The Learning Way:
Evaluating Co-operative Systems

Magnus Ramage

Abstract

Co-operative systems are combinations of technology, people and organisations that facilitate the communication and co-ordination necessary for a group to work together effectively. The technology involved is often, but not necessarily, groupware / computer-supported co-operative work (CSCW). The evaluation of such systems is highly difficult, especially because of the widely differing views and disciplines involved. This thesis suggests that this difficulty has partly been caused by attempts to find a single measure of 'right' or 'wrong' in evaluating these systems. Instead, it advocates evaluation as the facilitation of an organisational learning process, with learning outcomes for all stakeholders arising during the evaluation as much as from the evaluation report. Such evaluation cannot be confined to a single perspective and means of analysis – concerning technology, organisational structure or culture – as to do so ignores the inter-dependencies of the issues involved.

To solve these problems, I have developed the Systemic Evaluation for Stakeholder Learning (SESL) methodology. The principles and practice of this methodology are presented here. It exists in four main forms: as five steps to be followed in order; as a cycle of inter-relating steps; as a set of questions; and as a 'pro-forma'. The methodology rests on experiences in six evaluation studies – two of groups of academic researchers, and one each of academic administrators, health service professionals, students, and financial services administrators. It builds on evaluation studies reported in the CSCW literature; and on literature about organisational learning, systems thinking, evaluation research and stakeholder theory. Key findings from the studies include the limitations of ethnography; the importance of multiple perspectives; the value of multiplicity of discipline; the importance of encouraging learning for all stakeholders; the role of the evaluator as active contributor rather than passive observer; and the usefulness of cutting across organisational taboos.

Acknowledgements

Contents

Figures

1.   Introduction

1.1. The purpose of the thesis

This is a thesis about how to evaluate co-operative systems. These systems are combinations of people and technology within organisations, where the technology mediates the co-operation between the people in pursuance of their work. They are related to, but wider than, the technologies of groupware and computer-supported co-operative work (CSCW), although it is within the research domain of CSCW that this thesis is situated.

Such evaluation is generally regarded as a highly complex task (Ross et al, 1995; Bannon, 1994; Grudin 1988). It is difficult in terms of method (because the methods of usability evaluation have developed in the context of single rather than multiple use), practically (the effects of the change can only be seen in the longer term and often in different places), psychologically (systems designers and evaluators must make a Copernican-style shift from just-the-technology to technology-in-use-in-an-organisation) and politically (for once one begins to look at the evaluation of organisations, many issues of priority and conflicting interests arise).

This is particularly apparent given the social embeddedness of the computer systems: in information systems in general, co-operation takes place *using* the computer; in co-operative systems, that co-operation takes place *via* the computer. This means that any organisational conflicts or clashes in individual personalities and cultures will not only be more readily apparent but will directly affect how well the system works. It may well be the case that a computer system will be designed perfectly, with all the right sort of software engineering procedures, requirements analysis, and usability testing, but that the system is introduced insensitively, or it cuts across the way people have become used to working, or it changes the power relationships between workers at different levels of the organisation. All of these have been well documented in case studies in the CSCW literature (e.g. Bowers et al. 1995; Grudin 1988; Ramage 1994). In such situations, it may be the case that the system is not used by sufficient people to attain the critical mass Grudin points out is necessary for some systems to be useful. Alternatively, it may be that people devise work-arounds that bypass the computer and feed it information later, as in Bowers' study. It may be that people grumble about the "Big Brother" systems but have no choice but to use them (see section 3.1).

One of the things that makes CSCW evaluation so fraught is the wide range of different perspectives that need to be brought to bear: usability, individual psychology, group dynamics, the efficiency of communications, the effects of and on organisational structures and cultures, and many others (Ross et al. 1995). All too often, most of these are excluded, and a narrow disciplinary view is taken (according to the backgrounds and skills of the evaluators). This leads to the importance of multiplicity of theories brought to bear on the evaluation, to counter this and ensure the whole range of experiences with the system be considered.

A further problem is the dominance of the views of experts – the archetypal 'scientists in white coats' – over those of the people who actually use the systems. In fact, both are necessary: an awareness of the theoretical, analytic side to the evaluation (based on scientific analysis of users' experiences) but also of the users' own perception of their experiences. An earlier piece of work I was involved with proposed a framework based on multiple perspectives – those of users and evaluators – on evaluation, and methods to allow both to be reflected (Ross et al., 1995).

The question of multiple perspectives needs further unpacking, however. The category "user" is one that has become increasingly disputed, as ignoring the reality of people's work and shoving them into a category whose focus is on the computer (Grudin 1993). Particularly interesting are the different needs of the many groups of people with a stake in the nature and effects of a CSCW system; and how these different needs change the way they evaluate it. For the print-room staff in Bowers et al. (1995)'s study, the workflow system imposed upon them interfered with their work, and made it less efficient and interesting; but for their managers, it provided useful data about how they were working. The question of multiple stakeholders in the evaluation thus arises.

A third kind of multiplicity that can be seen to be a problem is the range of activities carried out under the banner of evaluation. For some people, it is a process of making things better during the development of a (software) system; for others, a process of telling whether a piece of research is of sufficiently high quality; for others a process of deciding what piece of software to buy; and for still others a process of figuring out whether a change in one's organisation and/or technology has had an effect, and what that effect is. Few evaluation methods seem to be aware of all of these potential meanings of evaluation, rather giving preference to one or another. There is nothing wrong with this in principle, given that most methods are limited in their scope; but there is a problem that they often tend to act as if theirs was the only kind of evaluation that existed, and to ignore the other kinds.

This thesis will outline a methodology[1] to overcome these difficulties, and describe my experiences with the use of that methodology. First, though, it is necessary to examine the concepts of evaluation and co-operative systems in more detail.

1.2.   *What is evaluation?*

Evaluation is a word that requires a certain amount of caution, meaning many things to different people. A good starting point is a definition of Elliot Stern (quoted in Sommerlad, 1992):

> Evaluation is any activity that throughout the planning and delivery of innovative programmes enables those involved to learn and make judgements about the starting assumptions, implementation processes and outcomes of the innovation concerned.

Stern is concerned with the evaluation of educational, social and organisational programmes rather than of computer systems. However, the extent to which computer systems are embedded in, and shaped by, the social systems within which they are situated means that such an approach will be of considerable utility to this sort of evaluation.

In computing circles, evaluation tends to be used to mean something like "studying a computer system [in the narrow sense] in use with a view to making it better" or "determining whether a computer system fulfils a certain set of criteria". (These two meanings are often referred to as *formative* and *summative* evaluation.) The criteria, or thing being made better, may relate to software engineering questions (efficiency, the fact the system works etc.), in which case it is pretty close to testing; or they may relate to usability issues; or they may relate to the meeting of the requirements specification. The way in which the evaluation is performed depends on the tastes of the evaluator, but it often tends to go on in isolation from the 'real world' of work, in labs and the like. For example, software companies tend to do

---

[1] The word *methodology* is used here to denote a meta-method, a method for constructing other methods (instead of meaning either the same as method or to mean the study of methods).

their usability evaluation in specialised labs with one-way mirrors, semi-realistic tasks and the like.

A slightly wider use of the term comes from those areas of computing that have a direct concern for the world of work, principally HCI, CSCW and some schools of requirements engineering. In those fields, the term evaluation often refers to the same kinds of goals as those given above, but with criteria that are based on the effects of the computer system on the organisation in which it is used. This gets much closer to the systemic perspective discussed above.

A rather earlier use of the term comes from those who evaluate educational and social change programmes. Again, the chief distinction here has been between formative and summative evaluation – making the programme better or determining whether it has been a success according to some criteria. An enormous amount of debate has gone on among practitioners about the appropriate kinds of methods; some of this is summarised in section 2.3.3.

On the whole, this debate has not had much influence on evaluation in computing, with the exception of work in computer-aided learning – this is principally due to its close links with education research, that has led to work such as the MEDA toolkit (Machell and Saunders, n.d.). There has been evaluation of the large IT research programmes, such as Alvey, but this has tended to be summative and led to little direct feedback to the research community.

This model of evaluation in this thesis is essentially a formative one, with an aim to contribute to the building of better co-operative systems. It may be the case that the evaluation does not *directly* contribute to a given system, but all evaluation is ultimately formative in that it feeds back into future systems design. This approach is clearly at the organisationally focused end of the evaluation spectrum in computing; but is also informed by the introduction of ideas from programme evaluation into work in computing. The perspective on evaluation being presented here can be summed up in the phrase: *evaluation is no good if it just considers the computer. The situation is also up for evaluation*.

1.3. *What are co-operative systems?*

To evaluate a class of objects, it helps to have some understanding of what those objects are. As this thesis is about the evaluation of co-operative systems, it is necessary to consider the meaning of this term. In its usage in Lancaster, it is simply used to denote the computer programs sometimes called groupware or CSCW systems, and it has purely arisen as an easier and less loaded term than those. Sommerville et al (1993) define the term as: "systems which [are] essentially co-operative in the sense that they [are] team-based".

However, it is clear that this usage is a neologism, and that both 'co-operative' and 'systems' have a long history in various areas of academic and practical discourse. Looking at that history briefly will show the meaning behind the words.

Co-operation, first, literally means "working together" (from the Latin "co", together; and "operare", to work). This is at the root of all modern uses of the term – it denotes some kind of activity conducted between two or more people, usually to common gain. Much work has gone on in social psychology in studying the subject, and from this Michael Argyle (1991, p.4) offers the following definition of co-operation:

> acting together, in a co-ordinated way at work, leisure or social relationships, in the pursuit of shared goals, the enjoyment of the social activity, or simply furthering the relationship.

In this descriptive sense, then, co-operation refers to any sort of activity that two or more people conduct together. Studies of such situations have been made by sociologists (e.g. Heath and Luff, 1991), by anthropologists (e.g. Hutchins, 1991), and by social psychologists (e.g. Axelrod, 1984). A point that must be made about the last of these three groups is that the studies of co-operation in social psychology have tended to use artificial laboratory situations where only a limited amount of co-operation is permitted (such as the much-used but limited – because of its artificiality – Prisoner's Dilemma), whereas the sociologists and anthropologists study real situations of co-operation.

Nevertheless, the following general principles about co-operation arise:

- it requires communication between those co-operating;

- it requires the establishment of shared understanding (cf. Ross et al., 1995) and goals between those co-operating;

- it is often (but not always) the case that there is some kind of benefit, not necessarily material, for all those taking part ;

- effective co-operation requires effort on all parts to maintain;

- awareness of others' actions, thoughts and feelings is important;

- if it works well, it can be immensely productive and personally satisfying.

This is essentially descriptive. There have also been those keen to be prescriptive of co-operation as a good in itself: this has happened among social justice campaigners, assertiveness and negotiation trainers, and that part of the political left who have remodelled business as a partnership of mutual owners, under the name of the co-operative movement. It is often contrasted with conflict, as illustrated by a cartoon used by a number of justice groups. Two mules, joined by a rope, are shown, each with piles of food. The rope is too short to allow them both to eat from their own piles at once. They strain against each other for a while, resulting in neither of them eating anything, then realise their folly and join together: first they both eat from one pile of food, then from the other. The caption reads: "Co-operation is better than conflict".

This contrast is not always helpful. As Howard (1987:176) has written, "to the degree that the idea of 'co-operative work' neglects or underestimates the role of conflict in working life and in the design and use of information technology, it leaves out a major (indeed, central) dimension of work experience". That is, if one ignores the fact that within systems of work that involve co-operation, there will also be conflict between the different people and groups involved, one will be in danger of ignoring the needs of many of those groups. This was precisely the criticism that Ehn and Kyng (1987) levelled against the socio-technical approach: that in the name of industrial harmony, it perpetuated the existing power relations and imbalances. This question is considered below in terms of the multiple stakeholders relevant to a co-operative system.

Co-operation is also not entirely positive. A great deal of co-operation, of precisely the sort described, can be found within armies, terrorist groups, the Mafia and other organisations which are at best of questionable value and more often morally wrong. In these organisations, co-operation takes place *within* the group, but it is harnessed to ends that are harmful to those *outside* the group.

Moving now to the term "system", it is important to distinguish carefully between two uses of the word when considered in this context. The word is in use by computer scientists and organisational theorists to denote a collection of computer hardware, software and networks –

one talks of a computer system and means this kind of technological mixture. Indeed, many researchers on computers in use collect themselves under the label of "information systems". This overshadows in computer circles an older and more general use of the term, which has been in use since around 1940 – the use of systems thinking.

Systems thinking is an approach which views the world in terms of models of real-world situations which have the common property that "the whole is greater than the sum of its parts". That is, there are properties of the entity viewed as a whole that are not to be found by considering the constituent parts of that entity. A good example is given by Lewis (1994:44), who considers a bicycle. This is composed of a number of pieces – two wheels, frame, handlebars, chain, saddle and so on – but taken separately none of these has any particular meaning. However, by combining the pieces together in the right way, it is possible to create a system that affords transport. That is, the ability of a bicycle to carry someone to work (given motor power from their legs) is an *emergent* property of the complete system. As Senge (1990:68) puts it,

> systems thinking is a discipline for seeing wholes ... a framework for seeing interrelationships rather than things, for seeing patterns of change rather than snapshots.

This older notion of system – from which the computer sense arose – is intended as a general perspective on all kinds of entities. It arose in biology, through the work of Ludwig von Bertalanffy (1969) and while to some extent, as Morgan (1986:45) comments, it can be seen as a "biological metaphor in disguise", it is taken by its theorists to be considerably more general.

Two distinct kinds of systems thinking have arisen as the discipline has developed. The first is mathematically-based, likes to create formal models of situations, and finds it useful to draw analogies between human and social systems and the better understood mechanical and cybernetic systems; it has been dominant in operational research and in systems analysis, and through the latter continues to be influential in the design of information systems. Another strain can be found in the work of those who take a holistic perspective on the world, requiring that mind and nature not be separated, and likewise that situations can best be understood by studying them from all sides. Checkland (1981)'s soft systems methodology stresses the importance of this kind of holism, and how vital it is that multiple perspectives be brought to bear. This approach also stresses a crucial point about systems: that they do not exist as such in the real world, but are rather ways of viewing the world – they belong to epistemology rather than ontology.

This is crucially important, as it emphasises that different systems will be identified by different people, and so what is meant by it must be clear. For ex4ample, some use the phrase "the computer system" to mean a standalone PC (box containing chips plus monitor, keyboard and mouse); others include software on the PC (perhaps just the operating system or perhaps also the word-processor etc.); others include networking to other PCs; and others also include fileservers on a local network, the files on them, the other computers, and even (implicitly) the technicians that keep the system running.

Finally, a mention of the phrase "socio-technical systems" must be made. Deriving from the work of the Tavistock Institute of Human Relations in the 1940s, it refers to "the interdependent qualities of the social and technical aspects of work ... these aspects of work are always inseparable, because the nature of one element in this configuration *always* has important consequences for the other" (Morgan 1986:44).

So to summarise the above material – co-operation is a process of two or more people engaging in an activity for shared gain, supported by communication and co-ordination; and a system is a collection of objects with emergent properties, here involving people and technology. Putting these together gives the following definition for 'co-operative system':

> a combination of technology, people and organisations that facilitates the communication and co-ordination necessary for a group to work together effectively in the pursuit of a shared goal, and to achieve gain for all its members.

A few comments can be made about this definition:

- technology is seen as one part of the system, but co-equal with the organisational and human structures also necessary for co-operation;

- the role of the system is the facilitation of other activities – it is a means to an end, rather than an end in itself;

- unless there is a shared goal and gain for all members, it is questionable whether co-operation is taking place. This arises from Argyle's definition of co-operation (above), and the observation that people seldom do things unless there is some gain for them (although the gain may be, for example, a feeling of well-being when acting altruistically, or the prevention of undesirable results).

The difference between co-operative systems and the more traditional term 'information systems' might at first sight be a subtle one. Are there really such things as information systems that do *not* involve co-operation? The shift in perspective here is one of focus –to be concerned not with the *content* of the systems (which is still information), but with the *process* by which people use the systems. This view looks at relationships rather than things, arrows rather than boxes. The interest then rests on people and their moment–to–moment activities, rather than on abstract and well–defined entities that can be labelled. *Inter alia*, this means that the shift from information systems to co-operative systems recognises the significance of the whole socio-technical system surrounding the software: for if co-operation between people is the purpose of such systems, then the people, the work they do and the setting in which they do it become even more crucial. With information systems evaluation in general, it is possible – although foolish – to only look at the technology. With co-operative systems evaluation, deeply embedded in organisations and their members, this becomes nonsensical.

1.4.   Outline of thesis

The rest of this thesis will be attempting to answer the question:

> Why are existing CSCW evaluation methodologies so problematic? How could they be better?

The brief answer to this question, to be demonstrated here, is:

> They are problematic because they ignore the needs of multiple stakeholders, concentrate on the technology rather than the whole co-operative system, and fail to recognise the different purposes to which evaluation is put (while assuming that it is essentially a scientific activity). They could be better by a refocusing of evaluation as a process of facilitating learning among all stakeholders, taking into account their different perspectives; by looking at the technology-in-organisation as a co-operative system; and by an awareness of the many purposes of evaluation.

Five chapters will follow this introductory chapter. They are particularly concerned with a development of a methodology to solve these issues: SESL (Systemic Evaluation for Stakeholder Learning). The chapters have the basic structure of background, development of the methodology, description of the methodology, evaluation of the methodology, and conclusions.

Chapter 2, *Background*, covers lessons learned from elsewhere. It discusses existing methods of evaluation, from CSCW and other fields, categorised by the types of data they incorporate: ethnography (and interpreting frameworks for it); other qualitative methods (interviews, questionnaires, focus groups); psychological methods (lab testing, analytical approaches); systems building methods (iterative prototyping, beta testing, user testing, heuristic evaluation); taking advice (fashion, consumer reports, marketing literature); meta-methods (such as SSM and PETRA). Following this, the chapter describes studies from the CSCW literature which have been carried out using these methods, categorised by the purposes the evaluations were carried out for – studying effects, formative evaluation, conceptual, focusing on people, and buying new software. Finally, it brings in lessons from elsewhere to solve the problems of co-operative systems evaluation, particularly from stakeholder theory, organisational learning and evaluation research.

Chapter 3, *SESL under development*, discusses the way the methodology developed in practice during the course of the work. It first describes four short case studies which showed the need for SESL, which especially demonstrate why evaluation means different things to different people; why it needs to be as broad as possible; and that the organisational context can completely determine the usefulness of a computer system. Following this is a description of a much longer study which led to further developments of SESL, especially in the widening of the use of stakeholders (from just the binary opposition of PETRA) and the consideration of evaluation as a set of multiple learning processes for many different stakeholders.

Chapter 4, *SESL: a methodology for evaluation* describes the methodology in its full form. This begins by discussing the theory behind the methodology, its essential principles (systems thinking, stakeholder perspectives, and evaluation as learning), the integration of which into a single framework is a key part of the novelty of this thesis. The methodology itself is described in several forms: a simple five-stage process, a more extended form of this, a 'pro-forma' guide, and a set of simple questions. Finally an extended practical example of the methodology is given, through the most mature case study conducted during this research. It clearly shows the need for the tools of the methodology.

Chapter 5, *Evaluating the evaluation*, moves to evaluate the SESL methodology. This is done in two ways: practical and comparative. In the first place, it discusses implications from the various studies, learning points for future work and use of the methodology. Second, the chapter outlines (and justifies) a set of criteria for the evaluation of a evaluation method or methodology (e.g. usefulness, practicality, coverage), show how these apply to the various methods described in chapter 2, and then how SESL fulfils the criteria.

Chapter 6, *Conclusions*, completes the thesis. It begins by restating the case for SESL, by describing its characteristics and what it offers to a practising researcher. Following this is a discussion of future work which might be undertaken on the methodology. The chapter closes a discussion of the journey undertaken so far in the development of SESL, how the methodology has changed, and how it is still changing.

There follow six appendices. The first, appendix A, is a short guide to SESL in the form of a set of questions, which is introduced in section 4.2.4. The other five, which take various forms (a published paper, two reports given to the organisations being studied, a detailed description

not given to the organisation at the time, and a piece of text from an earlier work), are more detailed descriptions of the case studies discussed in chapters 3 and 4.

2. Background

This chapter discusses the literature that has formed the background to this research. Part of this literature has directly fed into the development of SESL; other parts of it have informed SESL through reaction to its ideas. First, it discusses existing methods of evaluation, most of which have been used within CSCW; and then the evaluation studies within CSCW which have made use of those methods. It then moves on to consider literature from outside CSCW which has proved useful in the development of SESL.

2.1. On existing methods of evaluation

2.1.1. Ethnography

The dominant method of performing field studies in CSCW has been ethnography. This involves intensive and detailed study of a real workplace over a prolonged period. Data collected include audio and video-tapes of work practices, field notes as to the most significant practices carried out by the participants, descriptions and diagrams of the work setting, and samples of various artefacts (such as documents) which illustrate the nature of work in the organisation. The time typically runs from a few days to a few months – shorter studies tend not to be dubbed ethnography, while longer studies are rare. Traditionally (in its anthropological origins), ethnography requires a long period of immersion – months or even years – in the study setting before the ethnographer can perform an informed analysis. Clearly, this is hardly ever practical in a real-world systems context. However, as Hughes et al. (1994) discuss, methods such as "quick and dirty ethnography" (a brief study, typically a few days, with specific questions in mind as to the nature of the work) can still provide useful amounts of data in a shorter time.

The main focus in the method is the social organisation of work and its effects on technology, but various theoretical frameworks from the social sciences have been used to inform, guide and interpret the studies. These are listed below. The exact relationship between the theory and the ethnographic data-collection varies: at the least, different methods require different kinds of data (distributed cognition will tend to focus more on artefacts while ethnomethodology on conversations).

- *Ethnomethodology*: This originated in claims by researchers of the 1960s (Garfinkel, 1967) that the sociology of the day reduced the people studied to mere ciphers, dehumanised pawns in the social scientist's attempt to prove a point. The aim of ethnomethodology is thus to focus instead on the detail of what people do on a moment-to-moment basis, in collectively constructing their reality. As used for evaluation (Randall et al., 1996), this approach can bring great benefits in understanding what it is that people actually do, although some of the views of its practitioners would tend more towards the giving-information-to-people forms of evaluation than the making-decisions forms.

- *Conversation Analysis and Interaction Analysis*: Subsets of ethnomethodology, these methods study real group interactions as revealed by their (directly recorded) conversations and actions (Woofitt, 1991; Ruhleder and Jordan, 1997). They focus on the detailed features of interaction (at various levels), either on conversations alone or on interactions between people and between people and technology. The analysis process is time-intensive, given the detailed level at which it operates, but may reveal areas of interest, which will not be shown by more surface methods.

- *Distributed Cognition:* Traditionally, cognitive science was concerned with thought processes within one individual's head. No allowance was made for the effect of the environment. Through the work of authors such as Hutchins (1995) and Rogers (1992), the focus has gradually shifted to 'cognition in the world'. This approach sees cognition as being mediated by a number of different objects, including social processes, the use of artefacts and environmental factors. It tends to analyse data at a very detailed level – Hutchins' book is based on a set of extremely fine-grained studies of artefact use in ship navigation – and as such may be too much work for many studies. However, the concepts of a wider cognitive science are very useful and are proving influential in other areas.

- *Activity Theory:* This approach (which derives from the work of Soviet psychologists in the 1920s and 1930s) holds that people never interact with objects in the world (in fulfilment of some task) directly, but always via a mediating tool. It thus has considerable overlap with distributed cognition. The main advocate of the approach within CSCW has been Kari Kuutti (Kuutti and Arvonen, 1992), but recently it has spread to a number of other practitioners (Nardi, 1996).

- *Structuration Theory:* The work of the sociologist Anthony Giddens involves: "the duality of structure, which relates to the fundamentally recursive character of social life, and expresses the mutual dependence of structure and agency" (Giddens 1979:69), where "structure" refers to the rules and resources that are properties of social systems and "agency" refers to the ability of human beings to act freely. It is thus an attempt to unify the constraints of society identified by positivist sociology with the social construction by free individuals identified by interpretivist approaches such as ethnomethodology. It has been influential in Information Systems, and at crossover points between that field and CSCW. Orlikowski and Gash (1994) used the approach to good effect in their early studies of Lotus Notes.

- *Breakdown Analysis:* A breakdown is defined as any incident where the user has cause to focus on the system rather than the task (Winograd & Flores, 1986). Breakdown analysis studies group interactions and conversation transcripts to highlight such breakdowns. This is a useful method for identifying important problems associated with user-system (or user-user) communication (Urquijo et al., 1993). However, the focus is necessarily restricted, disregarding many other interesting aspects of collaborative work, such as the distribution of roles and power amongst the group members. While popular a few years ago, it has had little recent use within CSCW.

- *Others:* Other frameworks have been usefully applied within CSCW to the analysis of ethnography. These have included the work of Gregory Bateson (1972) on levels of learning, and the actor-network theory from the sociology of technology (Latour, 1993).

2.1.2. Other qualitative methods

There are a large array of other methods used in various social sciences which collect qualitative data. They all have roughly the same pattern: ask people what they feel about something, record the answer and compare it to others' views on the same subject. The methods fall into three broad categories: interviews, questionnaires and group discussions. Interviews involve one-to-one discussion between a researcher and a person from whom data is of interest; they run from the highly structured (with all questions determined in advance)

to the complete unstructured (with only a few vague topics, and otherwise working like a conversation). Group discussions (often called focus groups, a particularly common form from market research) are like interviews, only with several people at once; they are often fairly structured in their format. Questionnaires present a more structured set of issues (sometimes asking for responses on a numeric scale, sometimes in a more open format) which are presented on paper to a large number of different individuals, whose answers can then be collated and sometimes statistically analysed.

Within CSCW, these methods are especially useful to obtain data about users' experiences with computer systems (either immediately or a little while after use); or to study more general questions of the nature of the wider system and how the technology works within it. Their subjectivity (in that direct user opinions are being collected) makes them useful, but also limited (although this can be guarded against by using a large group of people, if appropriate, and by wording questions so they contain various 'consistency checks').

2.1.3. Psychological methods

2.1.3.1. Lab experiments

Laboratory experiments are quite widely used to evaluate CSCW systems (e.g. Ishii et al., 1993; Wan and Johnson, 1994; Olson and Olson, 1991). These are used to collect quantitative data about a single specific factor, attempting to screen out other influences. However, as with user testing, there are significant problems with the decontextualised and artificial nature of these experiments. They tend to assume the kind of disembodied intelligences criticised by Suchman (1987) and others, not helpful given the situated nature of CSCW systems. However, they are useful when applied to very precisely defined questions as part of a wider study. Twidale et al. (1994) discuss their applicability in the early stages of an evaluation, suggesting they can be used with easily available users (students) to throw up the most glaring usability problems, allowing the more complicated issues to be dealt with by the real users (air traffic controllers), who may only be available for short periods.

