senior managers as appropriate. Further research continued in three of the organisations, one of which has been the subject of a doctoral thesis (Kain, 1999).

4. Five learning themes

From this extensive but often quite diverse experience we have drawn together five themes which taken together we believe will provide a basis for a deeper understanding of evaluation in practice. First, Learning Theme One continues a theme we first put forward in 1993. We want to position the evaluation of IS investment squarely in the context of the discipline of Evaluation (Farbey et al., 1993, 1999). Evaluation provides us both with a framework for considering the varied evaluation activity we have observed over the years and yields a number of practical insights. Second, we propose four further learning themes as significant. These have been chosen not so much to reflect what organisations were doing in practice to evaluate their investments but to accentuate the role and position of evaluation in the rough and tumble of organisational life.

5. Learning theme 1: evaluation theory

The literature is full of a wide range of evaluation techniques but research has shown that organisations use few of them, accounting-based methods being the favourite. A review of a number of potentially applicable techniques, and discussion of their strengths and weaknesses, is given in Farbey et al. (1993). The techniques have all been developed to enable organisations to cope with common evaluation problems. Each has its own distinctive characteristics which dictate the situations in which it can be used. There is no one method which is universally applicable and the problem is to select an appropriate technique to use in a given situation.

Every IT project has characteristics which influence the choice of a suitable evaluation technique. At the same time every evaluation technique has characteristics which point to
the set of circumstances in which it could be applied. In Project 1 we developed, empirically, a process for matching a system with an evaluation technique (Farbey et al., 1992). The process has three stages:

1. Represent the characteristics of the project as crosses on a $2 \times 2$ matrix. Each cross relates to one of six factors relevant to evaluation. In practice the technique has been used with 10-point scales to assess the factors and position the crosses on the matrix.
2. Locate each evaluation technique at some point on another $2 \times 2$ matrix. The evaluation methods were assigned to cells according to their dominant features.
3. Overlay the matrices to match project with technique.

The two axes of which were encapsulated as (see Fig. 1):

- “well defined vs. fuzzy” evaluation constraints;
- “conservative vs. radical” role of IT.

The location of the crosses, whether clustered or dispersed, is used to suggest the range of techniques that might be applied. In some cases all crosses may fall within the same quadrant giving a strong indication of exactly which techniques might be suitable. In other cases the crosses might be spread around indicating that the choice is not clear and several techniques could be used.

These matrices and the matching process are described in detail in Farbey et al. (1993). The method has proved to be effective for the projects we were studying and also for other researchers working independently. However, there is a theoretical platform which explains why the method works. The classification, which is due to Hellstern (1986), comes from the “evaluation research” body of work. It is also available in the IS literature in an article by Earl and Hopwood (1987) where it is discussed in terms of strategic decision-making. Taking all three classifications together, it becomes clear that they are essentially the same. In particular, the logic which underpins the Hellstern categorisation and the language which he uses provide a rationale for our empirically discovered method.

Hellstern is writing about evaluating programmes of public policy—major governmental programmes. They are social interventions and he writes that:

- “technical” program evaluations… may prove to be successful for well defined problems with clear goals and clear causal relations… most problems… are ill-structured… with substantive problems varying from actor to actor and over time; they are dynamic. Such problems are not mechanical and cannot be separated from the context.

This gives a different perspective to the argument we have been making. We advocated consideration of different methods principally on the grounds that a variety of methods was necessary to match the variety of situations encountered. Following Hellstern, it is possible to argue for variety on the grounds that the problems are not only often ill-structured but also that the lack of structure has two bases: lack of structure with respect to objectives and lack of structure with respect to knowledge of cause and effect.

Expanding the argument: if IS are complex and pervasive socio-technical systems whose life extends over several months or years then an IS investment can be seen as a social action, based on a complex technology and taking place over a substantial period in time. They are thus like the programmes of social action which are the subjects of evaluation research. In particular they generally present problems to the evaluator
which are poorly structured. As with more general evaluation problems, the lack of structure has two bases: lack of clearly understood and agreed objectives and a lack of knowledge as to the potential impact of the IS, and hence a lack of knowledge of cause and effect.