2.1.3.2. Analytic approaches

The technique of user modelling – the construction of quasi-mathematical models of interaction – appeared early in Human-Computer Interaction. The best known, and most developed, of these techniques is the GOMS approach (Card, Moran and Newell, 1983), which has been used extensively to answer questions about highly detailed issues of interface design – and thus for formative evaluation. These methods can be useful to address fine questions of detail, but are far too time-intensive to be used in general; they also (as with laboratory experiments) completely ignore the context in which the system is used. They have thus had little application in CSCW, although the AMODEUS projects (Blandford and Duke, 1997), which use the methods in combination with other techniques, took some interest in co-operative working.

2.1.4. Systems building methods

A different tack from the above methods relies upon first building a complete, working system, and then performing an evaluation of it. Such evaluations will usually be formative (intended to lead to further changes in the system). The methods come in several different varieties.

2.1.4.1. Iterative prototyping / participatory design

Iterative prototyping – building a system in small chunks which are then tested out with users and improved upon further – is often discussed within the HCI literature as a way of improving a system as you go along, as well as empowering users by giving them a say on what their system will do (Nielsen, 1993). It is thus related to the school of methods known as participatory design (an overview of which are found in the book edited by Greenbaum and Kyng, 1991, but which are discussed in detail in the successive proceedings of the biennial conference *Participatory Design*). Deriving from the Scandinavian industrial democracy movement, these methods stress user participation in the design process, tending also to simplify the process to enable this. They also tend to combine design and evaluation into a single cycle, as advocated by Bannon (1994).

2.1.4.2. Beta-testing/customer feedback

Another form of systems-based evaluation is the 'beta-testing' procedure followed by many commercial firms, where a semi-finished product is passed around a selection of users to gauge their reactions. Related to this is the development of the product from one version to another. An example of this is given by Abbott and Sarin (1994), who talk about the development of different versions of a workflow system according to the experience of users with it; the same process is discussed in the case of Microsoft by Cusumano and Selby (1995), who discuss that company's increasing reliance on the feedback it receives via customer helpdesks and comment channels, for further development of their products.

2.1.4.3. Heuristic evaluation

Heuristic Evaluation (Nielsen, 1993) relies on an evaluator's immediate reactions, intuitions and predictions, categorised under a set of design principles and usability attributes. These can be used as an intrinsic part of a heuristic evaluation, or as a useful framework for categorising interface characteristics after any evaluative method. For CSCW applications, additional issues must be considered, such as the set discussed in Ross et al (1995)[2]. To an extent, heuristic evaluation is an inevitable part of any system design process, as designers do something and then try to figure out if they like it.

2.1.4.4. User testing

Much advocated in HCI (Tognazzini, 1992), user testing generally takes the form of studies conducted by system designers with real users in a semi-realistic use context. The aim is to see how the system is used and what usability or functionality issues arise – typically, qualitative data are collected, to feed back into the design process. User testing is a useful part of the iterative design process, but care needs to be given to questions of context, as argued by Bannon (1991).

2.1.4.5. Semi-situated ethnography

Related to user testing are the methods that Plowman et al (1995) refer to as "semi-situated ethnography": those which lie somewhere between the unsituated lab experiment and the messy, real-world ethnographic study. They often involve having one's colleagues (or a similar accessible, controllable group) use a system for a prolonged period, before it is tried out on real users, and are well used within research settings. Such studies can suffer from being rather inward-looking, in that they end up focussing on their own research teams, and as

---

[2] See Appendix C.

Harper (1992:36) comments, research labs are "peculiar fish bowls" due to "the forms of working relationships one finds therein". However, such methods are often highly instructive in practice, given some degree of care as to their wider applicability.

2.1.5. Taking advice

One form of evaluation that is particularly useful for deciding which new software to buy is to ask others' advice. A whole host of different sources of advice are available in the commercial domain. It cannot be denied that one reason for deciding to buy one piece of software or another is simply fashion, but this section deals with slightly better grounded reasons.

2.1.5.1. Consumer and consultancy reports

Rather in the same way that the consumer magazines assist people in buying cars or cameras, computer magazines give reviews and analyses of what is best about the current set of groupware tools, Web servers, word processors or spreadsheets. For example, an issue of *Byte* magazine lists the relative advantages of Lotus Notes and the Web (Roberts, 1996). More commercially, I.T. consultants often publish lengthy reports with similar analyses but intended for a strategic business market, allowing them to make decisions about the impact of a proposed system on their profits. These reports are frequently sold for large sums. They are often very useful for business-people within their target market, but less useful for other readers.

2.1.5.2. Marketing literature

Software manufacturers and resellers will often provide information to allow decisions to be made. Lotus provide a "Domino Evaluators Guide" for potential purchasers of their office software (Lotus Development Corporation, 1998). This guide begins "Looking to evaluate Domino against the competition? These documents will help you discover why Domino is the best choice for your messaging, application and collaborative needs – across all platforms and supporting all standards", and goes on to list a set of reasons why they hold their product to be preferable to their competitors' products. Other software firms produce similar lists for their products, which of course contradict the lists produced by the other companies.

2.1.6. Methodologies

As the term is used here, a methodology is a 'meta-method': it is a method that lays down certain rules for the creation or selection of methods, which are then applied to evaluate a system. Two methodologies are discussed here. SESL, the methodology developed during this research and which is discussed in detail in chapter 4 of this thesis, also fits within this set of methodologies.

2.1.6.1. SSM

Soft systems methodology was developed in the management sciences by Peter Checkland (Checkland and Scholes, 1990) as a way to study and improve systems of all kinds. The 'softness' of the methodology relates to the fact that, unlike the traditional 'hard' systems methods derived from engineering and operational research, it doesn't make assumptions about the nature of the system to begin with, or the key questions being asked. Rather, it holds that constructing a view of what system one is evaluating, and understanding the right questions to ask, are important parts of forming an answer. It is often regarded as a methodology for problem analysis rather than evaluation, but as Checkland and Holwell

(1998:194) say, "since organisational interventions and social programmes are examples of would-be purposeful action, it is not surprising that SSM has been used in evaluation studies".

SSM has long been widely used within the discipline of information systems (with its roots in management schools), but within CSCW (with its roots in computer science), it has only recently come to be used. The chief criticisms to be levelled at SSM as a methodology are that it is often difficult for the untrained potential practitioner to gain enough of an understanding from the published work to make use of it; and that it requires a fair amount of tailoring to be useful in a given situation. However, for an expert practitioner, it is subtle, malleable, and very powerful as an approach to evaluation.

2.1.6.2. PETRA

The PETRA methodology (Participatory Evaluation Through Redesign and Analysis) was developed by myself and Susi Ross as a way of evaluating CSCW systems, taking into account multiple disciplinary sources for analysis and multiple perspectives on the evaluation. It is described in section 3.2 below, and more fully in Appendix C (a version of the paper by Ross et al, 1995). It is particularly useful in fairly apolitical, technology-centred environments which need a formative evaluation approach. In environments where the organisational politics are more noticeable, or where there are greater numbers of stakeholder groups, the SESL methodology described in chapter 4 is more appropriate.

2.2. Studies which use these methods

2.2.1. A categorisation

The breadth of the CSCW and cognate literature on evaluation means that a certain amount of categorisation is helpful to understand it better. This categorisation is made by forms of activity carried out during the process of evaluation.

To date, four main kinds of activity have been identified. No doubt there are others – this taxonomy is an open one to be changed, presented for the provocation of thought. In practice, they appear in combination and in modified form. The four types described here are not intended to be taken as a set of closed, complete categories – they overlap and may well be incomplete.

However, CSCW evaluation can get terribly confused (and confusing) unless one realises what type of evaluation one is conducting. Thus, there is nothing wrong in purely conducting a usability evaluation, for formative purposes, but if one believes one has completely evaluated the system and its potential effects on an organisation in that way, unfortunate consequences could result.

The following diagram, figure 1, and descriptions show the four types in broad form. They are also mapped on to a 2x2 matrix (figure 7), for analysing the type/purpose of an evaluation in SESL, section 4.2.1.3.
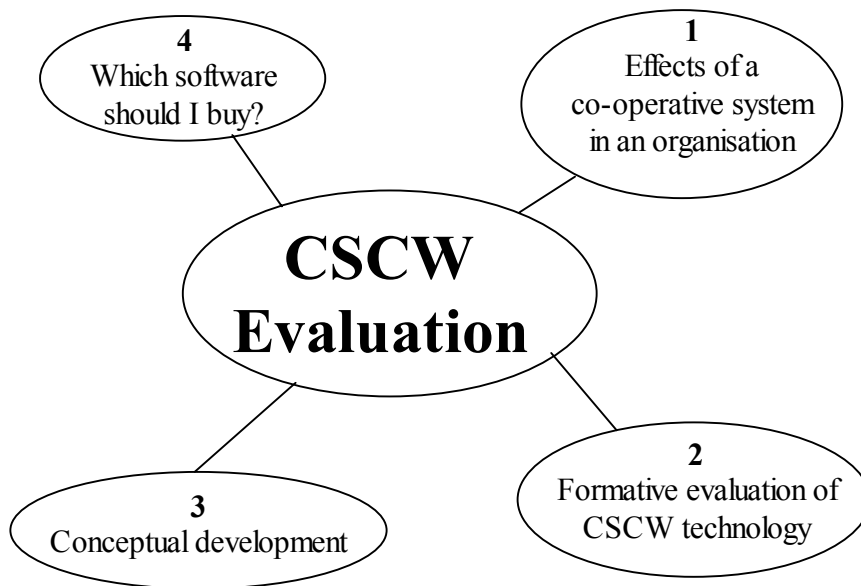
Figure 1: Types of CSCW evaluation

- *Type 1: Effects.* What happens when you introduce a new computer system into an organisation? How does it change the organisation's work, members and outcomes (such as profits)? How does the technology change for the needs of the organisation? These questions are those of evaluation in its classic form.

- *Type 2: Formative.* There are those within the CSCW community whose interest is to build systems (in the narrow, technical sense of the term) and have them be used, either for commercial purposes or within a research community. Their need for evaluation is to develop their systems further, making them more usable and more appropriate for the intended or actual users – their evaluation is a formative one.

- *Type 3: Conceptual.* Not all systems development concerns the design or use of products intended for 'real' organisations. Many pieces of development are conducted purely for their research interests. The evaluation of such systems is therefore not at a level where one is examining the effect of a system in use, or trying to redesign it for future use. Rather, what one is seeking to evaluate here are the *concepts* that underlie the system, and whether those concepts are applicable. There are four kinds of research one might wish to evaluate in this way: research projects carried out within commercial research centres; academic research projects; PhD projects; and externally funded research projects.

- *Type 4: Buying.* The question "should I buy WordPerfect, Microsoft Word or Lotus WordPro for my company?" is a form of evaluation: it is an exercise in examining alternatives, weighing up their pros and cons, and coming to a conclusion. So how does the hard-pressed IS manager make the decision as to which system to buy?

This typology will now be used to describe evaluation studies which have been discussed in the CSCW literature.

2.2.2. Type 1: Effects

The general form of this type of evaluation study is: the researcher either asks or is asked to come into the organisation; they 'hang around', watching what is going on, and perform interviews; they structure these ideas in terms of their preferred theory; and then they present

a conclusion to the members of the organisation. That is, they conduct a fairly typical kind of qualitative research.

It is usually the case in these evaluations that the researcher/evaluator is an outsider to an organisation, called in as a consultant or coming in for their own research agenda. This has all the standard advantages and disadvantages of outsiders: they can see things the organisation's members can't, as they are not bound by its norms and assumptions; they also fail to see things for precisely the same reason, and there are risks about them becoming the 'Scientists in White Coats' warned against in the introduction.

On the more positive side, the presence of evaluators in an organisation can be valuable of itself. They can bring valuable experience with computers and with organisational theory, and help members of the organisation learn about these things in little ways that enhance the use of the system they are evaluating. Indeed, the process of doing evaluation may act to provide organisational learning, by raising issues for the members of the organisation that they had not previously given thought to.

Similarly, members of an organisation conduct their own evaluations of the effects of a computer system: managers are keen to know what is going on, for example, so they can decide whether to keep it in its current form or to change it. Few of these are reported in the literature. Their form would reflect the kind involving outsiders, only with a different scope and more problematic issues about power (employees may be reluctant to tell managers a system is terrible if the instigator of that system is a powerful political operator).

A common theme of many of the papers here is that **there is a dialectical relationship between the social ordering of the workplace and the effect of new technology**: the technology alters the culture and work practices of the organisation, but is in turn altered by that culture and work practices. Given that CSCW has historically involved the entering of social scientists into an area previously mainly the province of computer scientists and engineers, the main thrust here has necessarily been one of sociologists trying to make it clear that systems cannot be evaluated effectively without consideration of the social ordering of the organisation.

O'Day et al (1996:160) make this relationship clear: "it is important to look at *social* and *technical* design elements together. There are social implications of technical design choices and technical implications of social design choices…as a designer, one makes decisions in both areas". These views are stated in a complementary fashion by Bikson and Eveland (1996), in a study of groupware implementation at the World Bank. It is clear to them that the moment when the system was put in place was not a single, concrete point in time, but rather a fluid process over a period of time of the changing of social and technical systems.

Harper et al. (1991) is a good example of this approach. The first of several papers by these authors and others concerning the London Air Traffic Control Centre (LATCC), it is a study of the use (or rather the lack of use) of an automated air traffic control system. They observe that the controllers made little use of the system at all at busy times, and when they did use it didn't trust it, and conclude this is because the designers of the system failed to take into account the way controllers worked at LATCC. They comment that "insofar as system designers wish to take co-operative working practices seriously, then they will need to know a great deal more about the social organisation of work" (p.232).

Another early paper considering such matters is that by Blomberg (1986) who argues that technology "must be understood in terms of the social environment into which it is introduced" (p.35). Studying the introduction of a photocopier interface design system (Trillium), she notes that its usage patterns are different according to whether the users were

working at the company before the introduction of the system or joined after it was introduced: those who were already there frequently used the services of a human 'mediator' between them and the technology, while those who joined post-Trillium seemed to expect that using the system was a normal part of their job. However, she comments that the organisational culture affected the technology also: it was redesigned to allow interface code to be produced directly from Trillium rather than the results being passed to a programmer for coding. This seems to have occurred not so much because of the existing relationships between designers (human factors psychologists) and software engineers, but rather with the intention of changing that balance.

A more recent air-traffic control study than the one discussed above is that by Twidale et al. (1994). They discuss their finding that the acceptability of technology is bound up with the acceptability of work redesign: the introduction of new technology often goes hand-in-hand with the explicit changing of work practices, especially given the current popularity among managers of Business Process Reengineering (Hammer and Champy, 1993). In a not dissimilar comment to the earlier paper, they say that "a 'situated' evaluation would need to address not only the capacity of the 'system' but the flow of work around it" (p.449).

Another common theme in several of the evaluations studied is **the question of by whose criteria a system can be said to be 'good'**: that is, efficacious for its desired purpose. In particular, there is often a clash of cultures between the designers of a system, especially if they are responsible for user support, and its users.

Star and Ruhleder (1994) illustrate this point well in studying the use of a system to support communication across the Internet by biology researchers studying the genetics of a particular type of worm. They observe that often a user support person will say something like "just throw up X-Windows and ftp the file down" to a user, who may not have any idea what is meant by ftp, or how to start up X-Windows. However, they also observe this kind of clash between different types of user with different knowledge not only of computer system but also of types of biology – those within the "worm community" and those outside it.

Orlikowski and Gash (1994), who studied the use of Lotus Notes in a large financial services company come to similar conclusions. They found that the IT department's main aim was simply to get the system running and to keep it running, thinking that the users would work out for themselves how to use it. Users, perhaps unsurprisingly, had somewhat different perceptions!

The same trend can be found in the ethnomethodological studies of the use of a workflow system in a printing firm (Bowers et al, 1995; Button & Sharrock, 1997). The introduction of workflow, they report, required a significant change in the working practices of the shop-floor staff, who previously worked according to local procedures based on the need to co-ordinate complicated sequences of work. In some cases the result was non-use of the system; in others a complicated series of 'work-arounds' were instituted to satisfy both the need to get the work done and the need to produce information that was useful to management via the workflow system. The questions of the cost of formalisation in workflow systems are also raised by Grinter (1997), who concludes that in certain settings (specifically those of professionals, trusted by management broadly to organise their own work), the systems may cause less friction than has sometimes been reported.

The question is turned on its head by Grudin and Palen (1995), who discuss the take-up of meeting scheduling systems, and conclude that groupware can succeed without management ordering its use, if it is easy to use and supports appropriate needs. It is important, though, that a "critical mass" of users is attracted to the technology.

Harper (1992), in a study of active badge use across two research laboratories, found that there was significant difference in its use according to the roles of individuals within the organisation: those whose job it was to keep tabs on people (especially receptionists) found them very helpful, while researchers' perceptions were that it was more of an intrusion into their working patterns.

Some studies observe that **the reason why technology was not used was in fact purely for technical reasons**. This was the case in both the studies by Goodman and Abel (1986) and by Tang et al. (1994), where the overheads for starting up a communication in the system were quite large: in the first study because it was necessary to shout to make someone come to the video wall (in a common room), in the second because the system took around 45 seconds to open up a simple 'glance' window on another user's screen. In the Tang et al. study, another technical reason for non-use was that the system was unreliable and often unavailable, so that it was not built into the standard practices of its potential users.

Ackerman (1994) comments that even in a perfectly useful system, the amount of usage may be low simply because of the nature of the system. His evaluation was of an organisational memory system designed to assist X-Windows users with questions – in many cases it may be quite appropriate, he comments, for users to only use the system once a month but still to find it useful and necessary to have to hand (in the same way as a dictionary might be used only infrequently but when it is used is important).

Finally, sometimes the focus of an evaluation study is on **issues other than the technological ones**. Clement (1990) and Sherry (1995) are two good examples: both present case studies of situations that involve CSCW use (an office automation system and a communications system among Navajo Indians); but both take as their main focus of interest questions of power balance, empowerment and authority. The case study is used as a justification for an academic discussion about a topic deeply important to the authors, on which they had explicitly worked with the research participants for purposes of their empowerment. Harper (1997) also fits into this category: his long-term ethnographic study of the International Monetary Fund has much to say about the use of information in organisations, which is highly ritualised as part of a process of decision-making, as much as it is gathered for its own sake.

2.2.3. Type 2: Formative

Formative evaluation often takes place without it being explicitly recognised. However, among those who do write about it, significant phrases recur, such as "initial experiences" and "some problems which need to be refined". These studies tend to be led by the needs of the designers rather than the users – they may be based on real work for methodological reasons, but will tend to take place in such settings as usability labs. It is seldom the case that in practice formative evaluation gets much beyond usability questions – it may *speculate* on the effects on people and work, but is not able to find out much about these, given the time limitations it works under.

More technical formative evaluations are increasingly reported in the CSCW literature, where a system has been developed and then is tested in some quasi-real-world setting; but where the results then feedback into the design process (as distinct from the studies in the next section). Examples of these have included the evaluation of the MAJIC videoconferencing system (Ichikawa et al, 1995), on which the authors conclude a number of shortcomings with the system (such as only one person per site being able to make eye contact with other users), and they say "we have improved the MAJIC system based on the conclusions of the experiments" (p.289). Likewise, Gutwin et al (1996) describe a study of the usability of a set of "widgets" designed to support awareness in a shared workspace, draw several conclusions

about the advantages of, and problems with, their designs, and report that they will "incorporate some of the participants' suggestions into new versions of the widgets … test the widgets with larger groups and on other kinds of tasks … [and] undertake further studies that use stronger measures of the widgets' effectiveness" (p.266) – so that here the ongoing evaluation is also formative of itself.

The ongoing experiences of the POLITeam project, concerned with the electronic linking of German federal ministries during the moving of the national capital, have been documented since its start (Klöckner et al, 1995). While they had assumptions about the nature of the work, they were careful to check it out before engaging in a lengthy design process – and these sometimes proved false, as with the assumption that all procedures in the German civil service are followed rigorously (they are in fact followed with as much flexibility as is necessary elsewhere). Likewise, initial prototypes were tested out, leading to considerable changes. Further experiences were then reported by Prinz and Kolvenbach (1996) and by Mark et al (1997), which in turn describe the resolution of successive issues through the creation and testing of prototypes in such a way that social conventions could arise as well as simply the technical issues being considered.

Another long-term self-aware piece of design/use/evaluation work (cf. Bannon, 1994) was the shared notebook created during the DUCK project. Turner and Turner (1997) report on the different stages they went through in testing out models for the shared notebook, changing it "in response to change requests and the evaluation exercises described". The exercises mentioned were based around an extended pilot test of the system across two test sites (one in London and the other in Glasgow, about 400 miles apart). Users of the system were progressively interviewed: before the pilot, one month after its introduction, three and then six months later. As their views progressively changed as they became more experienced with the system, the system was tailored by the project to suit their needs. In this way formative evaluation becomes (as indeed is explicitly stated by the POLITeam project) related to participatory design.

2.2.4. Type 3: Conceptual

This section refers back to the four kinds of conceptual evaluation discussed when introducing the four types of evaluation: evaluation of commercial research centres; academic research projects; PhD projects; and externally funded research projects. Examples of each will be presented.

The studies in commercial research centres are those that earlier work (Plowman et al. 1995) identified as semi-situated, that is neither completely situated in the 'real world' nor completely artificial and in the laboratory. Thus "real work is still under study, but it is the real work of researchers rather than of typical users" (ibid., p.311). These have been particularly seen within CSCW in centres such as the Xerox Palo Alto Research Centre (PARC), and its European counterpart the Xerox Research Centre Europe (XRCE) but also at Hewlett-Packard and IBM. A good example of these studies are those at Xerox concerning 'active badges' (Harper, 1992) and media spaces (Heath et al, 1995). Methodologically, they are often qualitative, based on users' experiences within the research centres over some time (sometimes with specific experiments added to test particular questions). The general model is one of "looking at ourselves" (Harper, 1992). Many interesting results have arisen from such studies at a theoretical level, such as how people co-operate at work, and how to build systems to support them, rather than directly marketable products. They also, by building up a corpus of work within the same organisational environment, have had a chance to test out hypotheses about the nature of collaborative working in ways that are not generally possible in the everyday work setting.

The academic research project, conducted for its own sake, will use evaluation as a tool to improve its systems. Many of the instances of this kind in the literature (Wan and Johnson, 1994) do just this, conducting experiments to find out how their prototype system is used and what is learned from its use – the subjects being either one's colleagues (similar to the research centres) or one's students (similar to the classic psychology experiments). The results are often of theoretical interest as well as improving the system – Wan and Johnson comment that "lessons learned through the design and evaluation of [the system] provide new insights into both collaborative learning systems and collaborative learning theories". Nor need the studies be prosaic and dull – Benford et al (1997) report the use of a public poetry reading to evaluate their shared virtual reality system.

Evaluation of a computer system produced as part of a PhD project is often felt by external examiners to be lacking. This seems to be for one of two reasons: the system produced may not have been properly studied (although this would have been appropriate); or it may have not been clear how the system could be evaluated. The former case requires more work, and a greater awareness of the importance of evaluation at many stages of a project; an example of good evaluation of this kind is Twidale et al. (1994). This was concerned with a toolkit for building co-operative systems, the evaluation of which is a particular difficult issue; this study solved the problem by building a sample system and evaluating that. This approach is a good solution if there is one main kind of system covered. For others, an evaluation of the toolkit according to theory and objectives seems a better option.

Evaluation for external funding bodies usually appears in the form of 'deliverables': weighty documents that show what the project has done, what papers and computer systems it has produced, and how it has met the goals that were specified at the start of the project. This last point is important: these documents are essentially rationalisations, proofs of the worthiness of the work conducted. Thus, they fit in neatly with the 'goal-based' model of evaluation. Another important feature is the evaluator's report – a good example is the evaluation conducted of the British Alvey programme of IT projects (Guy et al. 1991).

The subject of project has been considered at some length as part of the European learning technologies programme DELTA, which has a specific evaluation component (Cullen et al. 1993). The question here was "in what ways are participants in DELTA learning from the programme?" (Cullen et al, 1993:118). They suggest that learning technology innovation can be considered in four ways: technological innovation (i.e. the building of new computer systems and models), economic innovation (new activities of selling products and services by educational institutions), educational innovation (changes in the learning capabilities and methods provided by the technology) and institutional innovation (changes in the interactions between and within institutions). Considering all four kinds of innovation is to take a systemic view of the programme, i.e. to see "external environmental characteristics as represented within system boundaries" (ibid., p.122). The complexities of this perspective has led them to a 'contingency' model of evaluation, not prescribing methods but rather providing tools to allow evaluators to "tailor their evaluations to the specific nature and needs of their innovation" and to provide "guidelines for methodological choices appropriate to their particular innovation circumstances" (ibid., p.124).

2.2.5. Type 4: Buying

The occasional academic study describes the buying process in action, but by its nature this kind of evaluation is seldom documented, and if so mainly the popular computing press (weekly magazines such as *Computer Weekly* give frequent updates on the latest purchases of IT systems and services by large companies, and occasionally a view of why they made the purchase, but this seldom goes beyond the press-release level).

Fanning and Raphael (1986) describe experiences in the mid-1980s at Hewlett-Packard with computer teleconferencing, and give a good description of the process by which they selected one system and the criteria they applied to come to that decision. They list several systems with the reasons why they might have been appropriate, and why not. Interestingly, the selection was made by a team of seven people, "ranged from technician to lab manager", so perhaps the suggestion above that this is a task purely for the IS manager is an unfair one.

Indeed, the composition of the purchasing team is much of the concern of Green et al. (1991), who discuss the purchase of a library automation system by a team coming from all levels of staff. They explicitly talk about "systems evaluation techniques" – a combination of demonstrations of available systems by their suppliers, and visits by the team to other libraries where the systems were used (mainly the former). The process of evaluation for purchase "did open up systems selection and development issues to basic-grade women library staff" (p.41), without the usual intermediaries, and leading to a much stronger focus on staff and borrower needs than technical requirements (than might have been the case had the purchasing decisions been by managers). The issues of gender and empowerment that they discuss seem to have had a direct impact on the quality of the decisions made.

2.3. Lessons from elsewhere

The material above discusses some of the methods that have been used in CSCW evaluation, and some of the studies that have used them. This thesis also makes use of a body of material from outside CSCW, principally from management studies and social science. Four fields have been especially useful: systems thinking, stakeholder analysis, organisational learning and evaluation research. Systems thinking was considered above in section 1.3; this section considers the other areas.