Looking at the axes of our matrices, it can be seen that the distinction being made follows this same split. The vertical axis of the matrix represents situations where the objectives are not clear, whether the objectives of the system or the objectives of the evaluation. The horizontal axis reflects uncertainty about cause and effect. By definition, conservative action is action undertaken in known circumstances; radical action is not. Reverting to our matching process, and using Hellstern’s description of the contexts, we now assert that the critical features of each method are those which determine:

1. How far they contribute to, or are based upon, a knowledge of cause and effect.
2. How far they assume, or contribute to the formation of, clear and agreed objectives.

Evaluation theory has informed much recent work in evaluation. Examples are the work of Remenyi et al. (1997) in formative evaluation and Walsham (1993b) in interpretative evaluation. However there is not as yet as comprehensive a review and classification of evaluation methods approaches and techniques used for IS investment as that presented House (1980) for more general evaluation problems. House’s analysis, includes not only the methods and purposes, but their underlying philosophies and assumptions. This provides more detail than Smithson and Hirschheim (1998). We believe it would be both valuable and timely to build on Smithson and Hirschheim and progress to a structure similar to House, but with IS evaluation as the focal subject.

6. Learning theme 2: stakeholder theory

In both Projects Two and Three we were aware that external stakeholders can play a decisive and crucial role in many IS investments. Government policies change. In one case the news that the department of government in which we were hoping to conduct research had been abolished came (as a surprise to everyone) at the end of our introductory meeting with its principals—our shortest ever case investigation. Suppliers are sacked. Software upgrades do not work properly. Organisations are taken over. Customer organisations are made bankrupt, and so on. In more than one organisation suppliers or consultants were part of the Project Board. Being placed in the position of expert in the latest technology, perhaps because of a lack of self-confidence in the internal participants, the consultants were able to exert a dominating influence over decision-making. Indeed the literature cites many examples of IS projects in which multiple stakeholder groups are involved, with substantial influence.

This is the source of two surprises. First, established systems development and project management methodologies give little attention to the role of multiple external stake-holders. Second, there have been few attempts to analyse the phenomenon in detail. This is despite the fact that the concept of stakeholding, divorced from IS, is an area which is actively researched, principally in the context of environmental policy. There has, for example, been debate about the nature of stakeholders and whether they can be other than
individuals or groups of people, or even non-humans (Vidgen and McMaster, 1997). Can, for example, a nature reserve be a stakeholder (Kain, 1999)? There are clear parallels between these general debates and specific issues concerned with stakeholders connected to IS investments.

After the official close of Project Three, and with such thoughts in mind, we continued to research in three of the public sector organisations specifically to investigate external stakeholders. Preliminary findings from some doctoral research in one of the organisations are:

1. The participation of external stakeholders usually derives from regulation or the resources they command.
2. The relationship between stakeholders and the organisation was intrusive, affecting a number of discrete management activities as well as whole processes.
3. There were relationships between the external stakeholders, often hidden from the organisation, which exerted an internal influence on the organisation and the IS investment.

Turning to internal stakeholders we found that the roles played by each group: users, sponsors and developers varied. “Users” came in different flavours. The “user” on a project board was most likely to be a representative of the sponsoring department or a manager with experience of the area affected by the new system but no longer active within it, rather than a hands-on user. This could create problems if the user representatives were out of touch with current practice, or “went native” (Farbey et al., 1995). The role of the business sponsor was critical. A powerful sponsor could act either to rescue a project or to sink it. An investment in one organisation lost momentum when it lost senior management support. In another two investments were finally lost because they could not attract the financial support of a senior business sponsor.

However our most interesting findings in this area came from one organisation where we observed three projects. The stakeholders in each project overlapped, so that what we were witnessing was a series of similar “games”, using the term in the game-theoretic sense, with shifting participants, interests and coalitions. At times it felt as though one were part of a ‘segmented’ society (Gellner, 1991) in which small groups would be in opposition to one another in some contexts, say the project context, but would band together to fight a larger battle beyond it. At others it felt more Hobbesian “each against all”.

A significant consequence of these intense games was that many stakeholders were not aware of, or unable to fathom, the larger picture. Moreover, in both Projects Two and Three we found instances of two people standing in two different hierarchical relationships to each other. In Project Two, despite an initial good will and considerable intelligence on both sides, the position became unsustainable. One person was caught in a political pincer movement and withdrew. Internal competition and internal market organisation reinforced these internal tensions and multiple games.

In other words, relationships between internal stakeholders in a project were not necessarily confined to that project but extended beyond it. These relationships were ‘multi-stranded’ (Gellner, 1991). The project was one meeting ground among others, so that for example, if there were pre-existing tensions between IS and the business, or between two
user departments, they would carry through into the project. The project became a continuation of the war by other means.

Following this finding we suggest that field research not only notes and analyses the stakeholders, their interests and relationships within the context of the project, but also seeks to uncover the relationships outside it, since this may have a significant bearing on what happens inside the project.

7. Learning theme 3: evaluation as part of decision-making

A major objective of Project Two’s longitudinal research was to investigate the IT decision-making processes of which evaluation is a part. Early research on decision-making in general concentrated on conceptual simplifications which attempted to make some sense of it. For example, Simon (1955) developed the concept of “bounded rationality” and Lindblom (1959) saw decision-making as incremental, “mudding through”. The 1980s saw attempts to categorise different types of decision process. For example, Hickson et al. (1989) identified three types of process: sporadic, fluid and constricted. More recently research has changed direction, attempting to expose the full complexity of decision-making. For example, Langley et al. (1995) proposed a model of decision-making which incorporated a range of interacting factors which could affect the process. Our research, based on interpretivism, has led us in this same direction and a number of complex issues have emerged.

7.1. More organisations made use of formal evaluation methods but …

Project Two showed that formal appraisal and evaluation procedures were more common in the second tranche of organisations than they were in the first. All except one of the investments observed were formally appraised. The formal process was, however, not always decisive in determining the initiation or continuation of a project. While the platform was an accounting calculation supplemented by a statement of the expected contribution to strategic aims, other factors, political or strategic, entered into the decision and often dominated. Post hoc evaluation, was not well done although many organisations paid lip service to the concept. The most advanced post hoc procedures, post-implementation reviews, were in the public sector (Project Three) where they are often mandatory, although even here the implementation of post-project reviews was fraught with practical difficulties. On the other hand one health authority developed and implemented a well-thought out benefits management project.

7.2. Non-conventional methods

A range of evaluatory processes is in use. These are unconventional in the sense that they do not fall within the recognised techniques in IS evaluation, although familiar in other contexts. Examples of such techniques came from Project Two. A manufacturing organisation set up a ‘panel of experts’ to deal with the technical and financial evaluation of six product offerings; a financial services organisation which carried out a requirement by requirement analysis, with some RAD (Rapid Application Development), which
amounted to a form of Value Analysis. Adversarial methods were implicit in many organisations. In one organisation a “project board” had a role as arbitrator of disputes which could not be settled at a lower level. The resolution of these disputes remained adversarial, with each side putting its case for the board to decide. If a decision could not be reached, it was likely that more senior people would be drawn in, thus keeping the process adversarial.

7.3. Decision-making criteria

All of our three research projects took place against the backdrop of distinctive economic climates. Throughout, however, the need to remain competitive by cutting costs was seen to be paramount and often the way to cut costs was by cutting staff. But towards the end of the research there were some signs that the single-minded cost-cutting approach was changing and with it the criteria for formal evaluation. In the later interviews additional factors such as control, customer value and strategic benefits were part of the discussion. For example, one organisation was looking at projects which added value and/or enhanced control; another had a system for project selection and prioritisation which allowed the portfolio to track the current strategic thrust; another used criteria for acceptance which were adapted as time went by to reflect new strategic thinking.

In one IS project the decision-making criteria proved to be the heart of the problem. In this project different criteria were taken up by different sections of the organisation. As we saw it the battle was not so much about the criteria in themselves, but that one set stood proxy for the established view of the business, the other for change and a progressive view.

7.4. The levels at which decisions are made

Many evaluation methods focus on single projects. However in practice evaluation is conducted at many levels. Writers within IT who address questions at levels other than the individual project include Barua and Kriebel (1995, Single Business Units), Dos Santos et al. (1993, market value), CCTA (1992, portfolio analysis), Monk (1993, infrastructure), Parker and Benson (1988, Information Economics and linkage), Rivard and Kaiser (1989, Value Analysis), Ward (1990, portfolio), Weill (1992, firm performance). Many of these problems are addressed outside the ‘traditional’ IT investment evaluation literature. According to Galliers (1995),

The concept of considering investments of a synergistic nature as a ‘bundle’ rather than as individual, isolated investments within a bundle, is gaining considerable ground in accounting circles.