2.3.1. Stakeholders

The concept of a *stakeholder* is now an important one in a variety of areas, and is one of the main features of this thesis. Broadly, stakeholders are any individuals or groups who can affect an organisation's work or systems, or are affected by them. This section discusses the background to the term in organisation theory, and its use in a number of other fields. The use of the stakeholder concept within evaluation research is discussed in the section on that discipline, at 2.3.3.1 below; its specific use within SESL is discussed in the section "SESL as a methodology", at 4.1.2 below.

2.3.1.1. Organisational Theory

The considerable literature on stakeholder ideas in organisation theory is extremely well summarised in the paper by Donaldson and Preston (1995). The term was coined in 1963 by Robert Stewart, who defined stakeholders as "those groups without whose support the organisation would cease to exist" (quoted in Freeman & Reed, 1983:89). R. Edward Freeman has been the main proponent of the idea more recently, taking as a wider definition "any identifiable group or individual who can affect ... or is affected by the achievement of an organisation's objectives" (ibid., p.91): that is, to include employees, customers, shareholders, managers, unions, governments at various levels, the local community, pressure groups and even competitors. An interesting question is whether one should include non-human stakeholders. The environment might well be said to be a stakeholder in various heavy industries (and indeed better policies of these firms might result were it taken to be so). Vidgen and McMaster (1995) also discuss how artefacts can be considered stakeholders in a system, certainly in terms of their effects.

The stakeholder theory in organisational theory is now quite mature, to the extent that Evan and Freeman (1993) have presented a set of principles, based on Kant's principle that people have a right to be treated as ends in themselves rather than as means to an end, which they say underlies the stakeholder concept. They further argue that the so-called fiduciary relationship that managers have towards shareholders in British and American law – by which managers are said to hold a company in trust for its shareholders, and so have as a basic responsibility the maximisation of their profit – needs to be re-formulated so that there is a fiduciary relationship towards all stakeholders, that managers must act in all stakeholders' interests as their agent.

Donaldson and Preston (1995) justify this concept in terms of property law. It is often suggested, they argue, that a company can be taken to be the property of its shareholders, with the managers as their agents, and so theirs to do with as they please. This is to ignore, they say, the fact the all property ownership is necessarily limited, as part of the running of society. For example, landowners cannot build anything they like on their land, but must seek approval from local government – a limitation on their rights of ownership, but one that society needs to claim to ensure others' rights are also met. From this concept of limited rights, and the current theory of property rights that they term 'pluralistic', Donaldson and Preston conclude there is a moral necessity to take the needs of all stakeholders in a company into account.

At a closer look, the concept of ownership is not all that clear-cut. The workforce can be said to own the labour and sell it, or rent it out, to the company. Professionals can be said to own their skills and qualifications. These things can still be seen as commodities, although they are not tangible objects; but taking this view rigorously makes an analysis very complicated. Does management, for example, own the day-to-day running of the company? Do the neighbours of a chemical plant own the air that is filled with smells and noises?

2.3.1.2. Systems Theory

The second major area where the stakeholder theory has been used considerably is systems theory. The concept of an open system necessarily involves interactions between all the elements in that system, in a complex and interwoven way. This means that to ignore any of the human elements in the system is to lessen one's understanding of the system. In this way, stakeholder theory can be seen from a systems perspective to have an instrumental usage as much as a normative one.

An early use of the stakeholder theory is by Eric Rhenman (1965), who describes stakeholders as "the individuals or groups dependent on the company for the realisation of their personal goals and on whom the company is dependent for its existence" (p.25). His work is interesting because it is an early work in the Swedish industrial democracy movement, but also because he discusses how to resolve the conflicting interests of different stakeholders: his basic model is that these interests are not dissolved or subsumed but rather that co-operation between stakeholders is sought despite them.

The stakeholder concept was also picked up early on by Russell Ackoff (1974), one of the original 'soft systems' theorists; it has been used extensively by Robert Mason and Ian Mitroff (1981). They mainly use stakeholders to identify assumptions underlying corporate behaviour, which might clash between different stakeholders. They have also presented several ways of identifying who are the stakeholders in an organisation: those who have expressed an interest in it; those in formal positions in it; those who others feel are stakeholders; those who participate in the organisation (e.g. in meetings); those who shape opinions relating to the organisation; the standard demographic grouping such as age or race;

and those who fulfil a series of important relationships with the focal organisation in the study. They also talk of resolution of conflicting interests, using a series of dialectical processes whereby the prevailing views within a group are compared with their antitheses, to show how important particular views are to the group. Once a particular stakeholder group has identified its main views, these are then examined in a further 'dialectical debate' between the groups.

Related concepts to the stakeholder theory have arisen in the works of two other systems theorists: Peter Checkland and Peter Senge. Checkland (Checkland & Scholes, 1990) makes much use of a representation called the rich picture, which shows a system under study and the relationships within it, the result being very similar to the stakeholder map presented later in this paper.

Likewise, Senge, when calling for systems thinking (1990) in order to enable organisational learning, lists amongst other characteristics:

- shared responsibility within a system;
- multiple causes and effects;
- not necessarily "equal leverage in changing the system" (p.78);
- no scapegoating.

These are all important components of the stakeholder idea. Despite never mentioning the term 'stakeholder', Senge nevertheless includes vital features of a stakeholder approach in his theory. For example, when advocating dialogue and, as a precondition, "colleagueship", he warns: "colleagueship does not mean that you need to agree or share the same views. On the contrary, the real power of seeing each other as colleagues comes into play when there are differences of view. ... Choosing to view 'adversaries' as 'colleagues with different views' has the greatest benefits" (ibid., p245). Colleagueship, in turn, requires that the parties involved are ready to forsake the power of rank, hierarchy, seniority, or majority, that they need not have fear in expressing contrary views, i.e. it requires an atmosphere of trust where it is relatively safe to take risks and be open.

2.3.1.3. Information Systems

In moving away from the social sciences towards computer science, a first port of call must be information systems (IS) – which stands between social science, particularly organisation theory, and computer science. IS evaluation has been the main place of use of stakeholders in this literature.

A particularly full use of stakeholder theory is by Vidgen & McMaster (1995). Their account draws mainly on the systems theory work shown above (particularly that of Mitroff and Checkland). They show who are the principle stakeholders in their study of the introduction of a computerised entry system to a car-park – both human groups and technological artefacts – and how their interests differ according to perspectives they label Rational, Organisational and Individual. As they crucially point out, one feature of this kind of analysis is that some stakeholders are highlighted as having positive outcomes from the system and some negative.

A further important aspect of their discussion (apart from their interesting use of Latour's work) is their mention of the different knowledge interests – technical, practical and emancipatory – identified by Habermas (1971). As they point out, these relate usefully to conflicting interests of different stakeholders: in particular the conflict between the technological, rational focus of managers and designers; and the practical, everyday focus of day-to-day work.

Walsham (1993) has written extensively on IS evaluation, and uses the stakeholder concept as an essential feature of understanding the social context of information systems. Two qualifications he makes are worth repeating here: that it is not possible to incorporate stakeholder assessments in a complete way; and that it should not be held that all stakeholders views are equally valid (p.183). Thus he is brought to the question of the resolution of stakeholder conflicts, on which he offers the view that it is a problem and involves the view of the evaluator as a moral agent (cf. Ramage, 1996).

Jurison (1994) provides a useful "stakeholder-based value model" for IS evaluation, firmly rooted in an economic perspective, but based on non-financial benefits and costs as well as financial ones. He also discusses how one might measure these benefits.

2.3.1.4. Requirements Engineering

The study of the evocation of requirements for software systems also makes use of the stakeholder concept to quite an extent, in two different ways – directly in the more sociological literature, and under the title of "viewpoints" in the software engineering-based literature.

Eason (1988) wrote extensively of the importance of stakeholders in requirements capture and design, stating as a major proposition that "the design of effective socio-technical systems will depend upon the participation of all relevant 'stakeholders' in the design process" (p.46). He emphasises that he is not just talking about users here, but rather emphasising the many different groups with a stake, including designers.

The ORDIT methodology (Dobson et al, 1994) uses stakeholder analysis in two ways: at the start of a requirements process, the identification of stakeholders is crucial; and at the end of the process, to prioritise different stakeholders' needs, discuss the outcomes of the exercise with the different groups, and ensure the solution options being presented meet the needs. Similarly, Macaulay (1994) writes about "co-operative requirements capture", bringing together the different stakeholder groups to generate requirements co-operatively.

In software engineering per se, the concept of the viewpoint has arisen recently (Kotonya, 1994; Finkelstein & Sommerville, 1996). This is taken to be a combination of a participant in a software system together with their perspective on the system. Thus it might be held to be a combination of a stakeholder and their *Weltanschauung* (Checkland and Scholes, 1990). Because the viewpoint concept derives from software engineering, much of the usage of it has been of the 'hard' kind – rather stereotyped 'lines and boxes' descriptions. Researchers on the REAIMS project, however, have produced work on a softer model of viewpoints (Sommerville et al, 1998) that are more akin both to the stakeholder theory described here and to the *Weltanschauungen* of Checkland.

2.3.2. Organisational Learning

The phrase 'organisational learning' will be used considerably in later parts of this thesis, and therefore needs some unpacking. After all, it might seem plain that organisations, whether viewed as abstract entities or collections of people, are unable in themselves to learn, even if their members can learn. Yet the term is now well respected in academia – how is this so?

In the first place, the phrase is used metaphorically. Morgan (1986) presents eight metaphors used by organisations to understand themselves and other organisations. One of these is the image of an organisation as a brain, which has arisen from an application of the information-processing model of cognitive psychologists to an understanding of organisational life, and from the separate work of the interdisciplinary approach of cybernetics. A slightly more literal use might be attributed to the influence of the behaviourist school of psychology on these

approaches: as recounted in Wann (1964), that school tended to apply the term 'learning' fairly liberally to any process of adaptation in organisms or in abstract entities.

The phrase 'organisational learning' itself seems to have gained particular popularity through the work of Argyris and Schön (1996). Especially important has been their distinction between single-loop learning, where a set of actions are judged against pre-determined criteria as to whether they could be better; and double-loop learning where those criteria are themselves judged for effectiveness. Through this process, the organisation can be said to *learn how to learn*. This concept has its origins in Bateson's (1972) theory of levels of learning, and particularly of *deutero-learning* (Learning II), which he defines as "a corrective change in the set of alternatives from which choice is made" (p.293). Schein (1996) makes the distinction between maintenance and growth on the one hand and transformation on the other.

If organisations can be said to learn, then this surely applies to all organisations to some degree, although whether they tend to be good at rapid change within a limited set of assumptions or are able also to make changes in those assumptions, will vary considerably. However, the question was raised during the 1980s by various groups of academics/consultants: what would a company look like which focused on learning as its main way to maintain competitive advantage? Would it be clearly more successful than its rivals? How should it constitute itself? The parallel terms *learning organisation* and *learning company* arose for such companies.

One answer is given by Senge (1990), and by the subsequent work of the MIT Centre for Organisational Learning that he leads. His vision for a learning organisation is broad and utopian: "where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning how to learn together" (p.3). To achieve this, he argues, organisations must be experts in five 'disciplines': personal mastery, mental models, shared vision, team learning and systems thinking. The last of these is the key to the learning organisation, he argues – by gaining a picture of the whole system, learning can then occur more readily, and blockages to learning be overcome.

A complementary perspective is given by Pedler et al (1991), who talk of the learning company as one that "facilitates the learning of all its members and continuously transforms itself" (p.1). This dual view, of both organisational transformation – which they take to involve the day-to-day activities of the company as well as just the procedures and structures which enable that work – and individual growth, is also useful. It is reflected in the eleven characteristics they use (following validation in a series of research and consultancy projects) to define a learning company: the learning approach to strategy; participative policy making; 'informating' (cf. Zuboff, 1988); formative accounting and control; internal exchange; reward flexibility; enabling structures; boundary workers as environmental scanners; inter-company learning; learning climate; self-development opportunities for all.

Finally, various researchers have questioned the traditional model of learning – a form of knowledge transfer from an expert full of information to a novice lacking that information. They present instead a model that argues that "what is learned is profoundly connected to the conditions in which it is learned" (Brown and Duguid, 1991); that learning involves becoming a member of a "community of practice", and of coming to behave in ways appropriate to that community, in a form of apprenticeship. Brown and Duguid again write that "the central issue is becoming a practitioner not learning about practice". Further, as Lave (1993:8) argues, notions that might be assumed to be unquestionable categories, such as "bodies of knowledge", "learners", and "cultural transmission", must instead be understood in terms of their social context, produced in a particular setting and by particular people.

This approach to learning is paid a certain amount of lip service by organisational learning researchers and consultants, among whom the term "communities of practice", in particular, has gained a certain vogue. However, it has had little real effect, with most learning assumed to be essentially knowledge-based. The exception to this has been the work influenced by Nonaka and Takeuchi's (1995), who make a distinction between tacit knowledge and explicit knowledge – the latter being conscious, cognitive knowledge, and the former being the kind of practice-based experience that Brown and Duguid stress. There is a pragmatic difficulty here, of course – consultants want to be able to help companies change, and it is easier to do this if one is working with things that can be explicitly conceptualised. Nonetheless, a stress on unformalisable practice, and the way that that changes, is an important part of a full understanding of organisational learning. In this thesis, the later discussion of learning will be seen to have a lot to do with this practice-based approach.

2.3.3 Evaluation Research

There is a considerable body of literature on the evaluation of educational, social and organisational programmes, under the name of 'programme evaluation', 'evaluation research' or, nowadays, just 'evaluation'. This has arisen because of demand by funding bodies, typically governmental (local, national and international) – in the US, for example, all new educational programmes have been required to be evaluated for twenty years. Such evaluation tends to be of specific programmes rather than abstract theories, methods or tools. As well as a large body of evaluation reports, there is also much literature concerning how to do evaluation, substantially concerned with methods, methodologies, theories, paradigms, frameworks and so on. All forms of academic are present in the evaluation arena: from white-coated scientist performing quantitative experiments in laboratories to radicals helping programme users to free themselves from the 'chains' that bind them.

Thus, the first issue to be considered is that of paradigm. In general terms, evaluators tend to fall either into the camp labelled by Burrell and Morgan (1979) as "functionalist", or into the "interpretive" camp. That is, most evaluators are not in fact bomb-throwing radicals out to break the system but rather 'tinkerers' who want to see if the system is working out the way it ought (whatever that means) and to suggest how it might work better. The division comes in how to determine whether or not it is working: either one uses the methods of laboratory science, with controlled experiments to determine quantitative data on which one performs statistical tests that reveal 'The Truth'; or one uses the methods of ethnography, performing interviews, hanging around and using open-ended questionnaires to try to understand the situation and draw conclusions about it. That this relates closely to the division in CSCW between (broadly) psychologists and ethnographers will be clear – both come from the same theoretical debate.

The definition of "the way the programme ought to work" is also a subject of some debate. One can take either the perspective of those who designed the programme (i.e. their objectives) or are paying for it; or that of those who are working on it, using it or affected in some other way. This split is entirely separate from the above, but it *tends* to be the case that functionalism and objectives-based evaluation go together, while interpretivism and user-based evaluation go together.

These complicated divisions are rather often summed up under the banners of 'quantitative' and 'qualitative', which are usually referred to as methodologies rather than paradigms. That this is a gross over-simplification seems self-evident: it is entirely possible to conduct a qualitative evaluation of programme objectives, or a quantitative evaluation from the users' perspective, or an evaluation that mixes some or all of these things together. Nevertheless, a broad world-view can be seen.

Patton (1980:28) has summarised the difference as follows, relating to the use of questionnaires: "Quantitative measures are succinct, parsimonious, and easily aggregated for analysis; quantitative data are systematic, standardised, and easily presented in a short space. By contrast, the qualitative measures are longer, more detailed, and variable in content; analysis is difficult because the responses are neither systematic nor standardised. Yet the open-ended responses permit one to understand the world as seen by the respondents."

The other distinction that is widely made within evaluation is between formative and summative evaluations. Michael Scriven coined the terms in 1967 – he described them later as: "Formative evaluation is conducted *during* the development or improvement of a programme or product (or person etc.). It is an evaluation that is conducted *for* the in-house staff of the programme and normally remains in-house. ... Summative evaluation of a programme (etc.) is conducted *after* completion and *for* the benefit of some *external* audience or decision maker. ... The [distinction] has been well summed up in a sentence of Bob Stake's: 'When the cook tastes the soup, that's formative; when the guests taste the soup, that's summative.'" (Scriven 1981, p.63 and p.150)

This division is often held not to be very useful. Wadsworth (1991) emphasises that evaluation should be considered a part of the day-to-day working process. She identifies two types of evaluation – "open inquiry" and "audit review" (p.34), which resemble formative/summative but can happen at any time. She holds the cycle of evaluation to be Reflection-Design-Fieldwork-Analysis-Conclusions-Feedback-Planning.

Many writers present a taxonomy or framework of evaluation approaches. Easterby-Smith (1994) identifies five types of evaluation:

- experimental: based on formal scientific models, where a control group is compared to a group where the programme is being followed, and the effects measured in a quantitative, statistical manner;

- illuminative: using ethnographic methods, to determine the effects of the programme upon the participants in it;

- objectives-based: here there are a set of objectives defined at the start of the programme, and the outcomes of the programme are measured against these; this model has been particularly influential in government;

- goal-free: the aim of this is explicitly *not* to look at the programme's objectives, but rather at its effects;

- interventionist: here the evaluator explicitly seeks to become involved in the programme, to ensure the needs of many different groups are met.

2.3.3.1. The nature of evaluation

Some general issues in evaluation are worth considering from the literature, and particularly from the work of the Evaluation, Development and Review Unit (EDRU) at the Tavistock Institute. Their *Guide to Local Evaluation* (Sommerlad, 1992) summarises many such issues. In general (pp. 3-5), she says there are three kinds of issues in evaluation:

- Conceptual ("how the people involved think about the evaluation"): what is the evaluation for; who are the stakeholders; who is interested in the findings; what approach/model/framework will guide the evaluation; what are the main evaluation questions?

- Technical: what methods are appropriate? what will be the unit of analysis? what steps will be taken to ensure quality data? what kind of data will be collected, from whom and using what instruments? what evaluation products for which audiences? what mechanisms/opportunities are there for reflection and review? how should evaluation feed into decision-making?

- Operational: who will be responsible for the evaluation? who will do the evaluation? how involved in planning the evaluation are the people responsible for gathering or providing information? how will the evaluation be managed?

Why evaluation is conducted is an important question. Sommerlad (ibid., pp. 7-9) identifies three main purposes:

- To demonstrate accountability (i.e. to funders – cf. summative evaluation).

- As a means to improving programme implementation (cf. formative).

- Learning – "contributing to the professional self-development of participants through critical self-reflection; with encouraging systematic learning about complex problems and concepts and with learning about the process of managing change" (p.8). Cf. organisational learning etc.

Stern (1991) lists several more reasons why evaluation is performed: to make a case for funds, to prove to decision-makers that promises made have been kept, to show what benefits have been gained from the programme, even as a way to delay decisions until "all the facts are known". A common phenomenon in public policy is the "justificatory' evaluation": that which is conducted as a public relations exercise, to show funding bodies that the programme was worth doing.

Finally, Sommerlad (1992:11-13) lists the following stakeholders:

- those who care about the programme and its effectiveness (e.g. staff)

- those who influence programme decisions & environment (e.g. funders, administrators, management)

- those who are the intended beneficiaries

- users of the evaluation findings (e.g. academics)

Important points on these are that "stakeholders have different and possibly competing views about what is important, what constitutes success and how success might be measured". Also, "evaluation is 'contested terrain' and so the evaluator must address [stakeholders'] different interests, document the plurality of notions of 'success' and negotiate those issues that are points of contention." (p.11)

2.3.3.2. The nature of qualitative methods

As to an extent the history of evaluation over the past twenty years has been one of a paradigm shift (in Kuhn's sense) from the functionalist to the interpretive, with the latter's qualitative methods, it is worth looking briefly at what writers on qualitative (and formative) methods have said.

Patton (1980:22) has written: "Qualitative data consist of *detailed descriptions* of situations, events, people, interactions, and observed behaviours; *direct quotations* from people about their experiences, attitudes, beliefs, and thoughts; and excerpts or entire passages from documents, correspondence, records, and case histories. The detailed descriptions, direct quotations, and case documentation of qualitative measurement are data from the empirical

world. The data are collected as open-ended *without* attempting to fit programme activities or peoples' experiences into predetermined, standardised categories such as the response choices that comprise typical questionnaires or tests."

Patton also stresses the centrality in qualitative evaluation of a holistic, inductive and a naturalistic approach. That is, the researcher "strives to understand the gestalt, the totality, and the unifying nature of particular settings" (p.41), "begin[s] with specific observations and build[s] towards general patterns" (p.41) and "does not attempt to manipulate the research setting" (p.42). This last point refers specifically to situatedness of qualitative methods: they tend to be based in the 'real world' rather than in the laboratory, although as is argued in Plowman et al (1995), the real-world/laboratory axis is quite distinct in workplace studies from the qualitative/quantitative axis, and several excellent qualitative studies have been conducted in research centres.

Stern (1993) describes formative evaluation of a kind that he calls "real-time". This has the characteristics that: it takes place along with the programme; the evaluation is formative and intrusive, in that it presents its findings during the lifetime of the programme; and the evolution and implementation of the programme are of particular interest. This derives from three particular concerns of the Tavistock Institute's EDRU: the use of evaluation to support learning by participants and organisations, the increasing of mutual understanding; participants are actively involved in the planning and implementation of the evaluation; the evaluation is grounded in the wider context beyond the programme – in the organisation and in society. Real-time evaluation works well for long-term programmes with complex dynamics, difficult processes of implementation and uncertain objectives.

Patton (1981) identifies the need for creativity in evaluation, suggesting many participatory techniques for handling the evaluation process. These include: the use of metaphor; flow-charting; matrix thinking (e.g. 2x2s); experiential methods (games, simulation, role-play, demonstrations); use of acronyms and word-play to stimulate thought; story-telling; drawing pictures; the use of humour. He emphasises the importance of making evaluations useful and the many different roles evaluators play ("information broker, data manager, storyteller, artist, humorist or scientist", p.277).

2.3.3.3. Contingency evaluation

Many writers argue that forcing evaluation methods into particular categories, be they quantitative/qualitative, formative/summative or whatever, is self-defeating and leads evaluators to label themselves as one thing or the other to the exclusion of other methods. This resembles the argument that Ross et al (1995) make for multiplicity of methods in CSCW evaluation. Cullen et al (1993) describe the selection of methods according to the situation as "contingency evaluation".

Patton (1981:23) has written, "Like the shift in perspective among modern philosophers from a search for universal laws to a concern with situational ethics, the scientific revolution manifested by the emergence of the new standards of evaluation represents a situational approach to the conduct of social science. … The driving force that gave rise to this new evaluation ideology is that *there is no one best way to conduct an evaluation*. … A successful evaluation (one that is practical, ethical, useful, and accurate) emerges from the special characteristics and conditions of a particular situation – a mixture of people, politics, history, context, resources, constraints, values, needs, interests, and chance." Patton also writes (1980, p.20) of a "paradigm of choices" that "recognises that different methods are appropriate for different situations".

Similarly, Sommerlad (1992) argues that methods need to be chosen according to the particular situation, based on criteria such as stakeholders, resources, expertise, organisational politics, and the nature of the evidence required. "Our view [is] that the choice of method should be shaped by what the evaluation is for, the kind of questions being asked, and who will be making use of the evaluation findings. All too commonly, it would seem evaluation begins inappropriately with methodology ('let's do a survey') before clarifying the other questions. ... *What matters is choosing the right method for the particular situation.*" (ibid., p.31; italics in orig.)

One solution is to select methods *in toto* according to the needs of the situation – the other is to mix various methods. Patton (1980) discusses two forms of methodological mixes: triangulation, where a series of different methods are used in the evaluation of a particular programme, and the combination of parts from different methodologies. Quoting Norman Denzin, he gives four forms of triangulation – "(1) data triangulation – the use of a variety of data sources in a study; (2) investigator triangulation – the use of several different researchers or evaluators; (3) theory triangulation – the use of multiple perspectives to interpret a single set of data; and (4) methodological triangulation – the use of multiple methods to study a single problem or program" (p.108-9). However, as he points out, triangulation can be very expensive, and may even mean "a series of poorly implemented methods rather than one approach well executed" (p.109). He also discusses the mixture of different elements from different methodologies (as his consideration is qualitative v quantitative, specifically the combination of data types, analysis and research design from these approaches).

3.   SESL under development

In this and the following chapter, the focus of the thesis moves to consider the methodology, SESL (Systemic Evaluation for Stakeholder Learning) developed to solve some of the problems mentioned earlier with evaluation methods. The concepts and practice of the methodology are described in the next chapter; in this chapter are described the experiences of doing evaluation of co-operative systems that led towards the present form of the methodology.

Five evaluation studies are described in this chapter. Four were relatively small-scale (around a days' worth of fieldwork, although with much more analysis time added); the fifth took up some weeks of fieldwork, spread out over a year. Fuller details of the studies can be found in the appendices; the stress here will be on what lessons they provided that were useful for the development of SESL.

3.1.   *BigBank Financial Services*

The study from which this section derives was conducted as part of an MSc dissertation at the University of Sussex (Ramage 1994). Accordingly, it is only used as an example here; an extract from that work containing a fuller description of the study can be found in Appendix B.

In July 1994, I spent a day at the Peterborough office of a large financial services company (a subsidiary of a high-street bank), which offers life insurance, investment management and related services.

When I conducted the study, they were two years into an IT-led restructuring programme, Meridian, aimed to cut costs (from 90p cost per £1 of income to 50p cost per £1) by centralising operations, reducing staff numbers and using IT. They had already moved over from 30 regional offices to two functional ones (one for life insurance, the other for investment management), which had in total cut half of all administrators. They also had introduced a workflow system (based on the off-the-shelf Staffware product) that routed jobs requiring work from one person to another, storing information about the owner of the job, its current status, expiry date etc.

My main method was the 'quiet-observer' technique of ethnography. I had no illusion that participants were unaware of my presence (not least because they often referred to me or offered me information) but I tried not to interfere with their work. There were a series of questions I wanted to pursue (the nature of the organisational structure; how well they'd used business process re-engineering (BPR) and total quality management (TQM); how work and documents flowed around the organisation; how artefacts were used) but endeavoured not to confine my interest to these, to avoid imposing my pre-conceptions upon the situation. I took a considerable amount of notes during the day, including several chunks of conversation that I either overheard or initiated.

Key findings concerning the organisational structure and culture included:

- the organisation was extremely hierarchical, both officially and in practice; while there was talk of turning the administrators into professionals, attempts to make this change were being conducted within the context of a "machine bureaucracy" (Mintzberg, 1979);

- there was little interaction between teams, apart from the passing of clerical work from the administrators to support staff, or from generalist staff to specialist staff;

- while the company had made some use of BPR and TQM in their restructuring processes, the principles of these – flat, process-based structures and a company-wide concentration on quality – had only been taken up to a limited extent.