Just as for Evaluation Theory in general there is a case for applying this wider theory to IT investments. While in practice organisations attempt to carry out evaluations at different levels, the current IS literature tends to say little about evaluation in terms of particular IS related issues such as IT and IS strategies, infrastructure investments and closely interrelated projects.
7.5. The purposes of evaluation

Appraisal and evaluation are not simply undertaken for the purposes of deciding whether or not to initiate or continue a project. As well as feedback studies organisations may undertake feed forward and predictive studies. Examples provided by our studies included both a feed-back study assessing the performance of an existing system as well as a feed-forward study assessing what could usefully be brought to bear on a proposed new system from existing system components: hardware, software, training and experience (Canavet, 1995).

Evaluation may be summative, emphasising performance and attainment of objectives, or formative, that is designed to illuminate and learn. Methods like ROI are essentially summative, their purpose is accounting and control. Formative evaluation is a common notion within the discipline of “evaluation”. It has recently gained ground in IS writing most recently, as we have seen, in Walsham (1999) and Remenyi (1997). In Project Two formative evaluation was often asked for in meetings and interviews by senior managers. However in the constricted climate of recession (early 1990s) and severe cost cutting (mid-1990s onwards), and the consequent potential for job-losses which affected most of the organisations, this was not usually taken up.

As Walsham (1993a,b) and others have pointed out, evaluation serves many purposes. Decision making is undoubtedly one of them, but evaluation needs a more complex model. Such models exist, for example models of judgement and decision making in the field of social choice (for example in Elster and Hylland (1989) and Arkes and Hammond (1986)). Like Evaluation, these models need further elaboration in the context of IS evaluation.

8. Learning theme 4: project dynamics

An organisation may want to evaluate a project at any of several stages in its development and implementation. Before a project has been defined in any detail there is a need for evaluation when a decision to outsource an IT function is taken. The “outsource or not” decision provides the context in which projects go ahead and has a substantial impact on them. Likewise evaluation is needed when strategy is being developed. An IT strategy should be being developed alongside the business strategy and the role of IT will have to be assessed in general terms before either strategy can be finalised.

Once a specific project has been defined, detailed evaluation commences. At this stage the project has usually to be cost justified in the context of other capital investments. Further evaluations take place as internal and external changes may potentially affect the continuing feasibility of the project, and when it reaches the point of “sign off” when responsibility is transferred from the IT department to the user department. Immediately after the project is implemented, the system is checked to ensure that it is working as planned and delivering anticipated benefits. When the project has been in operation for some time, evaluation monitors the project’s impact, compares actual costs and benefits with planned, identifies unexpected benefits and costs and records lessons for the future. Finally, when the project is nearing the end of its life, the feasibility of replacement options is being investigated.
Kumar (1990) has suggested that organisations are not clear about the different purposes of evaluation and may confuse these stages. For example senior management may assume that the “sign off” incorporates a substantial evaluation when in fact it is a corporate signing ritual to give closure to the project. It may be preferable to make evaluation a continual process throughout the lifetime of the project.

One of the most rewarding findings of Project Two was the increased awareness of the vulnerability of projects over time, the volatility of their value and the ability to differentiate some of the causes. Projects were vulnerable to cancellation at any stage.

- At the initial (feasibility) stages they had to compete with other projects for resources, both money and skills, in a situation where there were ever-tighter controls on projects accepted for development.
- Only two of the projects accepted for development reached the implementation phase. One project fell when the organisational procedure it was supposed to support was scrapped at very short notice. A second could not proceed because the people concerned were not able to agree the numbers which were to go into the business case. A third could not continue following a review of several business cases.
- One project stalled when the systems supplier collapsed. Another was held up (although not for long) when management decided to change supplier.

The value of projects could change dramatically. The value of one system dropped very suddenly when the market which it was intended to support collapsed. Parts of the system were implemented, but the project had an operational rather than a strategic focus and finished up as a feeder system in a larger scheme. Its value in the end came from the learning experience it provided in developing systems and handling suppliers.

Investments took time to customise and deliver. Meanwhile value seeped away. For example, one project was intended to replace an existing, but outdated purchasing system. Delays to this project meant that savings from the new system were delayed too. In another organisation a system required major customisation effort at each local site, again delaying implementation.

Interest and attention from senior management fluctuated. For example in one organisation it was hardly surprising that senior management became pre-occupied elsewhere when the organisation was subject to a hostile take-over bid. In another the organisation itself was split in two parts and again senior management attention was “otherwise engaged”. Interest in evaluation itself and evaluation activity fluctuates across the life-cycle (Savolainen, 1995).