The main findings concerning the way the workflow technology impacted upon their working practices included:

- it acted as a 'process control' device, ensuring that work was done within set deadlines, and informing managers if this didn't occur;

- it acted as a co-ordination mechanism within a work group alongside other tools, notably conversation, 'tacit awareness' (Heath and Luff, 1992) of others' actions, and the use of a whiteboard to record completed work;

- the workflow system was very dependent on the correct working of other aspects of the computer systems – when the link to the printer went down, automatically produced letters that should have been sent were lost, with no complete audit trail made, and staff had to re-enter the letters later in the day;

- automatic reminders (for bill payments etc.) were issued by the system once a day at 1.10pm. Although these were an important event in the workgroup's day, and a great feature of their folklore, there was little visible sign of this happening; the reminders appeared and later someone came to pick them up;

- views on the system differed: staff felt it allowed them to organise their work better (especially given the volume of the work). The extent to which the manager could see what was being done by her staff from time to time – what they had done during the day – was useful for the manager, but very much experienced as a control mechanism, a 'Big Brother' effect, by her staff.

Lessons for SESL

This study showed the following lessons that informed the development of SESL:

- the importance of being aware of links between computer and other technologies;

- the limits to ethnography: it isn't always possible to observe directly what participants report to be the most important things;

- the effect of organisational politics – Big Brother effect getting in the way of change and effective use of the technology

- the effect of organisational structure – how was the organisation and its relation to the technology different because of hierarchy?

3.2. ShrEdit

In March 1994, Susi Ross and I conducted a study of a group using a collaborative writing system, ShrEdit (developed at the University of Michigan), as part of MSc work at Sussex University. Following the MSc (and in the early months of this research), we wrote up the study as the paper Ross et al (1995), which appears as Appendix C. The paper presents a framework for evaluation, PETRA (Participatory Evaluation Through Redesign and Analysis), which has as its key themes the concepts of multiplicity of disciplines and of perspectives found elsewhere in this thesis. In this section are described the study and its

results, the PETRA framework created around it, and the implications it had for the development of SESL.

The study and its results

The study aimed "to investigate some of the processes underlying collaborative writing conducted via various mediums; and to evaluate the efficacy of ShrEdit as a tool for the support of distributed computer based collaborative writing" (Ross et al, 1995). It had three parts, using the same participants in each stage. The participants were other Masters students at Sussex: their collaborative writing assignment was to 'create and edit a definitive/revision answer' to a sample examination question.

In the first part, the participants collaboratively wrote this answer in a face-to-face setting, using paper; we made an audio tape of the session to allow for later analysis, and logged important events as the session progressed. In the second part, which took place a week later, the participants carried out a similar task (although on a different question) but this time in three separate locations, using ShrEdit and a telephone conference call to link them; we again made an audio tape of this call (which we transcribed), sitting one evaluator with each of two participants and video-taping the actions of a third. We sought then to compare these two parts, to see how the change to a distributed, computer-supported, environment made a difference to the effectiveness of the participants' working. Analytically, we particularly used the different forms of mediation (social, cognitive, technical) found in distributed cognition; and the concept of a breakdown. We were also particularly aware of the different ways that 'shared understanding' developed among the participants – how did they know what one another were talking about? How did they perceive and resolve differences in their understanding? This section of the study, which developed into the concept of the evaluator's analysis, was mainly my concern; the latter part was that of Susi Ross. Key findings here included:

- There were three main mechanisms to mediate shared understanding in the face-to-face context: talk, writing and consultation of external artefacts, of which talk was much the most significant. The talk usually progressed by one participant raising an issue, then others agreeing or disagreeing with their view, and these cycles seemed to make clear where their shared understanding failed to operate. While they took extensive notes during the session, they produced little completed text. There was considerable consultation of external artefacts, which given their co-location was generally a shared experience.

- The same mechanisms were used in the ShrEdit session, but to a different degree. Talking was less continuous – and seemed to be chiefly used to repair misunderstandings, writing occurred more and was more directly productive. In the paper, we suggested that: "co-writing via ShrEdit acts … as the primary mediator of shared understanding". The consultation of external artefacts seemed less pronounced, and the authority placed on them was less. More text was produced using this method, although it is not clear that participants always had as clear a common understanding of the issues involved.

In the third part of the study, we used a workshop technique inspired by participatory design ideas of low-tech design with users, which we referred to as 'Playschool' to suggest its experimental, playful nature. Participants were given a series of everyday materials – paper, coloured pens, scissors, glue, sticky notes etc. Over about ninety minutes, they were asked to redesign the ShrEdit system to their liking, our interest (which we made clear) being not so much in the resulting redesign as in the comments about the system elucidated through the session. One evaluator acted as facilitator of the session; the other operated a video camera, which we subsequently used to note key features of the session. Key findings here included:

- The participants' main concerns with the system concerned awareness, communication and ownership – they wanted to know what was doing what and where, and how it was altering their text; they also had a strong preference for communicating with a microphone via the computer system rather than by phone (seen as a secondary system). Various technical solutions to these and more minor problems were proposed, particularly the use of colour, information on the current status of other users, the introduction of an 'alert' facility (to make other users aware of one's presence),

- The ShrEdit system was not intended as a complete, commercial standard, collaborative writing system; but rather as an experimental program for academic use. Thus, a direct criticism of some interface features was rather unfair. However, participants' experience of the system was that some of the design decisions less consistent with other word-processors (e.g. the use of fonts) made the system less usable than it might otherwise be. For example, when one selected a font, the whole document changed font and not just the item selected ("a whole document in Dingbats – ridiculous!" as one participant commented).

The PETRA framework

The PETRA framework is designed around two main concepts: multiplicity of disciplinary frameworks, and the twin perspectives of the evaluators and the users of the system. The issue of multiplicity is raised in conjunction with the wide variety of different issues relevant to CSCW raised by the large number of disciplines which contribute to it (from software engineering to cognitive science), and the impossibility of taking all of those simultaneously into account in a study; we therefore suggest that one must be selective about which disciplines to draw on, depending on the nature of the system under evaluation. The argument for using the twin perspectives that we developed in the paper is the need to both combine users' experiences of the system (which provide subjective and thus questionable, but highly important, data) with the tools available from scientific analysis (using various disciplines) to gain a deeper understanding from an outsider's viewpoint.

The first perspective of importance in the PETRA framework is that of the analyst of the current situation: of the organisational system before the use of the computer technology, or of the system involving existing technology that is to be evaluated for its effects and possible redesign or replacement. This perspective will be based in some kind of theory of the interaction of people with technology, although not necessarily (given the stipulations here about multiplicity) a well-established and well-labelled theory surrounded by disciplinary walls. This concept arose from the first and second stages of the ShrEdit study, which are based around a simple before-and-after experimental model, using the tools of distributed cognition and breakdown analysis to analyse the resulting data.

Secondly, PETRA includes the redesign of the computer system by the users of the system, a form of participatory design. Thus, "users become active participants in the evaluation process, expressing their experiences, criticisms and ideas directly using the design materials, giving both themselves and the evaluators a richer appreciation of their experiences with the system" (Ross et al 1995). This perspective arose out of the 'Playschool' session – the third stage in the study – using redesign to facilitate an awareness of the users' problems with ShrEdit from the second stage.

Lessons for SESL

This was the first CSCW evaluation study I performed, and it was through the work on the PETRA framework and subsequent understandings of its inadequacies that SESL came to be

formed. Some of the lessons learned are therefore positive ones, others more ones that are reactive. Key lessons included:

- The dual perspectives of evaluators and users are an important part of the PETRA framework: our argument is that by looking from more than one angle, one gains a richer understanding of the system. However, it became clear towards the end of the work on PETRA that these two perspectives were inadequate, and that other angles needed to be taken. For example, the concept of 'user' was much too coarse (as well as too focused on the technology): different kinds of user will have different experiences with a system, and other stakeholder groups have other experiences. Thus the importance of the concept of stakeholders, which is discussed extensively in chapter 4, arose through, and in reaction to, the development of the PETRA framework.

- The concept of multiplicity of discipline – the importance of not getting 'hooked' on a particular analytic framework, but selecting appropriate parts of each according to the needs of the study. Later in this thesis, there is much discussion about the importance that evaluation be seen as a systemic process, considering the whole socio-technical system and not just a small part of it. This is an important issue to raise the awareness of those who have a strong attachment to one or another part of a system (such as an interest only in the computer technology). However, it raises the problem of too much detail – if one considered all aspects of the whole system, one would be never stop studying it. Therefore one needs to stop somewhere – but the argument from PETRA is that that stopping point should be governed by the nature of the system rather than by the pre-formed disciplinary background of the evaluator.

- Similarly, the different results that can be obtained from the use of different data-collection methods. The scholarly, measured approach of the analytic part of the study and the more down-to-earth, everyday approach of the Playschool session, gave very different information, almost entirely complementary (although compatible). It is doubtful that the same depth would have been given through one or the other method alone, unless considerably more time and effort had been invested in it than we had available. In particular, different levels of abstraction were involved in the data available from the two methods – the analytic study gave results in terms of fairly abstract concepts, while the Playschool session, although dealing to some extent with some of the same concepts (such as awareness or communication), gave much more concrete results at the level of changes to the user interface.

- The user experience with ShrEdit included substantial dissatisfaction with the extent to which it departed from standard user interface guidelines, such as in the issue of fonts discussed above. The interesting thing about this is that the designers of the system never intended it to be anything other than a research system, to facilitate experiments about the nature of synchronous co-operation. This gives rise to the observation that, whatever the *designers* of a system think it is intended for, the more important thing is the experience of the *users* of the system with it (although the designers, if they are wise, will incorporate that experience into their system).

- The importance of the interrelationships between different communications media. The interrelationship between different *cognitive* devices (talk, writing, artefacts) is stressed by distributed cognition, and formed part of the analytic background to the study; but the way in which the telephone and ShrEdit interrelated in the second part of the study was an interesting result, and proved to be a contributor to the later concept that a co-operative system includes all sorts of media which facilitate co-operation (as well as people,

organisations etc.) and should not just be viewed in terms of the computerised mediation of co-operation.

3.3. Health Centre

In May and June 1995, I conducted a set of interviews with local health service professionals in Lancaster, with the aim of investigating and evaluating the use of computer systems to facilitate co-operation in their work. There were four interviews in all – two with doctors and two with practice managers at health centres. Some of the people interviewed worked at the University health centre, some at a linked practice in town, with the doctors occasionally working between the two.

Two main issues of computerisation arose. One was the development of an electronic patient records system; the other was the development of a system, LINX, to link the local health centres to the regional Family Health Services Authority (FHSA).

At the time I spoke to the practice, they were keeping a partly electronic and partly paper-based system of storing patient records. Full records mostly remained on cards; but a general information file was kept on computer of contact details and main long-term information (and this had been done since 1989): immunisations, allergies, smear tests and long term conditions such as diabetes or asthma.

The main purposes of the electronic records included

- making appointments (allowing the receptionist to check for serious conditions to be noted, for looking up of paper records, and for confirming foreign names unfamiliar to the receptionist);

- providing statistics for the government, to encourage the increasing national emphasis on health promotion and preventative medicine. The system can easily record the extent to which targets for smear tests, child health tests etc. are being carried out. One practice manager described this as: "the government's way of measuring that the patient is getting that care ... what they'll do with those I don't know".

- facilitating the use of repeat prescriptions. The system stores a standard list of drugs which have been prescribed to patients in the past and which their doctor deems appropriate for repeat prescriptions. It also can easily control how many repeat prescriptions are permitted until a further consultation is necessary. One doctor reported that 90% of all prescriptions were held electronically in the first instance.

Some local practices, within the same Family Health Service Authority area, had by then already gone completely computerised. They also expected eventually to move to full electronic records, but there was a big overhead of entering past information from existing cards on to computer, and since the practice had a lot of students, with entering new information at the start of an academic year. They also eventually intended the creation of a computerised appointments system, to allow the receptionist to key in to find out who was free at a particular time rather than looking up a book; and to keep records of appointments that could be manipulated on computer, allowing for government statistics but also predictions of demand.

The views on the patient record system were mixed. It clearly entailed a lot of extra workload – one practice manager estimated that probably a complete person was used to do all the work on the computer. Another administrator felt it to be a bit pointless, as all records were still held on paper as well for the sake of Home Office regulations and also because so many of their patients were students, who would move on (and hence need their records passing on) in

three years. The ability of the system to assist the finding of paper records was definitely seen as a strong benefit, and in general one administrator commented: "at first it was difficult, but now the staff wouldn't do without the computer".

It did seem, though, that the ability to generate many statistics hadn't reduced the workload but rather meant that more things were done, particularly given the changing government requirements mentioned above. Finally, one doctor raised a concern about the introduction of a wider system – at another local health centre, the procedure when letters were received was that a doctor would go through them and picks out relevant bits, which were typed up and the letter then thrown away. This is fine in principle, he said, but how do you know you've got all the relevant bits? What if different information from the letter is later required?

A different issue of computerisation, although linked to the patient record system, was to do with the link-up of data between general practices and the FHSA. This was especially motivated by the changes in government statistical requirements mentioned above. In this area, there was a project called LINX, which would enable that data transfer to occur. This had several stages: to make sure the patient record details mentioned above are held in exactly the same form at the FHSA and the practice, which requires that existing lists be "rationalised" and then that future data be input by the practice and electronically passed to FHSA; to allow practices to claim fees for general and for certain specific consultations – such as for contraception and immunisation – electronically from the FHSA; to link up with the local hospital's information system; to link up all FHSAs in the country.

At the time of the study, the first stage was partly under way. When new patients registered, their details (somewhat more elaborate than previously) were entered by the practice, then emailed to the FHSA, although paper forms still had to be sent on registration as well. Fees for registration interviews were credited the next day rather than (as before) three months later.

Most of the people interviewed were fairly negative about the project. One doctor commented that "most of the benefits are for the FHSA" (and the administrators felt similarly). The implementation seemed to be poor in its then-current form: it was installed in September 1994, just before all the new students arrived, and for the first three months of use (while 1000+ new patients were being registered), the system was very unreliable. The extra information required had caused some resentment, given the extra workload it entailed.

The later stages – fee claim, hospital linkup and the nation-wide system – seemed more useful to them, but there were various concerns. With the fee claims, it seemed that they would still need to be authorised by signatures at the FHSA, which would slow down the process. With the other two stages, timing was a bigger concern – no-one seemed quite clear when they would happen, although the prospect of patients' records being emailed from one FHSA to another as they moved areas was very desirable for them.

One final issue which arose was that there seemed to be a clear cultural divide between the medical and the administrative professionals, which mirrors that between the academics and administrators at "Poppleton University" (section 3.4 below). This was reflected on by members of both 'sides' in interviews, although in each case I only talked to the more senior staff in each hierarchy (doctors and practice managers, rather than nurses and receptionists). Both groups were clear that there *was* a them-and-us issue, but their view of the precedence of the groups was different – the doctor who reflected on the issue felt that doctors were definitely in charge, with administrators secondary, while the administrator felt that the two groups were performing different but complementary roles.

Lessons for SESL

- benefits don't always accrue to the people doing the work (as in the case of the collection of the new statistics) – this point was also made by Rogers (1994) under the title of "distributed knock-on viscosity", but was quite clear in this case;

- the introduction of new technology doesn't always bring less work – sometimes it's just different work, or it enables new work to be done that was previously difficult or impossible but now becomes desirable to external stakeholders;

- the start-up costs of moving a system over from paper-based to electronic can be so substantial that sometimes another impetus is needed to make it worthwhile;

- again, the interrelation of paper and computer technologies can significantly enhance both, as in the case here of the administrators looking up computer records to find the location of paper files;

- *when* new technology is installed can be a further environmental factor, contributing to the stakeholders' experiences with the system – here the fact that the first stage of LINX was installed just before the students arrived made for significantly greater problems for the administrators;

- even if organisational structure and culture means that there are clearly discernible splits between different groups within the organisation, it may still be the case that those splits are perceived in different ways by various people.

3.4. *University Management Information System*

In February 1995, I conducted a short evaluation of the new management information system at a university (referred to here as the University of Poppleton, not its real name). The system provided financial information of all sorts – processing invoices, managing budgets, issuing cheques etc. – throughout the university. It was one product of a joint project between a consortium of UK universities (the "Common MIS Project"), intended to standardise the computer systems being used for accounting, student records, facilities management, and other administrative tasks, so that information would be more comparable between the universities and more useful for the funding bodies.

The system at Poppleton – which is here called Cassandra – was principally intended for online, real-time use (via telnet), although it also had a facility for printed requests. It replaced an older system, which was notorious for its lack of useful reports and its difficulty to use, and didn't have the online capabilities. Both systems are clearly co-operative systems according to the definition here, though, as they facilitate co-operation between various groups within the university.

My research methods in this case mainly involved interviews. I conducted interviews with six people: from the finance department of the university, from other administrative departments, and users of the system. I also attended a meeting of a group set up to consider the user-interface problems with the system – and by extension the Common MIS systems yet to go live – but which also discussed other problematic aspects of the system.

Key views on the development of the system that emerged included:

- Poppleton was the first university to go live with Cassandra. One person involved from early on commented: "it's certainly not been easy…it's been a nightmare in fact". It took huge amount of person-hours to get it running – they tackled it systematically, implementing one piece at a time and checking that it ran correctly, running old and new systems simultaneously for a time. The whole implementation

process cost around £100,000 (including time taken up by workarounds on the system).

- There was a widespread view that the system had been imposed upon the users. One administrator said that it was "imposed on us, lots of staff don't like it because their backs were put up when it was first put in". Another said they had received only half a day training, then their boss left for maternity leave and they hadn't received any since. This was described by another administrator as: "they couldn't tell us how it worked, took ages to figure out shortcuts … it was really the blind leading the blind". The view was summarised by still another person who said: "How can they make decisions like that without consulting people? They're there to provide a service." An accountant felt that Cassandra had been "produced by computer people without thought of what accountants want".

- It seemed from some views that this had a lot to do with the organisational culture at Poppleton University, and especially the increasingly strained relationship between administration and academic departments. Members of the administration departments often reported a secretive culture, "always having meetings behind closed doors", and a rather confrontational one: "management attitude to staff – we get screamed at when things are late, it's not our fault". On the other hand, their links to the user departments seemed tenuous at times and based on a strong academic hierarchy – one member of an academic department said that memos tended to be addressed to Heads of Department rather than the people who need them, and another said that people who need financial regulations don't have copies of them. These aspects of secrecy, confrontation and hierarchy appear to have created serious obstacles to the implementation.

Key views on the way Cassandra was used day-to-day included:

- Considerable changes in working practices were required. For example, a lot more direct input of data by user departments was now being done, which reduced the workload of the central administrators but required the users to learn a lot of new information about cost codes etc. These had caused many problems. One person had (due to lack of information) been coding a particular item wrong for a year, meaning the university couldn't reclaim VAT on those items. In another instance, someone required a code for key cutting. The Finance Department said they had no code for key cutting (although they had one for grass cutting!), but rang the Estates department who did cut keys and had an internal code – which could be modified for the new use. It seemed very laborious, and the user commented that most other people probably just coded key cutting under Miscellaneous.

- Reports caused other problems. One person said they weren't what departments wanted to see. Another described them as "bloody difficult" to produce. The amount of information they produced was variously seen as helpful (for example because each research project appeared on a separate sheet) or giving too much detail. One person said "I know all the information's there…but as far as I'm concerned there's too many figures", and another that "they can't give you the information you need, it's all clogged up with information you don't need".

- The system seemed to be slow and difficult to use on many occasions. When many people were accessing it, it slowed down appreciably, taking 3-4 seconds to move to the next screen. For another user, the good and bad were mixed up: they found inputting slow and laborious, and printing tricky, but looking up information on screen

useful. However, apparently smaller things often ruined this: one person said her assistant was off sick, and so she was unable to get done any of the work her assistant would normally do, as only one person could own a report at once.

Overall, the views were mixed but generally negative:

- Advantages included the ability to generate monthly top-level management reports on a university-wide basis; the chance to display information on screen; the potentially greater control given to user departments.

- More negatively, someone in an academic department said: "We're not accountants…it's probably wonderful for accountants…". This view was commonly held – "why can't we have something like a bank statement?" was the way this was put by another person. Another view was that it wasn't making much of a difference, and that they just "had to live with it".

- The general view was expressed at the users' meeting I went to, when someone asked "why don't they just throw it out?", to which the response (from someone else more closely connected with the system) came: "they've already spent too much money on it".

The chief output produced was a short (2.5 pages) note to the "Common MIS User Interface Requirements Working Group" set up by the university to deal with some of the issues concerning usability which the Cassandra system had raised, both for it and for the further systems (then as yet uninstalled) from the Common MIS Project. The full note can be found as Appendix D. The chief point made was that "the methods currently intended for use for the study of user requirements for the Common MIS systems, both Cassandra and future systems, is not likely to provide an adequate solution". The note argued this from the common finding in CSCW (and requirements analysis) that traditional information gathering methods tend to lead to a lack of appropriateness of the system to user needs, and a lack of user acceptance. It therefore advocated an approach combining workplace studies (using ethnographic methods and unstructured interviews) and user involvement in the design process.

The report had little immediate impact. However, it is helpful to note that later implementations of the Common MIS systems were considerably less problematic at Poppleton, particularly the student record system (an earlier version of which was described by the university underground newsletter as containing "scores of useful templates that locate, collate and make usable the most commonly sought kinds of information … completely replaces student record cards"). This evaluation was one of the factors which contributed to the changes at the university that allowed this.

Lessons for SESL

- The study strongly reinforced the importance of organisational culture – many of the issues raised weren't technical problems but to do with the *way* the system had been introduced.

- As can be seen in the discussion above, much of the feeling about the system was negative. A certain amount of reinforcing feedback seemed to be going on in this – as people talked with their colleagues in other departments about the system, who tended to have negative things to say, their own nascent negative views were supported and strengthened. My set of interviews may have contributed to this process.

- There were radically different views about the system, though, and this was particularly noticeable according to the different positions in the university people held – the finance department member, as one of the 'owners' of the system was a lot

more positive than users from academic departments, although the picture wasn't as clear-cut as just dividing into centre = positive view / periphery = negative view (cf. the academic department member who said they had to just live with it). Nonetheless, the concept of different stakeholders having different views as to the rightness or wrongness of a system seemed to gain some weight here.

- Some learning seemed to occur for the people interviewed through the discussions I had with them, and an awareness of this later fed into the longer consideration of this phenomenon during the RICS study.

3.5. *The development of SESL in the RICS project*

The longest study conducted during the course of this research, and the most influential upon the development of SESL, was a one-year study of the learning processes of a research team. The team was researching learning processes, and specifically organisational learning, within the surveying profession (funded by the appropriate professional body, the Royal Institution of Chartered Surveyors, RICS). It took place at a university, here called Lakeside University[3]. The project was primarily carried out by four people: a full-time researcher (Fiona), and three other staff members acting as advisers to her (Mark – the project leader, Dennis and Mary). Around two months into the course of the project, Fiona realised that although the project team were looking at surveyors' learning processes, they had made little consideration of their own. As she knew of my interest in evaluation, she therefore asked me to come in as an outsider to evaluate their work. I started this process in late November 1995, and conducted my final interviews in early November 1996, shortly after the end of the project.

This evaluation of this project turned out to be highly influential in the development of my ideas on the nature of evaluation and what were the main issues to focus on in the development of the methodology. This section describes the methods used in researching the team, the main results arrived at, and how this impacted upon the development of SESL. Again, a fuller discussion, giving more details of the study and its results, can be found in Appendix E.

3.5.1. Research methods

The best characterisation of my research style during the evaluation is probably one of 'participant ethnographer'.

My main methods for conducting the research were a standard set of ethnographic ones: interviews, attending meetings and generally being around. My role was enhanced somewhat, as well as giving useful additional data, by additional activities I carried out, principally the programming of a database/spreadsheet combination to allow for easy entry of the data from a questionnaire; and providing feedback at various stages, but especially the summary I prepared of an exercise during one of the steering committee meetings and the stakeholder map co-created with the project team. My role was also frequently to be a "human reminder" of their commitment to the process of change and of conducting their research in a more reflective manner.

During each of these activities, just studying them was often enough: by conducting interviews, I acted as a transmitter of ideas from one person to another; and gave each

---

[3] All names have been changed.

individual space to reflect on the workings of the project with someone else familiar with the project.

Reflecting on my continuing presence in the project as it went on, it was clear that my main role consisted of being there as a 'human reminder' of their goal to be a learningful group, and hence of facilitating organisational learning. I also gave them feedback on paper (at least ten documents over the period of the study). Some of the questions I asked provided them with space to reflect on their experiences – in the stakeholder mapping exercise, for example, this certainly seemed to be the case. I was around quite a bit, especially at formal meetings, and I was not really playing the unobtrusive ethnographer: at Steering Committee meetings, for example, I did not wear a suit like everyone else but rather the standard graduate student's uniform – casual trousers and a shirt. My actions, too, were different: I sat in the corner writing, rather than talking about the project like everyone else. So just by being there, I had an effect, a further encouragement not to get stuck in the old ways of thinking about research projects but rather to move forward in new ways.

3.5.2. Key results from the study

Four questions came out through the study as most important – they are summarised here and discussed in more detail in Appendix E: how was the project 'learningful'? How has the research been better than the traditional model? What has been the effect of the topic upon the research style? How has technology facilitated their learning?

On the first two questions, the phrase 'learningful manner' refers to both the project's willingness to innovate and its ability to learn from its mistakes. This was shown in several ways:

- *Conduct of Steering Committee meetings*. The project was overseen by a committee of representatives of the various stakeholder groups – especially from the RICS and surveying firms, but also people with knowledge of organisational learning – which met five times over the course of the project. The meetings were fairly innovative. Although they used a standard committee format, several changes to this format were attempted to give space for more reflection. Most importantly, from the third meeting onwards, the data so far gathered in the project (variously, questionnaire responses, graphs, statistical analyses of quantitative data etc.) was put on the walls of the meeting room for comment by the Steering Committee members, allowing much more time for reflection than if this information had been handed out at the meeting. Next, the minute-taking was conducted in a way that allowed space for reflection on, and ownership of, the items under discussion before they were concluded (at the end of each item, a scribe drafted a minute, read it out for comment and changes, which often led to the re-opening of a discussion). In one of the meetings, an exercise was conducted, at the project team's instigation to determine how the project was progressing so far, with eight sheets of flipchart paper put up around the room labelled with titles like "What I give" and "What this Steering Committee lacks". This seemed a particularly good form of self-evaluation, and I wrote a summary of the sheets, most of which can be found within Appendix E.

- *Self-awareness*. How aware was the team of its own collective learning, and its own need for learning? It was about two months into the project that Fiona realised there was an anomaly between its focus on organisational learning and the lack of focus of the team on its own learning processes. There were two immediate results of this: the "go-round" sessions held at each team meeting and my involvement as an evaluator of their learning. The former were so called because they consisted of ten minutes or so at the start of each meeting going around the team with each member saying how they had been feeling about

the project recently. Of the team meetings I attended, these were a wild mixture of different styles. The extent to which they took place waxed and waned, but when they did happen, the go-rounds seemed to be evidence of an attempt by the team to act in a way entirely appropriate to the learning organisation: to integrate their work and emotional lives. They also gained an awareness of their own interactions through my role as an external evaluator: being there to remind them of their aim to be reflective, occasionally asking questions and generally furthering their organisational learning.

- *Open approach to team interactions: more aware of conflicts?* If one of the team's basic aims was to be open in their interactions, how does this mean they handled conflicts? The view of how much conflict was around, and how it was handled, seemed to change for different individuals. Two examples of this conflict, and the ways they were successfully resolved, can be found in section E.3.3. However, this discussion of the surfacing of some conflicts must not be felt to mean that *all* conflicts were resolved so openly. Mary commented in an interview after the end of the project that the lack of close day-to-day working by the whole project team meant that there were potential areas of conflict which didn't really arise – particular nascent conflicts identified by the project members were on gender, seniority and the divide between 'academics' and the 'practitioners'.

- *Presentation style, talks and reports.* They had a problematic dichotomy between wanting their reports to be read by sceptical surveyors, and yet being radical enough to make a difference. Their solution was an interesting one. The bulk of the report forms a reasonably standard academic project report in the management sciences, yet it is framed by an extensive set of appendices which contain graphs in abundance, codified data and statistical analyses and a set of mini-dialogues between an authorial voice and three surveyors. The graphs seemed, at one Steering Committee meeting I attended, to convince the surveyors present that they were doing something right, producing real data. The "Surveyors' Tale", on the other hand, is aimed to change their assumptions about the world through its style as well as its content – in Bateson's (1972:1) term, it is a metalogue. In other forms of presentation style, the project team attempted to challenge assumptions through multiple kinds of message, particularly using the wall-displays mentioned above, giving quantitative data but also qualitative responses to questionnaires, discussions about the project team such as the stakeholder map, and so on. It is interesting to speculate that the interrelatedness of the multiple media the team used for their communications processes may have had some effect on this use of multiple messages, in a systemic rather than directly causal fashion.

- *Openness to change in research style.* The team seemed on occasions to be more learningful than a more standard research style. Particular examples included changing their mind some way in to the project about their definition of a successful surveying firm (on which the whole project rested); and restructuring their plan extensively in the second half of the project, to deal with unexpected delays in the early stages of the data-gathering process. Whether these changes were planned or merely expedient is open to question; but they at least show that constant change was part of the team's experience.

- *Development for all.* The self-development needs of the project team members were all different, of course, and so they were met in different ways. For the project leader, who had been little involved in research before, seeing how a research project worked was a learning experience. For the full-time research assistant, going to conferences and meetings on behalf of the project was beneficial to her own development – but also significantly benefited the project, and brought back new ideas, skills and knowledge to it.

For the other two team members, their main benefits were measured through more standard academic criteria: papers published and the chance to work with useful data.

Moving now to the effect of the research topic upon the research, it is tempting to conclude a causal link. They were talking about organisational learning – did this make them focus on their own learning as well? An examination of the data and discussions with the research team leads to the conclusion that such a link would be inappropriately positivist in its attempt to establish causality. Instead, it seems likely that a systemic relationship took place between the membership of the team, their backgrounds, and the fact that they were looking into organisational learning. This is expressed below in figure 2 and then explained (as well as in more detail in section E.4).
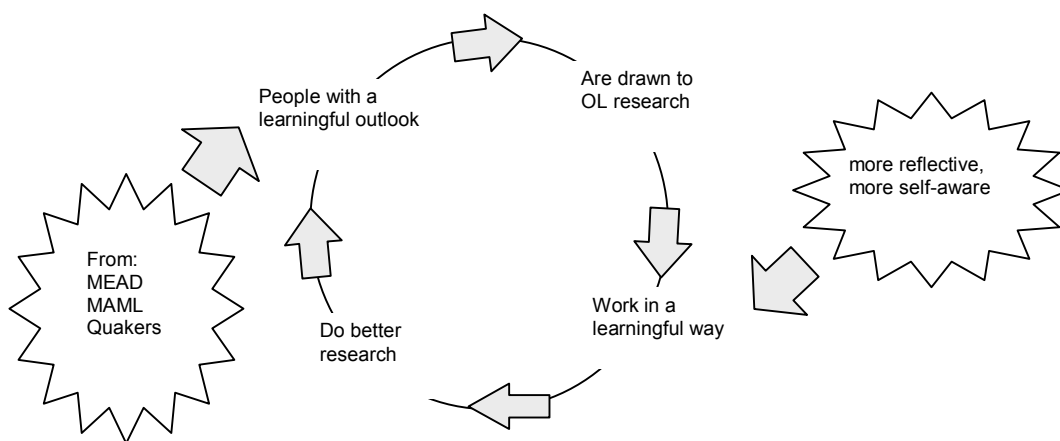
Figure 2: The learning loop of the RICS project

The first, and perhaps the key, point on the loop here is the item "people with a learningful outlook" at the top left. It seems that the team brought together were exceptionally well focused on learning as a good thing in itself, in their lives and working processes as well as a research topic. Their history comes in particular via the Department of Management Learning, Lancaster University. Three of the team were linked with this – one took the MA in Management Learning, the other two were on the board of the journal *Management Learning* (formerly *Management Education and Development)*, run from that department; and Mary is

51

also on the board of the journal. Why is this significant? The department has a long history of focus on learning in all areas of life. Two of the team are also Quakers, a group that also emphasises learning as a key part of life and as a lifelong process; and also refuses to separate different areas of life from each other. Another member of the team had been applying concepts of organisational learning practically in the department he runs for some time before the project began. It seems to be the case that the fact of being learners by nature has drawn each of them to organisational learning research.

That they worked in a learningful way has been discussed above – in particular, they were more reflective and more receptive to self-study. Was the research better as a result of this? The research was certainly strong, and has been subsequently well regarded by those knowledgeable in the organisational learning field. As to whether it was better through the effects of organisational learning, that is also hard to show. The link is a complex one, but thinking more about what they were doing and why they were doing it certainly seems to have been of benefit.

A further question raised throughout the study was the extent to which technology helped them to build a learning organisation. Their main communications technologies were email (Microsoft Mail), plus the various 'old technologies': fax, phone, and post, which were often integrated with each other. There were three main uses: community building, by keeping the group alive as a group between meetings and discussing various subjects; passing-around of documents; and dissemination of interesting information.

Since the team was based on three sites (although in the same city), it was necessary to send around documents at various times. They could be passed around as attachments more quickly than via the quirky internal mail. This was hampered by the lack of a single email system at the university, and by the frequent crashes of the internal network, but was useful both for dissemination of finished documents and to facilitate editing of draft documents. One occasion where it did become more useful was in the selection of which companies to which to send a questionnaire: email (via Microsoft Mail attachments) was used to allow the team members to select companies and see each others' choices. This raised one technical issue, as two of the team were unable to receive such attachments, and so their input to the selection process had to be by faxes to the research assistant…

The use of email for community maintenance was quite important to make appointments, set dates and times for meetings, and get messages through when people weren't at their desks. It also kept people in contact between the three sites of the project. However, one person felt that the difference this made was not significant. As time went on over the course of the project, there is little doubt that they became more sophisticated in their use of electronic communication, sending messages more productively using it – this was the first experience of co-operative working using such media, and so there was a considerable learning curve. There was also a spin-off to other places – Mark, who was a departmental head, fed back his email experiences for use within his department.

As for the use for disseminating information (sending around messages from mailing lists), this was felt to be of mixed benefit. One member commented that it had "been useful to help develop the rest of the team's understanding, though without reading it all" (interview, 4/11/96); another that it was "fun and interesting" (interview, 4/11/96), but on the whole hadn't read a large proportion of them, not having the time.

There were also more negative aspects of email mentioned by team members. For four months, no messages got through between one site (where one of the team worked) and

another (where two more worked); nobody realised what was happening until it was over, and just assumed no-one was replying to their messages. For another person, speed was an issue – she couldn't always comment immediately on documents; but if someone else did, it sometimes happened that a revised version went out before she'd had a chance to comment on the first version, and this led to a certain amount of resentment. Dennis, on the other hand, could, so that Fiona then sent out a revised copy before Mary could comment on the first. She said (interview, 4/11/96): "Well, why should I bother? Where am I in this?", and that she couldn't cope with the speed technology puts on her, so was less inclined to do the work. Finally, it was an issue for some people that it didn't really save any paper, as still preferred to print out most messages for later reading.

In summary, Fiona remarked (interview, 6/5/97) that email use involved a mix of what one might say on an answer-phone and what one might send by post:

> You can see the role of email if you think about how life in the project team would have been if we hadn't had it. We would have had to send it all by mail. … The collaborative writing process – and there was a lot of it … e.g. the questionnaire changes, comments & annotations on the minutes – all that would only have been possible with days & days & days in between; and then it would have to be collated on paper. … It would have been a lot slower, and wouldn't have been so collaborative – which might have been an advantage at times! – but with this team, email was quite an enabling device. … If I look at all those things we did with this email stuff, it would have been far more time-consuming and a hassle to work together as we did. Certainly, the chunky text bits, they would have been just less collaborative full stop.

3.5.3. Lessons for SESL

The study proved very important to the development of SESL, and probably had the biggest single impact upon the method and the ideas behind it (along with the early work arising from the PETRA method described above). Two of the main areas of SESL became increasingly clear during the work – the importance of stakeholders and the view of evaluation as learning.

The development of the stakeholder perspective through this case study was a curious one. Very early on, I became interested in the question of roles in the project – who was doing what? How was it defined? How did they manage to keep track? It seemed clear that I needed to gain a picture of their understanding of themselves as a group so that I could see what it was they were doing together. My initial interviews, conducted in December 1995, had this as a key question – a typical version (from interview with Fiona, 12/12/95) was: "the starting point I'm interested in is how you see the different roles that the different team members are playing at the moment, and how those relate to their previous – and how those have changed over time". Her response to that was relevant to later work:

> Well I have a sort of mental image, a mental diagram, of where we are rather than what we do, and that has Mark, Dennis and me in the middle, in the centre. … And it has, it's a bit like concentric circles, and some people are more at the core as it were, the centre of the circles, and others are on the outer circles. And that's why I said it's still a Gang of Four, with Mary as a half member, and Terry K as a half member. Which made Dennis pull a face, but I still see them at the centre, or close to the centre.

A month later, at the project team meeting, I raised the question in a different way. Taking the list of key people that had arisen from those interviews, I wrote their names on little pieces of paper – I initially had also intended to add little phrases describing how they saw their roles, but did not have clear enough information to do that. The team then moved around the pieces of paper on a large flipchart sheet (adding more as they went along), positioning them to reflect the relationships of power and influence between the different individuals and groups

connected with the project. The result was the stakeholder map that appears in the next chapter.

So the use of the stakeholder mapping exercise made the question particularly important for me; for the project as a whole it was an important concept too, as they needed to be constantly aware of who wanted to get things out of the project, and to what extent those collided. The concept had also come to the fore through reflecting on the experiences in writing the PETRA paper, and realising that more than just two different perspectives upon an evaluation were important. As it happened, the concept of stakeholders was around in many different places those first few months of 1996 – the phrase became something of a buzzword in the centre-left of British politics, following Will Hutton's (1995) book and a speech by Tony Blair; but it was also at the heart of the RSA (1995) *Tomorrow's Company* report and papers from the management sciences like that of Donaldson and Preston (1995). In all these different ways, then, the concept of stakeholders increased in importance during the RICS study.

The next important piece of learning from the RICS study that fed into SESL was the importance of learning itself. It was clear that the different kinds of learning arising during the project changed according to time, according to the various stakeholders, and according to the activities conducted with them. Thus in working through the stakeholder mapping exercise with the project team, they learned about their nature as a group and the environment in which they were working, and they had a chance to reflect on the way they interacted with other groups. By contrast, the interviews conducted were more individually reflective, giving people a chance on a personal basis to consider how far they had gone with the project. In a different way again, the work I did in setting up the Excel spreadsheet and Access database for entering the questionnaire data helped Fiona to learn about the issues involved with those programs. Thus at different times my involvement as an evaluator worked in different ways, leading to different results.

Learning was also important for the more direct development of SESL. An important concern that arose was the risks involved with seeing evaluation as a process of making judgements in a vacuum, of forming an expert opinion. The concept of organisational learning fits nicely as an alternative paradigm for the methodology. It is clear that this particularly arose from my involvement with the project, and consequent increased awareness of the considerable literature surrounding organisational learning.

A further issue that was particularly important during the RICS study was the value of a systemic perspective: the notion that the key issues aren't necessarily the technical ones. As detailed above and in the appendix, the use of computer technology for co-operation in the project was fairly limited, although important; but many interesting issues concerning the co-operation going on during the project and the extent to which they were learning arose. I followed these rather than the technical questions, which changed significantly the nature of the study, but seemed important to do. Had I just focused on the use of the email technology, I would have learned considerably less about the way they worked; and in the process probably learned less about the email as well, as its use was so bound up with the other aspects of the study.

Finally, it was important that my role in connection with the RICS project was not just one of passive observer. I played a part in making an active contribution to the project. Some of the activities I performed during the project are discussed in section 3.5.1 above, and it should be clear from that list that my contribution was as an almost-team-member as much as an external evaluator. This shows the importance of 'getting one's hands dirty' in the course of an evaluation, of playing an active role rather than being a passive observer. There are risks

entailed in the extent to which one is then able to be objective, but the extra data obtained thus make this worthwhile. This might be summed up in the slogan: *if you want to learn, do.*

3.6 Summary

Five case studies which led to the development of SESL have been described in this chapter. A discussion of the extent to which each affected SESL can be found in section 5.1. In summary, the cases were:

- *BigBank Financial Services*: A large financial service company had introduced a workflow system. I conducted a one-day observation of the use of the system. Key findings included the use of workflow as a 'process control' device and as a co-ordination mechanism alongside other tools, and that there was little visible sign of important events.

- *ShrEdit*: With a colleague, I conducted a short study of a group of students using a collaborative writing tool, to compare collaborative writing processes using various media. Key findings included the importance of talk as a mechanism in the face-to-face setting, but with the tool being used rather more in the distributed setting. The study led to the PETRA framework (Ross et al, 1995), which SESL has extensively built upon.

- *Health Centre*: I conducted a set of interviews with local health service professionals in Lancaster, with the aim of investigating and evaluating the use of computer systems to facilitate co-operation in their work. The systems seemed to entail considerable extra workload for little gain to those doing the work (rather to government). There was also a clear cultural divide between the medical and the administrative professionals.

- *University MIS*: I conducted a short evaluation of the new management information system at a university, chiefly providing financial information. Implementing the system had required considerable work from users, who perceived it to have been imposed from above (given the increasingly strained relationship between administration and academic departments). Thus considerable changes in working practices were required, and the system was still slow, difficult and inflexible (often not producing the information needed).

- *RICS project*: A one-year study of a research project team. Using various methods within the 'participant ethnography' tradition, I looked both at their co-operation using computers and using other media; and at issues of learning within the project (which had learning of others as its theme). I concluded that the project was fairly innovative in a number of different ways, but that it was difficult to tell what the causality of this innovation might be.

4. SESL: a methodology for evaluation

Both the theoretical and practical need for a systemic, stakeholder-focused methodology for the evaluation of co-operative systems have been outlined above. This chapter will describe a methodology developed to facilitate such evaluation. The chapter explains the principles behind SESL, the instantiation of those principles into various forms of methods, and gives a case study that describes the full use of the methodology.

4.1. SESL as a methodology

SESL is as much a way of looking at evaluation as it is a way of doing it. It is an attempt to look at evaluation in a different light. In this sense, it is a methodology (a meta-method, as discussed in section 1.1). Thus, this section describes the 'heart' of SESL, the real thing that makes it tick. The four words of the title describe the form of evaluation under study: it is systemic (that is, it looks at the whole system), it takes account of stakeholder needs, and it has learning as its chief focus. Standing outside of that phrase but very much within the ideas of SESL is the notion that what is being evaluated is a co-operative system, rather than a piece of co-operation technology. This section will describe these crucial issues. The sense in which co-operative system is used here was discussed in section 1.3, and so is not repeated here.

4.1.1. Systemic Evaluation

As discussed in section 1.2, the term and concept of evaluation means many different things to different people. The form of evaluation being discussed here is a *systemic* one. For something to be systemic, it must be based on the whole system, taking a holistic view of the situation under study.

Problems with this must briefly be addressed. As already discussed, the word 'system' has become appropriated by technologists to denote things to do with computers. In a sense this is the same kind of holistic thinking as mentioned above: a whole is identified that has properties that do not exist in the separate parts. However, this use does confine the system to only the artificial parts of the mixture. A second problem of language is that while the principle of looking at wholes rather than parts has been followed by a number of different schools of thoughts in different ways, all too often the practice has become conflated to a particular form of study. This involves a set of tools from mathematics and electrical engineering: feedback loops, input/output, control centres, and the like (sometimes, but not necessarily, expressed in equations or simplified "lines and boxes" diagrams). This is seen, for example, in the semi-popular work by Senge (1990). To be systemic is to consider the whole system, to look at something overall and be aware that it contains properties that the individual parts do not.

A further issue is what one might call the shopping list problem: there are so many perspectives to be taken into account, so many disciplines to choose from, that one ends up with an enormous long list that then needs to be fulfilled on each occasion. Either this takes a massive amount of time or corners are cut. One must be careful not to try to cover every detail, but rather only those appropriate for each evaluation.

The core of the evaluation process is the interaction between the evaluator and the people who are part of the system; and to a certain extent, with other aspects of the system such as the technology. There are many approaches to such interaction – which are referred to here as observation, although this is perhaps too 'hands-off' a term – but ultimately they can be stripped down to three basic styles: talking to people, watching them, or getting them to do things. Sometimes two of these come together, but these are the basic forms. The *names* under

which these forms are encountered, and the philosophy behind them, varies: but interviewing, ethnography, and experimentation will serve to summarise them.

Whichever of these approaches is taken, the same basic cycle between observation and analysis can be seen: the evaluator carries out some activity with the people in the system to learn more about them, then thinks about what they have found (using whatever method of analysis is appropriate to the situation and comfortable to them) and then observes some more. This process goes on as long as necessary and/or feasible.

Of course, different approaches have different views on the relationship between observation and analysis – in laboratory-based experimental methods, a research question is established in advance which is then tested against; in ethnomethodological ethnography, by contrast, key issues from the evaluation arise from the process of observation, and are not to be defined in advance. Of course, both these descriptions are stereotypes – experimentalists take issues from the situation and ethnomethodologists have some questions in their minds before they start – but there is a broad scale. While SESL is, in principle, compatible with methods along any point of this scale, it probably will tend to sit in use somewhere in the middle – the key issues to consider (discussed below) are more 'etic' (that is, defined from outside) than is preferred by ethnomethodology, while the focus upon effect is more 'emic' (arising from the situation) than preferred by experimentalism.

Any of the basic approaches will fit into SESL: the spirit of the methodology is in the questions that are asked and in the issues that are examined, not in the specific way questions are asked. Nonetheless, it is worth examining each of the basic approaches in turn and seeing how they fit into the methodology.

First are the approaches whose main activity is talking to people. This can range from individual interviews, through focus groups, to group exercises like Stakeholder Mapping (see below). Typically, the data collected are qualitative – verbal accounts or graphical representations, the discussions take place at or near the workplace, and they are moderately intrusive in that they require people to stop working for some time (but often only a short time). They are often semi-structured, with a few basic issues in mind – and this makes them especially compatible with SESL, which takes a semi-structured approach.

Next are approaches based around watching people doing things (that they would be doing anyway). Broadly, this is known as ethnography by its practitioners (although strictly that approach requires a long-term study). The relation between the method and the analysis is heavily prescribed by the ethnomethodologists who work extensively in this way – but other analytic strategies are possible. Watching people is a very useful form of acclimatisation to a situation, as a way of forming views of what is happening without the interpretations of the participants colouring these[4]. It fits well with SESL in these ways, although the principles of SESL would be in conflict with analytic approaches that wanted to bring in no external 'categories' (the sociological term for ideas used to structure one's understanding of a situation). This is not the place to argue the possibility or otherwise of such an approach, but it is not compatible with the assertion of this thesis that evaluation must, if it is to be successful, be structured around notions of the whole system, the needs of all stakeholders, and the support for ongoing organisational learning.

---

[4] While useful, this a moderately risky stance, open to an expert's arrogance; but it fits the realities of situated research.

The third main sort of observation/analysis is that of getting people to do things. This is maximally intrusive, as it removes them from their normal activity (and, sometimes, their normal *place* of activity) and has them perform tasks set by the evaluator. This is useful in two main circumstances: when one is considering a very limited, well-defined problem, which can be sensibly removed from its context; and when one is studying a computer system on its own, devoid of normal context, for example before releasing an off-the-shelf product to market. This is the province of the experimentalist, who claims great generality for his/her method because it is context-free (or, rather, the only context is the neutral one of the laboratory). No doubt this approach is useful in these limited settings – for single-user usability testing in particular (at the level of "what should my icons look like?") it is vital. For wider-scale evaluations of the kind envisaged within SESL, it is less useful. It can, however, be incorporated into one of those evaluations as a part – there is no reason whatsoever why an experiment should not fit into a SESL study. Indeed, the ShrEdit study which led to the PETRA methodology was partly an experimental one, insofar as we put the participants into a situation of our devising (albeit familiar) and had them perform tasks dictated by us; although we didn't have explicit hypotheses to prove, as is usually the case with experiments.

These are the three main activities of observation and analysis, then. The discussion above tacitly assumes that data thus collected are qualitative. Have quantitative data any part to play in the methodology? The answer is similar to the one above concerning experiments –they may well be relevant as tools within an overall qualitative framework. For example, cost-benefit analysis is crucial to many evaluations in business. It cannot be denied, though, that the paradigmatic assumptions of SESL are those associated with qualitative research – it deals with people and their needs of and experiences with co-operation via computers. The stakeholder approach as used here is especially incompatible with quantitative data. Ultimately, the forms of observation and analysis used in SESL will depend on the evaluator and the situation.

4.1.2. Stakeholders

The literature in various fields which uses the concept of stakeholders was discussed in section 2.3.1. Here, the relevance of stakeholders – individuals or groups who can affect an organisation's work or systems, or are affected by them – to CSCW evaluation and to SESL is discussed.

A few typical stakeholder groups in a CSCW setting are those who use the software, their colleagues and managers, the software developers and retailers, the Information Systems department of their organisation (if appropriate), and perhaps the customers of the organisation. Wider groups (such as trade unions, parent companies, employers' associations, shareholders and governments) may also be stakeholders. A typical stakeholder map that combines these groups and individuals would look like the following:
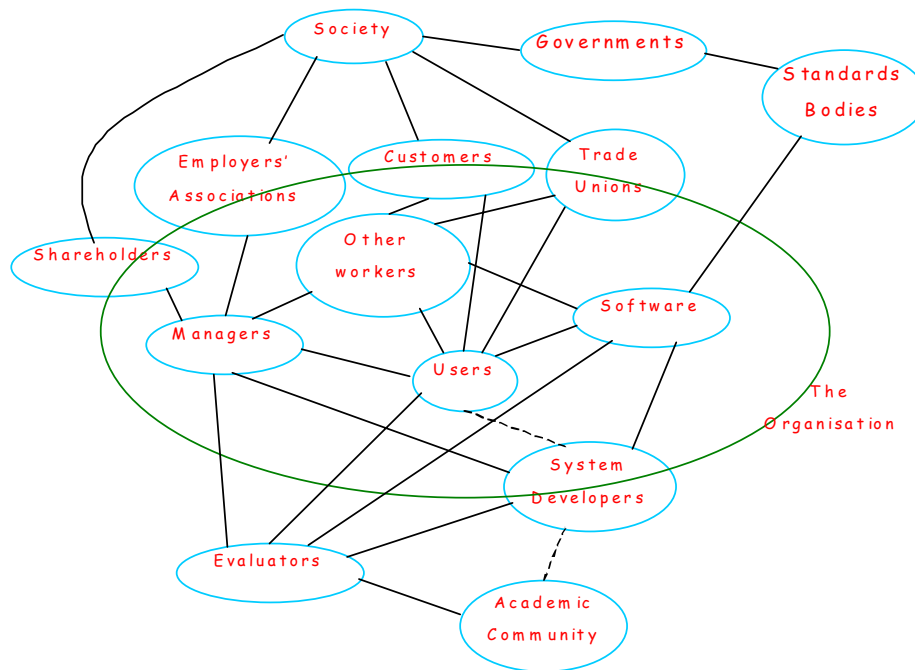
Figure 3: A generic stakeholder map for co-operative systems

The reason for the popularity of the term, in a variety of domains from management to economics and politics, is that as well as being descriptive of the groups within the system, it also is prescriptive in that it argues that certain groups and their interests are important. In this way, the stakeholder perspective replaces views such as those which wish only to focus on the interests of managers and shareholders. It is an attempt to be inclusive in one's view of the organisation.

The aims of such inclusivity vary according to the person using the term, but two reasons are typical: out of a sense of 'natural justice', or for instrumental cause. First, the concept is sometimes used by those who believe in workplace democracy, the breaking down of hierarchies, the "quality of working life" and so on. For such people, the consideration of all stakeholders' interests is an end in itself. On the other hand, there are also people who seek to use the stakeholder concept as a means to a different kind of end, typically the survival of the organisation in an ever-changing world. In the report *Tomorrow's Company* (RSA, 1995) for example, the stakeholder model is used to illustrate the way in which companies need a certain coalition of interests to be behind them in order to survive – such as employees, customers, shareholders, managers and governments. If the needs of some of these groups are neglected, that group may withdraw its approval for the further survival of the company, leading to its ultimate demise.

The problem of conducting evaluation in a situation where there are multiple stakeholders, of the kinds given above, all with potentially conflicting interests, is then relatively clear: If one is evaluating according to some pre-defined criteria or objectives, then whose criteria? Those of the employees may contradict those of the shareholders, for example. Again, if one is looking at the effects of a particular system, then the effects upon whom? This makes finding a system good or bad – forming a single, binary, measure of its appropriateness – extremely difficult. If a workflow program installed in a particular company (e.g. Bowers et al, 1995) is said to be good for the managers (because it gives them access to detailed information about what their employees are and have been doing) then it may be bad for those employees (because they are scrutinised and so feel threatened, or because their work is altered by the overheads necessary to keep the workflow system fed with information). Indeed, it may be

good for some customers – those for whom price is an issue, and so increased financial control may assist; and bad for others – who have to deal with unhappy staff or who get work from them that is slower or less well done than before.