9. Learning theme 5: management learning (or not learning)

A number of the projects we investigated were to some extent failures. Often and surprisingly, the primary causes of failure stemmed from factors and errors which are well known and often feature in development methodologies. In practice they are sometimes ignored with disastrous results. Here is a list of the main ones:

- Project boards (whose existence was mandatory in the public sector projects of Project
Three) could learn from analogous projects but we came across virtually no attempts to do so.

- Business cases should include all benefits: tangible/intangible, certain/uncertain. Then the benefits can be managed and realised. But benefits are frequently omitted because they are difficult to handle, politically sensitive or hinder the approval procedures.
- We found few formal processes for realising benefits (cf. Ward (1990)).
- For a major strategic project, affecting the core of the business, evaluation needs to be a more or less continual process.
- Costing estimates should include the full cost of ownership not just the hardware.
- Organisations should respond to known risks. For example, one organisation identified a series of risks in a new strategic plan but went ahead with it without making adjustments or developing contingency plans. It was as if noting the risks was sufficient.

This inability to learn from the experiences of others has also been well-researched. Even when the lessons are clearly put before the relevant managers they prefer to go their

<table>
<thead>
<tr>
<th>Theme</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation theory</td>
<td>Many techniques are available, each suitable for particular circumstance</td>
</tr>
<tr>
<td></td>
<td>We have developed and validated (theoretically and practically) a matching method</td>
</tr>
<tr>
<td>Stakeholder theory</td>
<td>Wider stakeholder theory is relevant to IS</td>
</tr>
<tr>
<td></td>
<td>External stakeholders can play decisive part</td>
</tr>
<tr>
<td></td>
<td>Internal stakeholder play different roles; but roles can change from project to project; stakeholder relationships outside project affect project</td>
</tr>
<tr>
<td></td>
<td>There are few attempts to analyse the stakeholder role</td>
</tr>
<tr>
<td></td>
<td>Systems development and project management methodologies too often pay little attention to stakeholders</td>
</tr>
<tr>
<td>Decision-making and evaluation</td>
<td>Formal evaluation methods in increasing use</td>
</tr>
<tr>
<td></td>
<td>Non-conventional methods also used</td>
</tr>
<tr>
<td></td>
<td>Decision criteria change over time and within organisations</td>
</tr>
<tr>
<td></td>
<td>Decision-making happens at various levels: projects should be seen in bundles not in isolation</td>
</tr>
<tr>
<td></td>
<td>The “go ahead” decision is just one of many purpose served by evaluation</td>
</tr>
<tr>
<td>Project dynamics</td>
<td>Project value can seep away as strategy changes</td>
</tr>
<tr>
<td></td>
<td>A project is vulnerable as its value changes</td>
</tr>
<tr>
<td></td>
<td>the attention of senior management to a project fluctuates</td>
</tr>
<tr>
<td>Management learning</td>
<td>Managers/organisations do not learn from the failure of others</td>
</tr>
<tr>
<td></td>
<td>There is little intra organisational learning from project to project</td>
</tr>
<tr>
<td></td>
<td>Incomplete evaluation contributes to project failure</td>
</tr>
<tr>
<td></td>
<td>Strategic thining/alignment falls off over time</td>
</tr>
<tr>
<td></td>
<td>Central project boards perceive projects differently from regions/branches</td>
</tr>
</tbody>
</table>
own way, learning only from their own mistakes of which there are inevitably many. See for example Armstrong’s (1998) review of this literature.

Other emergent learning needs are more fundamental and would require significant shifts in mindsets and current thinking about projects. Three common assumptions about IT projects need to be challenged. First, strategic thinking, i.e. considering the whole business not just IT, was not always maintained throughout projects, or even present at the outset. Senior management’s assumption that responsibility for projects can be delegated to less senior members of staff on a project board often means the original strategic thinking, and the ability to maintain it, is lost. Second, project boards which are often located centrally within an organisation usually perceive projects differently to the ultimate users who may well be on the periphery i.e. in branches or local offices. The dissonance between the two groups can be the root of failure; a cultural analysis of organisations/offices in which the system is to be implemented would be helpful. Third, notwithstanding the inevitable and numerous RFCs (requests for change), projects are usually seen as “fixed” from the moment the go-ahead is given. Therefore the task is seen as being to implement the specification on time and within cost; but major projects exist in a fluid environment and methodologies would better focus on emerging strategy and benefits, rather than requirements and costs.