While this problem has not been addressed explicitly in the CSCW literature in the past, various solutions have been taken to it. Most of these involve defining it out of existence, pretending it is not a problem and either hoping it will just go away (the strategy of many ethnographers); or explicitly favouring the interests of one stakeholder group or another – typically managers (as in the systems analysis methods) or workers (as in the participatory design methods). This latter solution (privileging one group) is entirely legitimate, as long as it is recognised as a political move. The former solution (ignoring it) is a little more problematical: rather in the same way that someone who claims they are acting on common sense is all too often simply acting on the prevailing worldview of society, it tends to favour the status quo – as this often has sufficiently strong momentum to require some effort to push it in a different direction.

A solution often taken by those influenced by HCI is also worth mentioning: the labelling of two stakeholder groups as designers (in which are included evaluators) and users, and then the examination of the relationship between the two. Many discussions have unpicked the concept of user, pointing out that there are many different types of users, but the dualism often remains. In the development of the PETRA framework (Ross et al, 1995), much time was spent considering the question of multiple perspectives, of evaluators and users, and effectively creating a dialectical relationship between those two perspectives. However, as is often a problem with dialectical approaches, in the process we defined these as the only two perspectives and as opposite to each other. This is clearly incomplete: there are many more stakeholder perspectives than these two, and indeed the categories we presented are more complex than possessing just a single perspective, as there are many different kinds of user.[5]

The first solution to this problem is not seeing it as a problem but rather as part of the nature of a complex co-operative system, and building one's considerations around that awareness. Thus, any evaluation of a co-operative system should be designed with the multiple, possibly contradictory, interests of different stakeholders in mind. It should early on contain some kind of mapping of the relevant groups, and some attempt to address the issue of who one regards as the key stakeholders. Thus, the evaluator must to some extent be a facilitator of a negotiation process between the different stakeholder interests (Fisher and Ury, 1991).

This process may itself be enough, but ultimately it cannot solve the problem of there being more than one definition of goodness within and for the system. A further solution to this is to nullify it by reframing: changing the definition of evaluation to allow different viewpoints to be heard. This is the subject of the next section.

4.1.3. Learning

The view of evaluation advocated here is one that sees it not as a task of performing a study and giving back results, but rather as part of an ongoing process of learning among the stakeholders and the organisations to which they belong. For different groups, this learning will take place in different ways and at different times.

---

[5] The same argument can be readily followed against trying to create a dualistic opposition, whether resolved dialectically or not, between the interests of managers and workers (cf. Axelrod, 1984, on the superiority of the 'tit for tat' solution to the Prisoner's Dilemma Game).

So for system developers, the learning that goes on in an evaluation (whenever it takes place) may concern how their system could be changed to better fit the needs of real or hypothetical people; but it may equally well be at the level of transferable information, such as the 'post-mortems' that Microsoft are reported to hold at the end of a project, to allow lessons learned to be used in the next project (Cusumano and Selby, 1995). For managers, the learning from a system implementation may be technical ("don't use AppleTalk, use Ethernet instead"), strategic ("don't buy from company X, their products are terrible"), organisational ("groupware will really undermine our hierarchy") or various others. For evaluators, learning may be in terms of methods that did or did not work (what you might call action research). And so on along the list of stakeholders – but for each, the key is that learning takes place.

Why is this a useful perspective on evaluation? In the first place, it gives space for the various stakeholders to have different outcomes from the evaluation at appropriate times. Second, it recognises that much learning goes on simply because of the evaluator's presence. It is now commonplace for various kinds of social researchers, especially ethnographers, to take a reflexive stance – to be aware of their own effects upon a situation under study (Plowman et al, 1996). It is useful for this to be extended to evaluation. In the RICS study for example (see section 3.5 above), my experience was that the fact the project team had an evaluator connected to the project, hanging around and coming to meetings and asking questions, was enough to constantly remind them of their goal to be reflective as a team and to learn from their own group processes.

Of course, this is very much a hands-on role for an evaluator to take. One has to give up one's metaphorical white coat – the symbol of a disinterested scientist-expert, able to stand back and offer dispassionate advice from a position of omniscience – and replace it with a grey suit or whatever else is required to become part of the team. One's first allegiance will eventually be to the project rather than to science, although one cannot be a total insider if the evaluator role is to be useful. In the RICS project, it was certainly the case that my presence, and peripheral membership of the group, formed part of the effect of the study.

As discussed above, such evaluation-as-learning plays a considerable role in the facilitation of individuals' learning. It is also interesting that it ties in with a larger process of organisational learning. If one is conducting some kind of study of a system in use, then a certain amount of self-examination must have gone within the organisation before the evaluator enters the situation – to decide upon the system, to set it up and so on – and that reflexive process will continue once the evaluator has gone, as the organisation continues to learn from its experiences. Thus, the learning that the evaluator facilitates by their presence and their actions is part of a longer process.

This focus on process is crucial. Organisational learning (and indeed any learning) has very little to do with being presented with the results of something, but rather to do with the process that is undertaken to reach that goal. As the Chinese proverb has it: "Tell me and I'll forget. Show me and I may remember. Involve me and I'll understand." So it is that the most important thing in the evaluation is the doing of the evaluation, not the result reached (which should simply be a confirmation of what has already been learned). This means that each stage of the evaluation process becomes a learning experience for those participating in it – and this has much to do with the ability for this evaluation-as-learning to offer a variable pace to the various stakeholder groups.

This is very compatible with the sense of learning as an active process discussed in section 2.3.2. But organisational learning is not just about changing behaviours. Given the importance of double-loop learning, it is better thought of as being about changing the patterns that underlie those behaviours. Such pattern shifting is notoriously difficult and fraught,

particularly if it involves long-held patterns (Price and Shaw, 1998); but the results can be well worth it. This kind of evaluation will tend towards looking at first-order learning, as the aim is usually to improve rather than to alter radically; and most evaluation methods reflect this. However, it is certainly open to the possibility that the learning that is facilitated could be of the second-order: the systemic focus of the process, the awareness of all stages of the evaluation as learning opportunities, and the space given to all stakeholders makes this an option if it turns out to be necessary.

4.2.   Instantiations of SESL

This section moves away from principles towards practice. I have used SESL in various forms during the period of this research, and these forms can be seen as methods which instantiate the methodology. Four such methods are described in this section: the core version based on five stages, an extended version of these stages, a pro-forma version of the methodology, and a set of guidelines for evaluation derived from the methodology.

4.2.1.   The five step method

The following steps form an approximate picture of SESL as a set of rules to follow. There are five in all:

1. identify the type and purpose of the evaluation;

2. decide what is the system to be evaluated;

3. determine who are the stakeholders;

4. observe & analyse (the heart of the process), concurrently formulating a set of key questions;

5. encourage various forms of learning (such as reporting back to stakeholders).

Figure 4 shows the steps in a certain simplistic order – three need to be completed before the observation process, and that before the reporting back. There are risks in this diagram because of the false clarity it implies. Like the waterfall method of software engineering, the diagram shown here bears scant resemblance to real life. In real evaluations, one is *always* gathering new knowledge about who are the key stakeholders and what their stakes are, during the observe & analyse phase. Even more so, given the focus of learning in this work, to suggest that the only time any feedback should be given to clients be right at the end is unrealistic. Therefore, the diagram is purely for illustrative purposes of one way of looking at SESL.
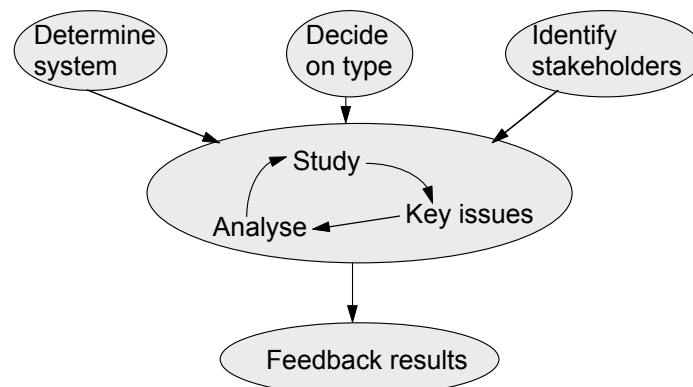
Figure 4: SESL as five steps

The rest of this section ventures inside some of the bubbles in this diagram and describes them in more detail. The following section, 4.3, gives an example of the methodology in use, in an evaluation of a research unit.

4.2.1.1. The study-analyse cycle

It was discussed above that the forms of observation/analysis available have many different features, but that all broadly share an iterative cycle of finding out about the nature of social reality, then making judgements about that finding-out via some theoretical or experiential framework, then going back to further data-gathering. Clearly related to this iterative process is the formulation of the questions one is trying to pursue. While the data-gathering approaches discussed in 4.1.1 differ in this, it is generally useful to know what one is trying to find out.

How to determine such questions? Some will be obvious from one's existing knowledge of a situation or an initial conversation. They are also likely to change over the course of the evaluation, as happened in the Metropolitan University study (see section 4.3). This changing of the questions over the course of the study is crucial: they should be used as the basis for each conversation or other activity, and modified as good ideas come along. That is in keeping with the principle of learning that underpins SESL – the learning of the evaluator about the study they are conducting.

What of other ways for determining key questions? One approach is the collating of all expert knowledge about the kind of system into a guide that tells which questions should be asked – and what answers should be expected. The 'heuristic evaluation' approach of Nielsen (1993), in usability evaluation, is a strong example; design style guides, such as that of Apple, are another. This is rather premature in the case of co-operative systems evaluation, since there is insufficient collective knowledge about what to look for; and the production of such a guide will have to wait for future work. One formulation of these are what might be termed the Eight Key Questions of evaluation. These are:

1. Does it work?   *(functionality)*
2. Does it work well enough?   *(efficacy)*
3. Is it workable with?   *(usability)*
4. Does it follow the *standards* laid down by various bodies?
5. What does it do to those who work with it?   *(individual effect)*
6. What does it do to their work?   *(group effects)*
7. What does it do to those they work with and for?   *(organisational effects)*
8. What does it do to the world beyond work?   *(societal effects)*

They can also be seen from figure 5, in which the inner rings correspond to the earlier questions above, and the outer rings to the later ones. The concentric circles illustrates the point that to some extent the later questions depend on the earlier ones (e.g., one can't determine the usability of a program if it habitually crashes); although this does not necessarily make the inner rings more important than the outer ones. It is interesting to reflect that most evaluation practice does not reach beyond the fourth ring, standards.
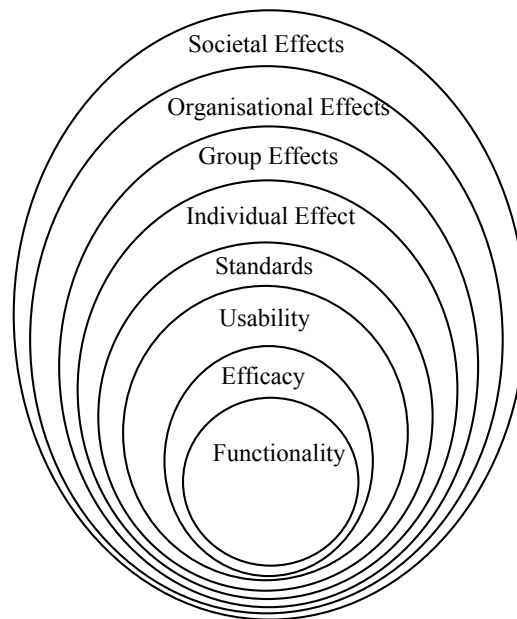
Figure 5: Eight layers of evaluation

The final possible approach is based on the type of the evaluation being undertaken. The question of types of evaluation has been addressed in the general case above; its specific use in SESL will be discussed below. The types lead to a basic question for each, as follows:

1. ACTUAL EFFECTS:
   What are the system's effects upon: a) the work of the group using the system? b) the life of the group? c) the life of the people in the group? d) the life and work of the people outside the group? e) the organisation(s) of which the group is a part? f) society?

2. ACTUAL OBJECTIVES:
   What are the system's objectives (from the different perspectives of the various stakeholders)? To what extent are these being met?

3. POTENTIAL EFFECTS:
   What are the potential effects of the system upon: a) the work of the group using the system? b) the life of the group? c) the life of the people in the group? d) the life and work of the people outside the group? e) the organisation(s) of which the group is a part? f) society?

4. POTENTIAL OBJECTIVES:
   What are the objectives for this new system (from the different perspectives of the various stakeholders)? How well are they likely to be met?

Again, these general guidelines need specific modification for the needs of a situation. In particular, the references in questions 1 and 3 to various groups (while related to the concentric circles above) may well be affected by who are considered the key stakeholders in the situation. Ultimately, only asking oneself, "what am I trying to find out here?" (or, better yet, asking it in collaboration with others) will give a useful answer. However, some of the above might be pointers along the way.

4.2.1.2. Stakeholders

The next question to ask is, who are the stakeholders in the system and what are their interests? Many of the answers can be found in the discussion on stakeholders in section 4.1.2 above. In practice, the most helpful approach is to ask who affects, depends on or can influence the system; and likewise who is affected by it or is influenced by it. This gives one set of ways of looking at who are the stakeholders.

Another way is to take a list of typical stakeholders for the relevant domain and see to what extent they apply to a given system. Typical stakeholders in CSCW are those listed above: those who use the software, their colleagues and managers, the software developers and retailers, the Information Systems department of their organisation (if appropriate), and perhaps the customers of the organisation. Wider groups (such as trade unions, parent companies, employers' associations, shareholders and governments) may also be stakeholders.

A third way is to have a representative group collectively construct a stakeholder map: a diagram showing who are the key people relating to the system by showing the system in the centre of the map and then the distance of others from the centre, showing their importance to it. This diagrammatic form can help to identify groups who might otherwise have been forgotten. An example of such a map is that which was drawn up by the project team and myself during the RICS study (figure 6, below). Note how the proximity of the stakeholders is used to show closeness to one another (even to the extent of overlapping) and to the centre; and how three groupings were ultimately identified. Names in this diagram have been changed.
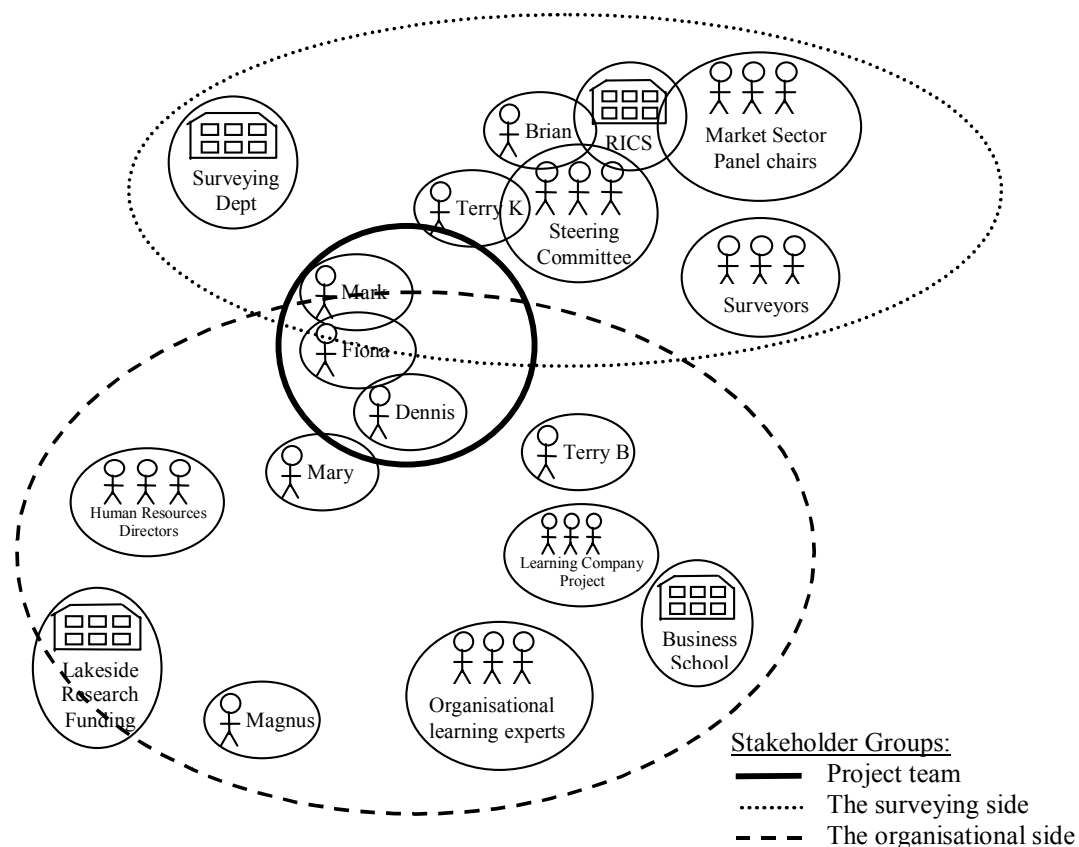
Figure 6: Stakeholder map from the RICS project

Researchers on stakeholder analysis, especially Mason & Mitroff (1981:94-103) have described several more methods for identifying them, including:

- taking standard demographic groups (sex, age, seniority in company etc.) and considering their relevance;
- asking various people who they think are the key stakeholders;
- studying (or creating) ethnographic accounts to find those who express pertinent interests.

Another significant question (which is assisted by the process of stakeholder mapping) is who are the key stakeholders: whose views need most to be met? Whose involvement is so vital

that the system will break down without it? It is important to make it clear who these are, as this will alter how much weight should be given to different aspects of the evaluation. It is also useful to be clear about what are the perspectives of these key stakeholders upon the system.

It must be stressed that the lists of stakeholders, their interests and relationships gained by these methods are subjective: different methods (and different participants) will produce different lists, and they will change over time. As Korzybski wrote, "the map is not the territory" (quoted by Bateson, 1972:449). For this reason, it is greatly desirable to produce them collaboratively, preferably involving both insiders and outsiders to the organisation.

4.2.1.3. Type and purpose

In section 2.2.1, the notion was discussed that evaluations come in several types – or, put another way, are conducted for many purposes. There is a crucial difference between a study of a co-operative system in place within an organisation, a study of a computer program under development for sale in the general marketplace, and a study of an experimental system developed for the sake of research (to name just three). The lack of recognition of that diversity of types causes many problems.

The following questions are a starting point for determining type:

- Are you looking at the effects or objectives of a system?
- Are you looking at the way things are now, or their potential?

The answers to these can conveniently be placed into a 2x2 matrix (using the first question as a horizontal axis and the second as a vertical axis), in which can be seen the main types of evaluation described in section 2.2.1:
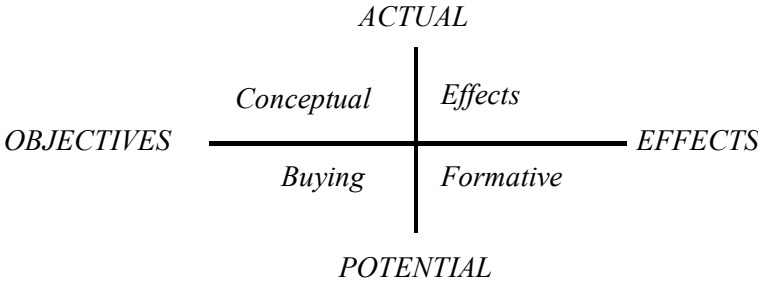
Figure 7: Four types of CSCW evaluation as a matrix

Why should this be at all important? One answer is the use of the key-questions list in section 4.2.1.1 above. Much more significant is to understand why one is conducting an evaluation: to what end? With what purpose? There may be many purposes. One way of looking at this is to say that the purpose of an evaluation is to satisfy the needs of the key stakeholders – in which case there would be as many purposes (not necessarily overlapping, sometimes not particularly compatible) as there are stakeholders. In that case, it is not necessary that these should be resolved in favour of a single purpose – that the evaluation will satisfy many needs is not a problem.

The chief type of evaluation described in this thesis is the Actual-Effects. Is this the only kind of evaluation possible with SESL? It is undeniable that the methodology is particularly compatible with this type, with its focus on a specific co-operative system embedded in an organisation. It is by no means limited to those situations, but my main experience with the methodology has been on such evaluations.

4.2.1.4. Systems and technologies

"What is the system under study?" might seem a trivial question to a computer-focused programmer, who is clear what they have built and just want to study it. In talking about co-operative systems, things become harder – the analysis moves from looking at a technology, which is affected (crucially) by its setting, to the whole combination of people, organisation, politics, technology and so on that is involved in the co-operative process.

This makes much clearer the remarks of Checkland (see above) that researchers construct systems rather than observing them. It is the evaluator's job to sift through all that complexity and determine *what* it is they want to evaluate: which people, doing what tasks, within which organisations or parts of organisations, to what ends. In particular, which co-operation technologies do they use to structure those tasks?

How is one to determine the system under study? As with some of the other questions, it may be rather obvious – if one has agreed to study the use of Lotus Notes in facilitating the co-operation of the Claims Division of the Warm & Friendly Insurance Society, then that would seem fairly clear – one goes to that division talks to/ observes/does experiments with the people there about their use of Notes, in the light of the way co-operation goes on in the organisation, and so on. And yet: does one just look at Notes-facilitated co-operation, or does one also look at co-operation facilitated by talk, meetings, paper, phone, post, fax or email (and the way this is influenced by the use of Notes)? Is only co-operation with people in the division of interest or is their work with customers, other parts of the Warm & Friendly, regulatory bodies, society etc. also of interest? (And if so should one then be working with people from those places too?) If the study seems to broadly indicate that co-operation in general – the organisational culture's attitudes to working together, their sense of 'we', their power games/struggles/structures, their myths and self-understanding – has more effect than co-operation through Notes, is that to become the main focus of the evaluation, or confined to a paragraph beginning: "Of course, effects are also seen of…". There are no rights and wrongs here. The ideal approach is as broad as possible, but this can be seen as a luxury many an organisation cannot afford – and it is quite clearly impossible in any case to study everything, so that choices will have to be made somewhere. The system has to be set boundaries, which are to some extent artificial but are necessary to limit the scope. This boundary setting must be recognised *as* a choice, for which alternatives could have been found; "the system" is always wider than a given technology – the only question is how wide.

Checkland and Scholes (1990:31) give some helpful advice on how to be clear what system one is studying. They suggest that two basic kinds of system can be identified: the *primary-task system*: "some organised purposeful action which could be reflected in the choice of a notional human activity system whose boundary would coincide with the real-world manifestation" (that is, the boundaries of an organisation); and the *issue-based relevant system*, which considers the working out of a particular conceptual issue (e.g. conflicts over resource allocation) as its boundaries. To assist in formulating these systems, they also suggest the use of metaphors to describe the relationships between different important parties within the system – e.g. master/slave, husband/wife, brother/sister, organism/virus etc.

4.2.1.5. Different forms of learning

The idea that the 'output' from an evaluation process is not just a report but learning for all the stakeholders at various times and in different ways, is one that has been stressed often in this thesis. The last diagram denoting the process of SESL above (figure 8) makes this clear: many different kinds of learning emanate from the whole process (at different times). The diagram is drawn thus to make it clear that it is not just the report that counts as learning, and

that learning doesn't just happen at the end of the process. Rather, it begins when the evaluator comes into the situation and goes on long after they have disappeared.

The bubbles in the diagram are not meant to be all the kinds of learning that can go on, but rather examples of particular kinds of learning that can occur. A few of these broader kinds are discussed below – others kinds also exist.

- *Stakeholder learning*: for particular groups, there will be learning outcomes specific to them: managers learn one thing, workers learn something different, the evaluator learns their own things, etc. It is crucial to recognise that this learning will differ from group to group – supporting and enhancing that is a crucial part of the SESL process. Other forms of learning will affect some groups more than they will affect others, but they are discussed together below.

- *The state of things*: learning how things are, could be, and should be are often crucial to groups taking things out of an evaluation process. The evaluator's job is to tell how things are (from their outsider's viewpoint – since it's the outsider's view, as well as particular expertise, that makes an evaluation especially useful) and to advise on how they think – from their knowledge of technology, other organisations etc. – they could be improved. Neither of these is necessarily correct just because the evaluator says them – only the people involved can say how things really are (although an outsider's view may well clarify the matter); and options for change are just that – other possibilities always arise. The evaluator cannot tell members of a group what they want to change, but the evaluator might be able to structure the decision-making process. A useful way of doing this is the Future Workshop (Jungk and Müllert, 1996) which structures discussion on a problem into Critique ("what's wrong?"), Fantasy ("what's the ideal solution?") and Implementation ("how do we get there?").

- *Ongoing learning*: the process of performing the evaluation often leads to a great deal of learning in itself – just considering questions about why you do things the way you do, who are the key stakeholders and such, *is* the learning from the evaluation. The evaluator can also often give practical advice along the way. If the evaluator is reasonably knowledgeable about a computer program being used, and the people they are working with are less so, then help with that program may well be appreciated. This kind of learn-as-you-go can often be rather valuable and, if the evaluator's primary interest is academic, can be a means of 'paying their way', and gaining trust, before the later stages of the process give more concrete outcomes. In particular, if the ultimate recommendations are more geared at management – which they should not be, but sometimes are – then this can be a still-useful learning experience for other workers. The same is true of co-operation skills – just a few mentions along the way of techniques like Stakeholder Mapping can prove more useful than the final result.

- *Formal reports*: the report-making process – whether that takes place on paper or in a talk – is another form of learning, of course. Sometimes it is the case that evaluators only conclude things once they think about them fully enough to write them formally – which has been my experience. Moreover, there are people who hear ideas best when those ideas are presented in a structured, official manner, which is then archiveable. So reports certainly have a place – but if they are the *only* learning output from the evaluation, then they risk simply not being heard because they are too threatening when unexpected. It may well be that many short reports along the way will produce more results than a big formal document. It is also harder to put an idea that has been surreptitiously spreading through an organisation courtesy of an unassuming little evaluator, on a shelf to gather dust – the meme is often mightier than the pen.

- *Other forms of learning*: who knows what else can come out of an evaluation? This will only become apparent as the evaluation goes along.

Determining which learning should be supported is not difficult – all of it. How to do this may also be easy, as often it will simply happen without a great deal of extra intervention on the part of the evaluator. Rather, the evaluator needs to be aware that it will happen, of the potential for spreading their ideas, and that the whole process of evaluation is a learning process – for the evaluator as much as others involved.

Learning is difficult to pin down and label. It may not be clear until long after it has happened. Nonetheless, the learning process that goes on throughout a SESL-based evaluation is very important.

4.2.2. The extended view of SESL

There now follows a diagrammatic view that, with its rather less fixed model of the process, better reflects the way in which SESL is done. At its heart is a basic process – a cycle of observation and analysis – that starts when one begins an evaluation and goes on until the end. Feeding into, and coming out from, that angle are four basic issues to determine:

- What is the system and its key technologies?

- What is the type and purpose of the evaluation?

- Who are the key stakeholders? What are their stakes?

- What key questions are we asking here?

Resulting from this process (and feeding back into it) are a series of different kinds of learning – one of these will be a final report on the evaluation, but there are many others. These begin from the start of the evaluation and go on until the end.

The following diagram modifies figure 4 (section 4.2.1) to make these processes clearer:
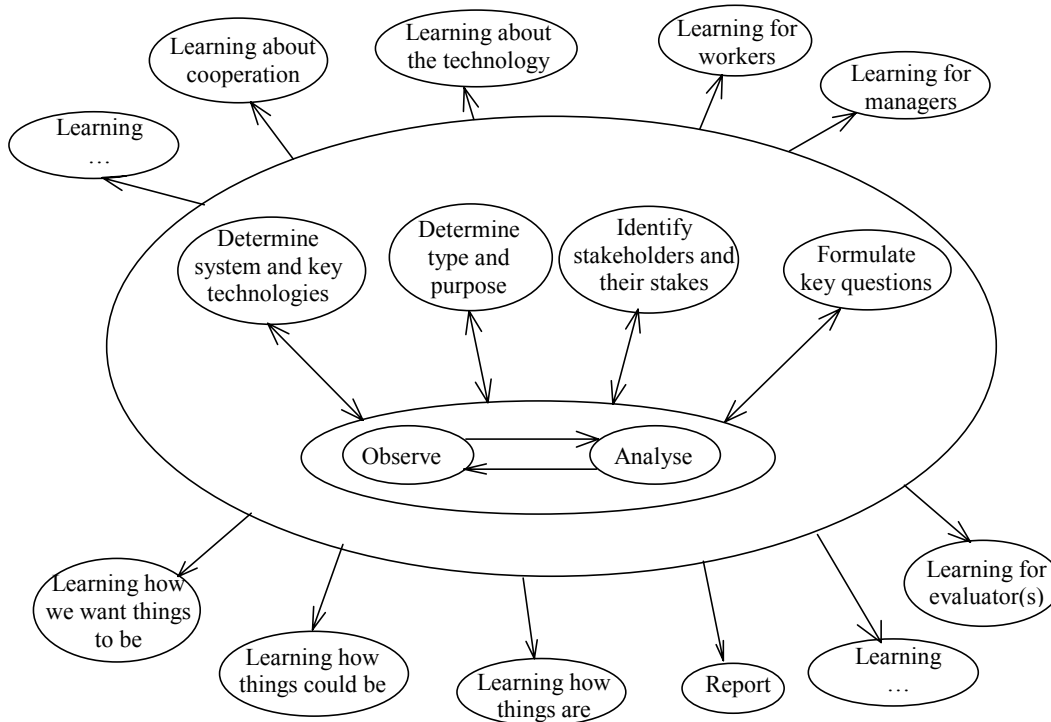
Figure 8: SESL – the extended view

Further extensions to this diagram are possible, but have not been pursued here. In particular, there is little sense of the feedback loops involved – all the processes shown by arrows take place over a long period, and there will thus be a complicated set of interactions between the different activities. An additional flaw in the diagram is that, in putting the observe-analyse cycle at the heart of the process, it might also appear to put make the work of the evaluator most important. SESL is a collaborative process between various stakeholders, in terms of input, ongoing activity, and output; and so such an 'evaluator-centrism' is by no means intended.

4.2.3.   SESL as pro-forma

The most basic way of using SESL is as a framework to guide ones initial thoughts. The following headings to do this summarise the main concerns of the methodology:

- System,
- Main technology,
- Type (of evaluation),
- Stakeholders,
- Methods,
- Key questions,
- Analytic framework.

This 'pro-forma' can be produced very quickly from initial discussions and provides an excellent point from which to discuss further the evaluation with the client. Some of the language may be unfamiliar – 'stakeholders' often causes problems, in my experience, as an unfamiliar term vaguely redolent of management jargon and of politics. This may be a helpful, although uncomfortable, form of 'reframing', challenging the assumptions of those involved through new language; it may alternatively get in the way of dialogue, in which case an alternative term (e.g. 'interested parties') may be useful. The list thus produced then proceeds to be a guide through the evaluation, to remind one what one has intended to look at; and provides a useful 'prop' in discussions such as interviews as part of the evaluation. It was used in the early stages of the Metropolitan University evaluation; this experience is discussed in section 4.3.1 below.

4.2.4.   SESL as a set of guidelines

The desirability (and difficulty) of producing a 'cook-book' set of evaluation guidelines was mentioned above. A further attempt has been made by Fides Matzdorf and myself, on behalf of a workshop of higher education facilities managers run by the Facilities Management Graduate Centre, Sheffield Hallam University[6].

The guidelines are designed to help managers evaluate management information systems within their organisations, which are defined in the guidelines as any system that holds and gives out information that is necessary and/or helpful to manage an organisation. This includes card-files as well as sophisticated software, people and their knowledge as much as technology – that is, co-operative systems.

The guidelines are organised in four main headings (in the form of questions), which are based on the main issues discussed in the five-step version of SESL in section 4.2.1 (*decide*

---

[6] The guidelines refer to Management Information Systems (MISs), as they were written for a workshop on that topic; but they understand MISs as co-operative systems.

on type and *feedback results* are included as one question here, as they are different aspects of deciding the purpose of the evaluation). Each main question is followed by brief explanations of the purpose of the question, and further sub-questions. The full guidelines can be found as Appendix A – in summary they are:

1. What is the purpose of your evaluation? Why are you doing the evaluation? What are you doing the evaluation for? Are you looking at the way things are now or their potential for the future? If you have been asked to do the evaluation by someone else, what is their angle on it? Will the main outcome be one of feedback or decision-making? Can the results be beneficial for some or for all stakeholders? How can the results be made available? How can the results of the evaluation be used to further best practice or be used as part of a benchmarking process?

2. Are you mainly concerned with technology, social processes or with a mixture of the two? Is your chief interest general or specific, broad or narrow? Which of the following aspects are relevant to your intended outcomes/purposes: technical performance, appropriate functionality, user-friendliness, meeting standards, effects on individuals, effects on groups/teams/departments, effects on organisations, effects on society…? What does the system do? What is it supposed to do? What is the interaction with other systems and with other media?

3. Who are the stakeholders and what are their interests? Who are the users/customers/clients of the technology? Who are the **key** stakeholders: whose views need most to be met? Whose involvement is so vital that the system will break down without it? What are the interests of the various stakeholders involved?

4. What are the key questions you need to answer? Who do the main users co-operate with, and how? What kinds of co-operation go on? What are the key MIS technologies, and what is their role? How could the technologies be better used? How do your MISs fit into the overall management of the organisation? What budgetary opportunities and restraints are there? Is there an overall IS strategy in your organisation, and how does it fit into your business plan? How could the technologies be better used?

4.3.   *Example of SESL in use: the Metropolitan University study*

This section gives an example of the SESL framework in use. This derives from my evaluation of co-operation, with especial reference to the PC network filestore, at a research unit ("the Unit") of "Metropolitan University"[7], in the spring of 1997. At the time, the Unit was a unit of ten full-time members of staff and one research student, and one part-time staff member based two hundred miles away from Metropolitan University, who travelled there regularly. They have two main activities: teaching on a number of courses (Masters and Diploma) for post-graduate students wanting further professional development; and conducting research projects for groups of professionals from public institutions such as the health service.

The Unit was then relatively new (3-4 years old) and thus had a culture of weakly-structured co-operation: until not long before the study, they were all based in one large office, almost all attended it during office hours, and so co-operation could take place in ad-hoc, informal ways (shouting across the room, listening to others' conversations, having a quick discussion as needed). Progressively, this changed: the Unit has become bigger in staff numbers, its tasks

---

[7] Again, names have been changed.

more diverse, and it had then come to occupy several offices (five in all). More formal forms of co-operation had thus become increasingly necessary.

In the spring of 1997, two members of staff became increasingly concerned about the way one technologically mediated form of co-operation was developing. Metropolitan University uses a Novell/Windows 3.11 network, with individual and group directories on the file-servers assigned to drive letters. For some time, the Unit had maintained a group directory, referred to as "the I: drive" from its Novell assignment, with the intention that it be used as a common filestore for sharing files and making them generally available. The drive was being ill used: some people were not putting files on it at all, while others were putting them there in a disorganised and inconsistent manner. In short, the drive was a mess: no one could be sure what was on there, and no one knew where to find things that they knew were there. As business expanded and the complexity of directories and sub-directories increased, this became more and more of a problem. The two staff members, knowing of my research, asked me to consider the matter, and to see if I could assist the Unit in re-organising the drive to make it more useful.

This, then, was initially a concern with the right use of technology. However, it was clear from early discussions that the matter was rather more complicated than just how they used the filestore: it seemed that their basic model of working was not one that was helpful to any kind of co-operative effort beyond the ad-hoc kind mentioned above. Once we had established this, I put the question of the study to the director of the Unit, offering him a form of 'free management consultancy' of mutual benefit, to initially look at the I: drive but with an awareness both of other forms of co-operation and of the attitude of the Unit to co-operation in general.

This section will describe how SESL was used as a framework to guide the evaluation. It is based on the stages found in the extended diagram version of SESL, figure 8 above. The full report was presented to the Unit in June 1997 – it gives more details on the case and can be found as Appendix F.

4.3.1.   SESL as pro-forma

The SESL pro-forma seemed to be a good way to start this study. From initial discussions, I produced a pro-forma description of the study; and showed this in draft (hand-written) form to the Unit's director when we discussed the potential of the study. These discussions led to a redrafting of the pro-forma, which in turn formed the basis of an initial email message about the study to the Unit's members, and of the questions asked of them during the study. The general outline of a pro-forma is discussed above: the particular pro-forma used here is as follows:

- System: co-operation within the Unit, via network drive
- Main Technology: Filestore [I: drive]
- Type: formative, in-use, socio-technical
- Stakeholders: members of the Unit; customers; IT support staff (Pete & Kev) as suppliers; rest of the School[8]; rest of Metropolitan University (e.g. Facilities Directorate)
- Methods: Short interviews; hanging around; Stakeholder Mapping?
- Key questions:
  - What kinds of co-operation go on?

---

[8] Metropolitan University is split into several Schools.

- What's the role of the common filestore?
- Does it help or hinder co-operation?
- How could it be better?
- Analytic Framework: Distributed Cognition [cf. the engineers described by Rogers, 1992], Common Information Spaces (Bannon and Bødker, 1997)

This was the initial version produced; various aspects of it, especially the key questions, mutated as the study went on (these are discussed under Key Questions below). That the focus on the System moved beyond just the network drive has been mentioned above. My changing view of the appropriate Methods and Analytic Framework are discussed in the next section.

4.3.2.   The process of observation and analysis

As discussed above, the cyclic process of observation and analysis are at the heart of a SESL-based evaluation study. Of course, this is both truth and untrue: they *are* the study, the whole of it and not just a component part. The observation process begins as soon as the evaluator first hears of the evaluation situation, and the analysis thirty seconds after the observation. (Whatever method of analysis one uses, however formal or rigorous, it always has some component that is intuitive enough to begin working in the evaluator's mind immediately.) This was so in the Metropolitan University study. I first heard of the problem several weeks before conducting any interviews; time passed before matters were discussed with the Unit's director and an email sent to staff members; then a few weeks more passed before the interviews took place; and it took a few weeks further before reporting back my findings to members of the Unit. It was thus a lengthy process: but throughout it, observation and analysis took place.

How was I observing? One method has been discussed: the kind of quiet observation that is the heart of ethnography (albeit rather brief in this case). This involved short conversations with the Unit's members and a process of acclimatisation myself as I tried to understand the Unit's organisational culture, not just in the details of what it did, but *why* it did things. What made it tick? Why did people come here? How did work go on day-to-day and moment-to-moment? I spent some time finding these things out. Partly this was incidental to other activities; some of it went on while sitting in the Unit's largest office waiting to talk to its members; partly it was by design to get a richer feel for their co-operation than their reports would allow.

This provided rich background data. However, the main form of *explicit* information was gathered through interviews. I spoke individually to all of the Unit's staff members (eleven in all). Depending on the rapport felt with the 'interviewee', I used the key questions above to varying levels of detail. When I felt little connection with the interviewee I stuck to them closely, using them as a template for the interview in a formal, structured manner; at other times, they merely served as a guide for our discussions, a set of issues that we must cover. In either case, however, the list of questions were made explicit to the interviewees, so that the interview became more of a conversation, with a shared responsibility for its structuring. I was able to do this without any sense of its being artificial by having sent out the questions some time in advance, so that by showing them to the interviewees I was 'just reminding them' what we would cover. I took notes in some quantity, but did not tape-record the interviews: this felt more natural, less intrusive on this (already considerably intrusive) occasion.

The interviews had one strange feature. In each of the two main offices, three people worked together closely and hence sat together. In interviewing each of these people (and this was especially so in the larger office), it was quite clear that their colleagues were listening – and

they even made comments on occasion as to what they heard. Obviously, this made our conversation less likely to yield personal remarks, especially criticism; but it did reflect their views-in-situ well. It was fairly clear both during and after these six interviews that conducting them in private (even if in the corridor!) would have been better[9], but somewhere between the dynamics of the situation and my own fears of transgressing what was acceptable behaviour didn't allow for this. As it was, the interviews were held at each person's desk, and hence in their natural surroundings. This meant they were so much steeped in their everyday interactions that they couldn't help but answer in a way that was very grounded in their working-day reality, unlike the separateness achieved in many interviews.

The third possible form of observation mentioned in the pro-forma was to conduct a Stakeholder Mapping exercise. This was discussed briefly above. It is a collaborative exercise developed for use in the RICS evaluation, which involves a key group mapping out who they think are the stakeholders in a situation/system/project, what are their relationships to one another and the core of the situation, what are their stakes, etc. It is an excellent way to encourage a group to discover things about itself; but in this situation there seemed, in the event, not to be the will to come together as a group and discuss things – they were neither enough of one mind nor likely to be willing to collectively give up that time. Therefore, I incorporated the question of stakeholders into each individual interview. As it happens, when the Unit *did* come together as a group and discuss my observations of them (as described in the Learning section below), it was a thoroughly worthwhile and animated discussion, and much was achieved. However, perhaps when I was conducting the interviews was not the right time for this.

The analysis was grounded in the way I had collected the data, according to the principles of SESL (not least an awareness that the observation/analysis process *is* cyclic). The analysis also rested on several sources:

- The Distributed Cognition analysis of Rogers (1992), concerning the use of notice-boards in ad-hoc co-operation by a group of engineers; and the Common Information Spaces concept discussed by Bannon and Bødker (1997), which talks about shared virtual spaces very much like the network drive used in the Unit. These were specifically highlighted in the pro-forma.

- The general CSCW literature. In particular, the discussion of information reuse and keeping records on this is strongly based in organisational memory and knowledge debates; the advocacy of a well-defined directory structure is based on the experiences of the PoliTeam project (Prinz and Syri, 1997); and the need for information to be public by default is informed by the work on Common Information Spaces.

- The concepts found in this thesis. The precept of socio-technical systems is found early on (the choice to look at co-operation in all its forms, aware of their relations to each other, rather than just technology); the concepts of stakeholders, organisational learning and systems thinking go right through it; a respect for the experiences and needs of the Unit's members is crucial to the report.

---

[9] This view is open to question; one of the Unit members subsequently pointed out that on other occasions, people in the Unit have shown themselves rather open to discuss their experiences with the Unit in anonymous settings, but rather reticent in more identifiable settings. My one-to-one interviews would have fitted into the latter category, under this view, meaning that no further information might have been available through such an interview technique.

- A knowledge of current computer technologies – the report talks quite extensively about the benefits of Windows 95 and Lotus Notes (or equivalent groupware tool) to them to replace the Unit's existing I: drive.

- A set of academic ideas from the social sciences and organisational theory. Uses of these included the illustration of who worked with whom (using a diagram common among social scientists, figures 10 and 11); the discussion of the key stakeholders; the numbers of offices they work in and the ad-hoc, face-to-face nature of their co-operation (using ideas from ethnomethodology); discussing formal co-operation and the organisational cultures found in the Unit and the rest of Metropolitan University (the former being rather graphically described by its director as "an island in a sea of crap"), using concepts from organisational theory.

- Three semi-popular management books: Mintzberg (1979), Covey (1989) and Peters and Waterman (1982) – these were principally used illustratively in the report to the Unit.

However, while these sources all contributed to the analysis, most were not used directly in reporting to the Unit. It was important that the report not be seen as an academic paper, but rather practically focusing on the issues for the Unit. Thus the theory appears only in the background.

4.3.3.   Key questions

Moving to the four input-bubbles to the observation-analysis process, it will not be a surprise that they turn out not to be inputs to that process at all, but rather integral parts of it which can be identified as having particular features. The first of these is the determination of key questions, which were so important in guiding my interviews and report-writing in the Metropolitan University study.

The use of such questions is discussed above; their formulation will be outlined here. I wrote them completely myself, modifying them from experience but still my words. This need not necessarily be so – the questions could usefully be co-created in the initial design of the evaluation – but in this case it was clear that their principal purpose was for my personal use, despite their usefulness in the interviews.

I wrote the first version of the questions as part of the pro-forma. I modified them further following my initial discussion with the director of the Unit, and then changed them further still over the course of the study. But most of the changes came in the early stages of the interview process: I realised after the first interview that a key issue was, for my understanding, to find out who was working together. (I was later able to use this understanding of who worked with whom to demonstrate how much interaction went on outside the two main groupings.) I also realised that my original thought of a Stakeholder Mapping exercise was not feasible given the constraints of the place, so incorporated the question of stakeholders into the issues to consider with members of the Unit. Finally, I realised from my discussions with the Unit's director that email was also a co-operation technology within the Unit to which some consideration would need to be given.

The final version of the key questions used (for comparison with the version in the pro-forma) was as follows:

- Who do you co-operate with, and how?
- What kinds of co-operation go on?
- Who are the key stakeholders of co-operation within the Unit? To what extent are their stakes being recognised by the current means of co-operation?
- What's the role of email, the common filestore (I: drive) and Web pages?

- Do they help or hinder co-operation?
- How could the technologies be better used?

4.3.4. Stakeholders

I placed less focus on stakeholders in this study than in some earlier ones, and their generation seemed to be less crucial. It was obvious to everyone who were the key stakeholders in co-operation in the Unit. The list produced for the pro-forma was quite a general one:

> the Unit members; customers; IT support staff as suppliers; rest of the School; rest of the University (e.g. Facilities Directorate)

As mentioned, the question of who were the key stakeholders, and to what extent their stakes were being met, was part of the interview process. There were not many surprises; this may have been because in many interviews it felt like a less significant question to ask in the limited time available, so I confined my question just to showing them the list above and asking whether they thought it seemed right, or whether any stakeholders were missing. Such a closed question is unlikely to elicit much response, so perhaps it is not too surprising that not much arose! The final list that came up, as discussed in the report, were as follows:

- Members of the Unit – one person said that if the Unit (and hence its co-operative processes) doesn't generate income, no-one has jobs. By contrast, a few people mentioned that the rapid growth of the Unit to date (and potential growth in the future) makes it increasingly difficult to respect all the stakes that Unit members have;
- Students on the postgraduate programme [as customers];
- Clients of the research projects and other research/consultancy work [as customers, and also as givers of input to various research projects];
- The rest of the School and the University [as suppliers and colleagues – also, as the Unit seems to be rather profitable, as financial recipients];
- Other higher-education institutions [as competitors but also as collaborators].

Curiously, the IT support staff were not mentioned by anyone. Perhaps this has something to do with the focus of the study being less to do with just the technology than when I drew up the initial list; perhaps because support staff in general were not thought crucial enough to mention; perhaps because they were on my list anyway and so the interviewees thought they weren't worth an additional mention.

An interesting lesson to be drawn from this list is the fact that the members of the Unit appear as a single group. This is very much not realistic in terms of stakes. The director of the Unit has a strong stake different from others, as do a number of other individuals. Not all their stakes seem to be compatible. I did mention this to some extent in the report, saying that, "*I have certainly observed that some individuals have stakes in the Unit which could (and do) come into conflict with other individuals' stakes (as I don't feel that naming any of those people or detailing their stakes are especially helpful or fair, I shan't pursue this question further)*". It might have been more educative not to duck this question, but rather to point out that – from my observations – there were some individuals with a strong view of how things should be and a strong personal investment in what happens.

I did not do so for two reasons: I have understood the Stakeholder Mapping exercise to be essentially reactive, taking the views of participants as the only useful guide to who are the important stakeholders. A piece of learning-for-the-evaluator about this study, then, is that the evaluator's view of who are stakeholders is necessary to be taken into account, especially if they're aware that there are hidden agendas in the situation which might be beneficially brought out into the open. (In the Unit, the equation between the director's benefit and the

Unit's benefit is often made; this does not seem justified from my experience.) The second reason not to discuss the question was political. I already believed myself to be taking quite a risk in what I was saying, and felt there were things they could usefully do which I had already outlined, and which would be jeopardised if the members of the Unit were antagonised by my discussion of the side of the Unit that they preferred to keep unspoken. In this, I was perhaps wrong, but it felt potent at the time; perhaps it is also an illustration of the extent to which anyone studying a situation can readily be caught up in it.

4.3.5.   Type and purpose

This study fits in the types discussed in section 2.2.1 as mostly one of Effects. That is, I was looking at the way co-operation actually took place now in the Unit, and the effects of the way the co-operation was constituted (including technology) upon the work of the Unit. To some extent, I was also looking at the way things might be (that is, Formative): I made some suggestions about new technology, use of technology and other ways of co-operating. Little reference was made to the Unit's objectives, either current or potential: the evaluation was of the things themselves, rather than the purposes behind those. (An objectives-based study might have considered the motives for the technology, the history of the ad-hoc co-operation, and other 'why' questions.)

In terms of the classical divide of evaluation theory, the study had definitely a formative aim rather than a summative one: the reason why I was there was to help the members of the Unit change the way they worked together (via technology and in other ways), not to determine for someone else's purposes how good or bad their co-operation was. The purpose of the study, then, was organisational and technological change: looking at the way they were working, and especially how they used their I: drive, and then giving advice on how it could be better. In my email message to the members of the Unit introducing the study, I wrote: "*I would therefore like to study co-operation within the Unit, how computers fit in, and whether it might be improved: do the computer technologies used by the Unit help or hinder that co-operation? Over the next few weeks I would like to assist you all in considering these questions.*" That, then, was the primary purpose from the Unit's point of view. It was fairly clear that at least three of its members had hidden purposes for the study – different kinds of change which they hoped would be supported by my 'White-Coated' presence (the director and the two who had initially discussed the research with me). Because most of the other team members knew of these agenda, I was fairly careful to make it explicit that I was working alone and would be influenced by each person as much as any other in drawing conclusions. (I also tried to make it clear throughout that my views were only personal observations anyway, no more or less immediately valid than anyone else's views.) From my angle, as outlined in the initial email, part of the purpose was that having developed SESL in theory and through a number of earlier studies, I wanted to use the methodology as a whole as part of an 'evaluation of the evaluation'. What I learned is detailed below in the section on learning, and in chapter 5. There were therefore many purposes and many hopes for the evaluation.

4.3.6.   Systems and technologies

Again, it is important to emphasise that there are not such things as systems in the real world; rather, they are devices which are constructed on an ad-hoc basis to make sense of the world. Thus in my understanding of co-operation-within-the-Unit, I was quite comfortable in changing – widening my view – quite widely.

In the first place, the problem presented to me was with the PC LAN filestore, which was disorganised and hard to navigate around. For other discussions, it seemed like a possibility that the general culture of the Unit was to 'blame', but in the first place, the focus was clearly

on the filestore. I then widened this view slightly, to encompass the other co-operation technologies in use in the Unit – email and the Web. However, from the first couple of interviews it was quite clear that, while the technology made a difference, it was so affected by the individualism (among other factors) of the Unit's culture, that the wider question of co-operation must be considered. Thus, in the final report, 5 pages were devoted to discussing co-operation in general, and 3 pages to the use of the technology; although my eventual recommendations split ¾ of a page to the general issues and 2 pages to the technology. There was no hidden agenda in this shift, just an openness to what were the important aspects of the co-operative systems – and it was clear that these were they. To another evaluator, with a stronger base in technology, management or the social sciences, different priorities would have arisen. Nevertheless, this openness will always be crucial.

One part of the system under study that did not change was that I remained focused on the Unit as such, and that as a unitary entity. In the discussion of relevant stakeholders above, it is mentioned that there are a number of groups with a stake in the co-operation going on within the Unit, who might be said to be part of the system. Members of the Unit lecture elsewhere in the University; they participate in decision-making in the School of which the Unit is part; they work with others on research projects (in their research projects, for example, their clients also act as co-participants in an ongoing cycle of action research); and they have many interactions of diverse kinds with students and other customers. It would have been a valid choice, then, to regard the-system-of-co-operation-around-the-Unit as a much wider one (in which the role of co-operation technologies would have been seen much more closely – email is used commonly with other members of the University, fax and phone with 'outsiders'). I did not make such a choice, because my initial remit was to look at *internal* co-operation, and the widening did not seem to add significantly to that (as distinct from the widening from technology alone to other forms of co-operation); of course, the amount of time available played a strong role in this as well.

4.3.7. Different forms of learning

Who learned from the evaluation of co-operation in the Unit? What did they learn? Some of the learning that went on (under the headings given in the discussion of SESL at section 4.2.1.5 above) included:

- Stakeholder learning: The stakeholders in the evaluation (as opposed to stakeholders in co-operation within the Unit, detailed above) comprised: its director, the two members with particularly strong views on co-operation, others members of the Unit split into two main groupings (research and teaching), and myself as evaluator. For myself, I learned that the pro-forma model of SESL can be very useful as a short guide through the evaluation; I also learned that the basic description of the methodology (as found in 4.2.1) is much more limited in its usefulness. The current descriptions of the methodology reflect that learning.

- Ongoing learning: My presence there seems to have made some difference as the evaluation proceeded – they seemed more attuned to what was good or bad about their co-operation, and by the time of the team meeting, the Unit administrator and another team member had already begun to reorganise parts of the filestore. I provided little 'technical support', as my time with them was short, but in one interview with the Website maintainer, I did suggest that she explicitly add their pages to the Web search engines.