10. Conclusions

The purpose of this paper has been to present a learning process which draws together different themes from our own research and the literature. The themes were summarised in Table 2. Overall we summarise by saying that the practice of IS evaluation has improved but is not yet a tame problem. A “tame” problem is “one which can be specified in a form agreed by any relevant parties, ahead of the analysis, and which does not change during the analysis”. A “wicked” problem, by contrast, is “ill-defined” (Rittel and Webber, 1973 as quoted in Rosenhead, 1989). The evaluation of IS is a little ‘tamer’ than it was only a few years ago. The technology itself is more familiar, as are its uses and applications. Competitive pressures have forced management to take a grip on IT budgets and to compare them with other areas of corporate expenditure. In the organisation we visited latterly, the mystique which protected some budgets is no longer credible. Evaluation is demanded and done.

What, then, are the continuing sources of wickedness? We suggest:

*Change*. There is usually substantial change within projects throughout the life-cycle. The project is an intervention in an organisational context and there are two types of change to take into account. The first is the ‘natural’ change that one would expect from a complex, creative activity, the kind that would be dealt with in routine if occasionally acrimonious meetings and RFCs. It is the nature of the intervention which changes. The second, which has shown up clearly in this research, is contextual change: to markets, strategy, structure and policy. Most of the projects were wicked on both dimensions.

*Multiple stakeholders*. There are many relevant stakeholders in a major project and this leads of itself to complexity. The research showed that the stakeholders often stand in multiple relationships with each other (multi-strandedness). The choices they make may
be influenced not only by the project in hand but by all the other connections. When there are many relevant stakeholders the decision that is being made is a social choice as opposed to an individual choice.

Multifactorial problems. Tracing cause and effect relationships is extremely difficult. The introduction of a new system is always accompanied by other changes, for example: new work practices in BPR, new marketing initiatives, new business relationships using EDI. Such changes interact with those implemented in the new systems and with each other. Singling out one factor, the new system, for evaluation is at best conditional and at worst counter-productive as managers attempt to control the part instead of managing the whole. Similarly, projects may have intricate links with other projects in a major programme. The value to the organisation is ultimately the value of the whole, but this may be too complex to comprehend. There is a danger of sub-optimisation.

Emotion. It is at least possible that emotion plays a greater part in decision-making than has been recognised in IS writing. At Project Boards people were there to represent their own patch and, where necessary, formed alliances and coalitions. However, the emotional overtones of some of the discussions were also evident. People were afraid of losing their jobs, or that someone close to them would lose theirs. Project board chairs often played on an emotional palette, not a political one, to gain commitment. Fear, embarrassment, pride were the levers, not power, or not only power.

At the time of writing IS initiatives in many organisations are dominated by the Year 2000 problem. This work must be done—there is no room for argument. As a result evaluation has taken a back seat. However, the issue will arise again. The continuing
wickedness of the area is still present and further research will prove valuable in the longer term. We now turn to the future. How might the area develop in the future?

The first step, which is a limited one, would be to consider further the insights provided by evaluation research. The approach would be to see it not just as a source of methodologies, or indeed a justification for any one kind of methodology, but as providing an overview of the whole panoply of evaluation methods, together with the philosophies and world-views from which they spring, the assumptions they depend on and the experiences of the evaluation profession in using them. Information Systems evaluation would as it were slot into the boxes, not only categorising methods in a sound theoretical context, but enabling the identification of gaps. The work of House (1980) indicates the direction of this research.

As a second step, we would suggest a change in perspective. We see a need to re-integrate evaluation back into its organisational context as a part of organisational behaviour. In this we are following, with a great deal of hindsight, the lead given by Dent and Ezzamel (1987) in accounting. In other words we want to propose that IS evaluation, like accounting, is not an isolated activity, or a detached intervention, or something that can be understood in the context of a single project, but one which is inextricably linked with other organisational activities. It is part of the fabric of organisation as is accounting or the coffee machine.

Besides the broad areas suggested in the paper our work has revealed several narrower issues for further research and these are summarised in Table 3. These issues are grouped according to the five learning themes and follow directly from the discussions in the relevant sections of the paper.

Acknowledgements

Some of this work was supported by the ESRC research grant R000 23 3758. The authors are most grateful for this support.

References


Emery, J.C., 1976. We Can Implement Cost Effective Systems Now, EDUCOM, Princeton, NJ.