- Formal reports: I gave a report to the Unit in two forms: a talk and, a few days later, a written report (the latter both in printed form and by e-mail). The talk happened at a team meeting at a time of great change in the Unit: at the same meeting, they discussed their

move to a new office separate from any other parts of the University (giving a real chance to make a separate culture for themselves) and the director of the Unit announced that he had decided to go to work part-time in industry and only remain part-time in the Unit. It was therefore an especially opportune time for me to be at the meeting, which was itself something of an innovation, only the second in a new series of team meetings (following the urgings of some team members – expressed in my interviews as well – for some more occasions of formal co-operation in the Unit). I gave a talk for about half-an-hour on my findings and recommendations, supported by two flip-chart sheets (these can be found as the last page of the report, Appendix F). This presentation really seemed to touch a nerve among those present, as it led into a great deal of discussion on each point as it was raised. As for the written report, aspects of its analytic methods are discussed above. There was little direct response to the report (except a few people saying they liked it), but it seems to have added to their changing understanding of co-operation together with the other forms of learning.

5. Evaluating the evaluation

As this is a thesis about evaluation, the question will naturally arise: how does one evaluate the methodology presented in the thesis? This chapter seeks to evaluate SESL – and hence demonstrate it to be a useful methodology for the evaluation of co-operative systems – in two ways. First, it looks at the *practice* of using SESL: drawing on the studies presented in chapters 3 and 4, it discusses how well the methodology works in real evaluation settings. Second, it uses a set of criteria for a good evaluation method/methodology, examines the extent to which SESL fulfils those criteria, and compares it to other evaluation methods.

5.1. *Practical evaluation: implications from the various studies*

Given that the methodology discussed here is based on a broadly interpretive paradigm (Burrell and Morgan, 1979), with a research style that stresses uniqueness and context rather than repeatability and precision, it is difficult to measure the exact effect of SESL upon the studies described earlier. The changes which I observed are therefore difficult to attribute to the use of SESL rather than the nature of the situation or my personal skills as an evaluator. One answer to these issues in general is the argument that the lessons learned for the stakeholders during the evaluation arose during the use of SESL, and that this is the main intention of SESL, so that any lessons which arose can be said to be fulfilling the methodology.

A more specific answer to the two alternative causes for the changes found in the studies can also be made. In the case of the nature of the situation, we can use the technique of *triangulation* (Stake, 1995), which says that if a similar result is found in a number of case studies, then they must be linked in some way; and thus in the discussion below of lessons arising from the methodology only issues which came up in a number of studies are used. Second, to answer the question that perhaps only its originator can use it, this sub-section ends with a discussion of a situation where the methodology has been used other than by myself, and the experiences found there.

The following issues arose for the development of the SESL methodology during the studies described in chapters 3 – that is, prior to the existence of SESL in its 'mature' form. In each case, the issues were seen in more than one study; in each case their lessons have been incorporated into the current version of SESL.

- The *limits to ethnography*: it doesn't always show what participants report to be the most important things.

- The importance of *multiple perspectives*, and especially more than just the evaluators and users, but also of all stakeholders. There were radically different views about the system.

- The concept of *multiplicity of discipline / a systemic perspective* – the importance of not getting 'hooked' on a particular analytic framework. The different results that can be obtained from the use of various data-collection methods. The key issues aren't necessarily the technical ones.

- The importance of *learning* itself. The different kinds of learning arising during the project changed according to time, according to the various stakeholders, and according to the activities conducted with them.

- I was concerned about the problem of seeing evaluation as a process of making judgements in a vacuum, of forming an expert opinion; and the concept of *organisational learning* fit nicely as an alternative paradigm for the methodology.

- My role (as evaluator) was frequently not just one of passive observer. I played a part in making an active contribution to the organisation.

The following lessons were learned during the use of SESL in its 'mature' form, in the study of the Unit described in chapter 4:

- the value of the pro-forma as a structuring device – to start off the evaluation, to conduct interviews, and to measure the changing nature of the evaluation as it goes on;

- interviews conducted in public places give a strangely different angle on the situation to those conducted in private… sometimes a group is so fragmented that group exercises don't work;

- sometimes it's obvious who the key stakeholders are, and you don't need to ask;

- sometimes it's more helpful to an evaluation if one is prepared to make observations which cut across their taboos, and thus are painful for them to hear and for the evaluator to say. The phrase *resistance to change* is sometimes misused (by managers whose interests collide with those of their employees), but there are often unmentionable topics in organisations which stop useful change occurring;

- the evaluator's view of who are stakeholders is still worth taking into account, especially if they're aware that there are hidden agendas in the situation which might be beneficially brought out into the open;

- it just happens sometimes that your evaluation is heard at the right time for the organisation, when they're going through a period of change anyway in ways unrelated to what you were evaluating.

As discussed above in section 4.2.4, one expression of SESL is as a set of guidelines. These were presented by myself and Fides Matzdorf at a workshop of the Higher Education Research and Application Forum run by the Facilities Management Graduate Centre, Sheffield Hallam University; held at Bradford University on 26 May 1998. The response was mixed. The delegates liked the second half of the guidelines, which present sets of questions on stakeholders, key issues and reporting – these they found helpful and useful. They complained about the first half of the guidelines, containing introductory material and questions on the purpose of the evaluation and the nature of the system under study – these they found overly academic in their tone and inappropriate for their needs[10].

Finally, let us examine the effects which the methodology (in its developing or finished forms) had upon the various evaluations conducted:

1) *BigBank Financial Services*: There was little resemblance to SESL in its finished form in this study, which was mostly a conventional quick ethnography. I certainly learned more by a combination of methods – observation, interviews and secondary sources – than just performing interviews. The effect of my being there seemed to be at least one of making the group I studied think about their working processes; and, through the report I wrote for BigBank Financial Services' I.T. department, influenced the thinking of the reengineering project to some extent.

---

[10] It is worth noting in this context the fairly adversarial relationship between most university facilities managers and academics – their customers but members of a very different world. This may have coloured their reactions to the guidelines, as may the setting – wet, cold and uninspiring.

2) *ShrEdit*: As discussed in section 3.2, the main contributions to SESL of the ShrEdit study, and the subsequent PETRA methodology, were the twin concepts of multiplicity of discipline and multiplicity of perspective. Both added a richer mix to the study than would otherwise have been possible. The first ensured that we didn't just work from one theoretical angle, becoming trapped by the things it allowed to be seen or not seen, but rather used a variety of approaches. The second worked similarly but with the views of people – it gave voice both to the evaluators and to the users of the system, enriching our views on how well it worked considerably.

3) *Health Centre*: Here the learning for SESL (see the end of section 3.3) was mainly in content terms rather than methodological – issues of the inter-relatedness of computer systems with other media and the fact that benefits don't always accrue to those doing the work were especially important. This study was influenced by an awareness of these issues as important ones, though, as they had arisen in the previous two studies to some extent.

4) *University MIS*: The developing importance of a stakeholder perspective in SESL was important to the nature of this study. It was quite clear that the different groups within the university had different views of what made it good or bad, and an orientation to this was crucial. The fact that the problems of user acceptance are as much to do with political issues of the way the system has been installed as the technology – an issue of systems thinking – was also important here.

5) *RICS*: The stakeholder perspective was also important here in making the team aware of the expectations different groups had of them. The fact that the evaluation was oriented towards ongoing learning, of different kinds, throughout the evaluation process, was also crucial – I have discussed above the 'human reminder' effect, and this was especially to the fore in the RICS study.

6) *The Research Unit*: Again, the study was much richer for its use of stakeholder thinking – the awareness here that the views of different members of the team on co-operation within the unit were so different as to affect the outcome significantly. This also linked in with the heterogeneous learning concept, and the various kinds of learning going on as the evaluation progressed. The formal framework of SESL, especially in its pro-forma state, was invaluable in this study for guiding the initial discussions I had with the unit members as to what the evaluation should consist of; and it served as a constant reminder of the changing purpose through the interviews. In this study, SESL also enabled a much richer view of the effects of other kinds of co-operation on their computer-mediated co-operation than would have been possible using a methodology that emphasised the whole system less.

5.2.   Theoretical evaluation: criteria for a good method/methodology

This section moves on to compare SESL to other evaluation methodologies in a more 'objective' fashion – based on a common set of criteria not selected to favour one methodology or another. The following criteria have been selected based on my experience of what is useful to look at with evaluation methodologies – they form a kind of 'buyer's guide' of the sort one might find in consumer reports (recommended as a form of evaluation by Scriven, 1996). The concerns expressed in them have nothing to do with the direct concerns of this research. The criteria are discuss below, and then a comparison of SESL against the methods and methodologies found in section 2.1 is presented in tabular form. There are ten criteria in all, which can be conveniently divided into being concerned with issues of

theoretical interest or practical ones. Of course, the criteria all inter-relate to some extent, and so the divide between them seems artificial at times – nonetheless, it is heuristically useful.

Criteria concerned with theoretical interest include:

1. *Claimed Benefits*: What does the method have to say for itself? What do its main practitioners claim are the benefits for the method? If the method is followed in its fullest form, what results are intended to follow?

2. *Scope*: Linked to the previous criterion, how wide is the method intended to be? It may be that a particular method only really applies to the technical parts of a system, or only to its social aspects . What limits have been set upon it by its designers? Specifically, how is I.T. treated in the method – as a key resource, as the only resource, or as one of many different parts of a system?

3. *Supporting Theory*: What is the theoretical basis on which the method rests? From what discipline or body of ideas does it derive its concepts? How do its practitioners analyse the data it collects?

4. *Attitude to People*: How does the method treat the people that it is studying? Are they just ciphers or are they regarded as important in themselves? Are their views taken into account in the forming of the evaluation?

Practical criteria include:

5. *Useful in Practice*: How has the method worked out as it has been used? What results have occurred?

6. *Usable*: Is the method quickly and easily comprehensible? Do the available instructions make sense? How much of an expert in the underlying theory must one be before one can use the method?

7. *Timescale*: How long does the method take before it produces useful results?

8. *Resources*: What material resources does the method require to work properly? Presumably it will need some human resources also – are these large or small in terms of time required? And what input is needed from other people?

9. *Adaptable*: How readily can the method be adapted from the purpose for which it was originally intended to some new purpose? Can it be sensibly mixed with other methods, especially those from other disciplinary sources?

10. *Experience*: What is the collected world experience with this method? Is this a recently introduced technique which has only been tried by a small group, or is it widely tested, with lots of available information about the things to do or to avoid?

Against these criteria are evaluated the five broad classes of evaluation methods which formed the subsections of section 2.1: ethnography, other qualitative methods, psychological methods, systems building methods, taking advice; and three methodologies also discussed there: PETRA, SSM and SESL. The differences between the issues involved with various of the methods found within these groups have been somewhat elided (for example, the laboratory experiments and quasi-mathematical analyses categorised under the term 'psychological methods' are in some ways rather different from each other); but under the criteria used they are mostly similar within the groups.

5.2.1 Theoretical Criteria

Class of methods	Claimed Benefits	Scope	Supporting Theory	Attitude to people
Ethnography	Getting to the truth of an evaluation, by describing it in detail. Taking context into account.	Whole system. Regards I.T. as one resource among many.	Varies (see § 2.1). Different views of relationship between data and interpretation. Use of theory often seems to be heuristic rather than normative.	Can be conservative and non-interventionist: basic orientation to the status quo[a]. Often more interested in detail than the Big Picture.
Other qualitative methods	Gaining an understanding of users' own experiences with a system, either on a large-scale or an individual basis.	Mostly, the whole system – tends not to distinguish between experiences with technology and general organisational life.	Fairly atheoretical – often carried out with an interpretive paradigm in mind, but can take a positivist turn (e.g. large scale questionnaires).	Regards peoples' views as main data sources. Sometimes a risk that they ignore the person beyond the specific research questions.
'Psychological methods'	Understanding the 'real truth' behind a situation. Most users believe that by filtering out the details outside of their specific question, they can discover what's really going on.	Mainly focused on I.T. – regards others aspects of system as unnecessary detail[b].	Typically well developed and well grounded. Most users share an orientation towards the reductionist paradigm, and most tend to regard context as irrelevant.	Tends to reduce people and their interactions to numbers, ability to do certain tasks, or 'lines and boxes'.
'Systems building methods'	To understand how well a system works by building it. PD methods also aim for increasing industrial democracy.	Mainly focused on I.T., but other aspects may also be of interest insofar as they affect the technology.	Mostly fairly atheoretical. PD methods sometimes have a political basis. Often some relationship with a social science (e.g. psychology for the user-centred methods)..	Fairly respectful of people's views in most forms; PD methods regard user involvement as crucial. Can be slightly patronising[c].
Taking advice	Learn from the experts; uses objective criteria to compare one system with another; decide sensibly which is the best system to buy	Entirely focused on technology, but looks at technology in (potential) use rather than for its own sake.[d]	Atheoretical, although consumer reports will have some sort of basis to their evaluation criteria.	Market-driven, so regards people as capitalism tends to – mostly rational actors in an economically-focused organisation
PETRA	"incorporating a theoretically-driven evaluators' perspective to investigate the collaborative activity, and a design-based, user-focused participants' perspective to evaluate the	I.T. focused, especially in redesign session; also takes into account how users interact with system and their co-operative behaviour	Conflict between two basic philosophies (PD and analysis) is ideally resolved dialectically.	PD side: people's views are most important. Analytic side: their actions are more important.

"supporting tool" (Ross et al. 1995)			
SSM	Encourages a process of whole-systems change through a collaborative defining of the nature of the system and of the problems with it.	Whole system, but views a system as an analytic resource rather than a thing as such; so that how widely an evaluation is drawn depends on user.	Based on 'soft systems thinking': whole system with a phenomenological paradigm. Increasingly, some awareness of political aspects.	People's perceptions seen as important in the development of the nature of the evaluation as well as data.
SESL	Emphasises organisational learning on a number of different levels. Based on needs and views of all stakeholders. Looks at whole system rather than one aspect.	By design, looks at whole system rather than particular aspects. Again, view of what is the system is very much dependent on the user of the methodology.	Systems thinking is at the heart of much of the methodology (systemic perspective, stakeholder approach, learning perspective). Most theories applicable to an ethnographic context can be of use here.	Conflict between different perspectives resolved by participants' own negotiations rather than directly via the methodology. Tends to respect their viewpoints, but only insofar as they are individuals' views, not as definite truth.

Notes on this table:

a. The basic orientation of ethnography to the status quo arises from its focus on description rather than prescription, on examining the way things are rather than changing them. Ethnomethodology in particular has tended to see its main task as reflecting back the organisation to itself. Ramage (1996) discusses an alternative orientation to ethnography, based on an interventionist perspective.

b. It is in the paradigmatic nature of psychological experiments methods that they only look at part of the situation under study, regarding the rest as detail which detracts from the main hypothesis. In the case of CSCW evaluation, they thus tend to ignore much of the context of use. This makes their applicability limited only to questions of narrow detail, although where these are concerned – particularly concerning usability – they have shown much commercial application.

c. Advocates of the user-testing school of systems design (e.g. Tognazzini, 1992) tend to write from a clear perspective of us-and-them: the knowledgeable designer and the innocent user. Under this perspective, the designer has a clear responsibility to study what users do, but this responsibility is framed from an angle of superiority.

d. The approach of giving advice tends to mainly look at the technology as a thing in itself, and to assume a deterministic relationship between its implementation and the effects caused by it to the organisation. It does however, particularly in the form of market analysis, often give a view on the wider super-organisational context that is missing from any other perspective.

5.2.2. Practical Criteria

Class of methods	Useful in Practice	Usable	Timescale	Resources	Adaptable	Experience
Ethnography	Very useful on context/collaboration. Frameworks tend to lead to 'blinkers' problem[a]. Issues of time & resource constraints.	Obtained by osmosis (the methodology) and deep reading (the frameworks). Fairly natural but also quite elusive.	Fairly long-term. Ideally months, but increasingly a week or so seems to be common.	Low: a tape-recorder and a notebook. Significant amounts of time required of busy people, though.	In terms of available analytic frameworks, highly. Also links in well with other methods, especially other qualitative ones.	A fair amount, especially within CSCW. Evaluation often seems to be a side-effect rather than a direct intention[b].
Other qualitative methods	Generally, most useful: quick and easy way to gain lots of basic data. It is questionable how rich and how accurate the data are[c], so additional methods are useful.	Highly – intuitively obvious (asking people questions), but lots of guides to assist understanding of more detailed issues.	Varies, but generally not very time consuming. A useful set of interviews can readily be carried out in a day or two.	Low: a tape-recorder and notebook, plus perhaps questionnaire resources. Less costly in overall time than ethnography, but more immediately intrusive.	Can be used in conjunction with almost any other method, and often useful as a secondary check.	Considerable, in many different areas of life. Not just confined to academia – helpful insights from market research and journalism.
Psychological	Splendid for single-user issues & those fairly divorced from work contexts. Fast, easy to do & analyse. Collaborative settings harder but possible.	Fairly usable if one has a strong background in maths or psychology. Mostly well documented.	Depends on the detail of the evaluation. A small issue can be covered in an afternoon, wider issues can take much longer.	Usability labs of some form are typical. "Budget" methods can be done with one computer. Analytic methods just need experts.	Other theories than cognitive psychology can be 'plugged in' – e.g. methods from participatory design.	Plenty of well-grounded knowledge, lots of theories, lots of experience. Used commercially.
Systems building methods	Highly pragmatic. Breaks down barrier between design & evaluation, so that the evaluation comes out of the redesign. Most appropriate with bespoke development.	There are lots of methods about, and most are fairly well detailed. May be a problem in being allowed to do it	Fairly long-term in principle, but in practice all kinds of modified methods make it available in an afternoon.	Relatively low-key: people (evaluators and potential users) and cardboard. Does tend to take up time of key workers, though..	Highly. Many available methods, and more formal approaches can be plugged into the basic framework.	Large amounts for design; less for evaluation but an increasing amount.
Taking advice	For deciding which system to buy, fairly so. Need to be careful about objectivity of	No special knowledge required – mostly provided by experts – although the results can	Time to understand report and relate it to one's own situation.	Only the reports themselves. If produced on a profit-making basis, can be	The results can be taken only partly into account by readers.	Fair amount in commercial settings – many consultancies work in this way.

	approach, though.	be jargon laden.		very expensive.		
PETRA	Explicit aim to avoid blinkers problem. Very good for studying semi-situated collaboration. Relatively easy.	Reasonably coherent guide in paper. For some people, what exactly the method entails – where it begins and ends, what is part of the method as such – is unclear.	Around an afternoon for the studies, plus a further time for analysis.	Computers to see system in use; coloured pens, paper etc for redesign sessions.	Highly; intended to be tailored to fit in preferred theoretical frameworks. Basic dual-perspective approach non-negotiable, though.	In this form, one study only, but it's solidly based on other methods which are well grounded.
SSM	Very useful – gives lots of valuable, practice-oriented information.	Practitioners tend to swear by it. Many books and papers written on subject. Basic methodology sometimes feels hard to track down.	Varies, but generally a few weeks for a full evaluation.	People, and a fair amount of time allocated by participants.	Intended to be adapted (definitely sees itself as a methodology rather than a method). Has changed considerably over its history.	Considerable. Thirty years experience of the method in various forms, now quite widespread.
SESL	Use of many viewpoints. Explicit consideration of multiple uses. Raises issues normally kept hidden.	Guides available at various levels for quicker or slower versions.	Varies, according to which parts of the methodology are used: from two hours to a few weeks (or longer-term).	Generally low-key: evaluator plus their notebook. May require reasonable amount of time from participants.	Designed to be adaptable according to situation. In my use, has often been tailored considerably.	As reported in this thesis. Mainly by myself, although versions of the methodology have been used by others.

Notes on this table:

a. The problem of 'disciplinary blinkers' – that many analytic approaches only look at certain small areas, and ignore the rest, is discussed in Ross et al (1995) / Appendix C. It is present as a risk for many of the methodologies discussed here; it is particularly noticeable in ethnography because of the discrepancy between a *data-collection* approach that is most subtle and all-encompassing, and the risks of an *analysis* approach that ignores many of the data available.

b. However, Randall et al (1996) and Hughes et al (1994) both discuss the applicability of ethnomethodologically-informed ethnography to evaluation.

c. The question of accuracy of qualitative methods is often put in terms of the fact that they rely on reports by people of their own experiences. However, many people rewrite those experiences to fit their own expectations and beliefs, either consciously or unconsciously. This argument leads some to suppose that observation is more 'reliable'.

5.3   Summary

This chapter has analysed the effectiveness of SESL. This has been done in two ways. First, the lessons to be drawn from the six case studies have been presented – the issues which arose in more than one study and which have contributed to the principles of the SESL methodology; and the effect of the methodology in its developing form upon the various studies. A number of important ways in which SESL made such a contribution were raised.

Second, an evaluation of a more traditional form was attempted: a set of criteria for what makes a good evaluation methodology were presented and justified, and SESL was judged on these categories against a number of other methods and methodologies which have been used for evaluation. SESL came out fairly well on these criteria, although for some of them the value of other methodologies can be seen more clearly; and the tables make it very clear that the advantages of one methodology over another will very much depend on the purpose to which it is put.

Rather more in keeping with the SESL's view of evaluation as an ongoing process of learning, two of the sections in the following chapter also take on an evaluative role. Section 6.2 discusses the future work which might go into the development of SESL; by implication, this demonstrates gaps in the current version of the methodology. Section 6.3, which closes the thesis, tells the story of SESL's development, and how my various experiences in conducting case studies and reflecting on the issues involved changed the methodology over time; as it is about journeying and ongoing learning, it is perhaps the only consistent way properly to evaluate SESL.

6.   Conclusions

SESL is a developing methodology, and will continue to develop further with each new piece of evaluation to which it is applied. Thus, this conclusion should not be seen as an end point, merely as the end of a stage or phase in this process. It will discuss the principal benefits of SESL, and then describe something of the journey involved in developing SESL, by way of a summary of the main issues involved here, and as a look to the future.

6.1.   *Restating the case for SESL*

This section discusses the characteristics of SESL and what it offers to a practising researcher, and can thus be seen as a sort of advertisement for SESL.

The *need* for SESL has been demonstrated above in both theoretical and practical ways. The key flaws of existing evaluation methods are all to do with their single-mindedness: they privilege the viewpoint of *one* stakeholder group over others (either explicitly or implicitly), they ignore the whole system and instead concentrate on the perspectives available from *one* disciplinary viewpoint, they focus on right and wrong. SESL is different: it works to evaluate the whole system, even the bits that are uncomfortable or seldom looked at, taking into account the views and needs of all stakeholders, and making evaluation part of an ongoing learning and change process.

The following features of SESL's nature make it distinctive:

- SESL covers the whole system, rather than just parts of it. Solutions to a problem within a narrow (sub-)system cannot occur from within it, but rather need to arise from outside it. In particular, it makes little sense just to focus on the technological aspects of a system, as the use of the technology is bound up with a range of issues including organisational politics, culture and the work people do.

- SESL 'helps the parts other methods don't reach'. Its whole-system and politically sensitive orientation helps to uncover hidden corners. It raises the awareness of organisational stakeholders for deeper issues than just the technical ones, thus enabling a process of deutero-learning[11]. Thus, instead of looking at superficial detail, SESL aims to look at underlying causes; it is able to deal with taboos, issues not normally allowed to be voiced, but which affect the way that people work together and use technology.

- SESL aims to work with an organisation's current learning, not against it. That is, it facilitates evaluation rather than dictating its results; it supports and enables the learning processes of an organisation. By building a climate of learning, SESL also strengthens and prepares the organisation for future change, through involving stakeholders, encouraging participation and ownership and providing motivation for ongoing evaluation as part of an overall strategy of organisational growth.

- SESL's unusual perspective brings refreshment and release from the more traditional scientific evaluations. Its more varied and interesting approach can help to revitalise an organisation, since it encourages excitement, experimentation and play. We can only learn through taking risks and trying things. SESL encourages people to try out new ideas, new ways of understanding their organisation, and new ways of working.

---

[11] That is, meta-learning, learning how to learn (see section 2.3.2).

- SESL looks to satisfy all stakeholders, bringing in many voices. It encourages 'buy-in' and support from a wide range of people. It respects all points of view, going beyond the adversarial thinking of "I am right, you are wrong"[12] to a collaborative thinking that recognises the existence of shared problems, needing shared solutions.

- SESL is based not on unchangeable rules, but rather on one core principle: dare to question (*sapere aude* – Kant, 1990/1784). Its recommendations for conducting evaluation, and the solutions it provides, are written on 'tablets of liquid' rather than tablets of stone – it is flexible and adaptable, rather than static. It is therefore expandable, making it more useful for future extensions.

6.2. Future work

From the previous section and from the very nature of SESL, it follows that it is an ongoing process rather than a finished model. There are a number of areas which to explore would have been beyond the scope of this thesis. Areas where future research is needed include:

- Making the steps involved more explicit as a 'simplified form'. Presently, the methodology can partly be understood from a short guide like that in appendix A (discussed in section 4.2.4). However, in its present form, it is best used either in a workshop-type setting or as an introduction to the main ideas in the thesis. A stand-alone guide, deriving from SESL, would be invaluable. The current guide serves well as an early version of such a guide, but need to be developed further through user testing. Presently it has had some exposure to others than myself, both in its current form as a set of questions and in an earlier form as a set of instructions. It has been rewritten to include some of that feedback; but it needs more. In particular, it has not been used yet by others as a guide while performing an evaluation, and modified in the light of those experiences. From this process, it will also become clear which aspects of the methodology are 'core' and which are 'optional extras'.

- The status of the methodology with respect to other types of evaluation than Effects (see section 2.2.1 on types of evaluation) needs clarification. It is clear that my main understanding of the methodology, and main fieldwork, has been in connection with situations of technology-in-use. However, there are other settings where it would be highly valuable to be able to perform multi-perspective, systemic evaluation. This especially applies to formative evaluations of systems to be put in place in real settings (either known settings, in an implementation situation, or potential settings, during software design); and to the evaluation of research projects, for example by PhD students. The latter often have difficulty in knowing how to go about evaluation. Widening SESL to consider these types of evaluation explicitly – and conducting further evaluations of these sort making use of the methodology – would be useful.

- Further explicit development on the learning outcomes for system designers would be useful. Presently it is clear that they will be part of the general feedback loop formed by the learning processes; but their special concern of how to harness this learning to build systems that take into account all stakeholders' needs is not explicitly addressed. There is a certain similarity with the gap between the rich and detailed ethnographic accounts of work and the rather less useful "implications for systems design" which often follow them

---

[12] Cf. de Bono (1991), who writes of 'water logic' – fluid, quick to transform, difficult to stop; instead of 'rock logic' – hard to change or move.

(cf. Plowman et al, 1995). Finding a 'bridge' for that gap has long been a keen aspiration of ethnographers within CSCW, not least at Lancaster, but it is a difficult problem of reconciling inherently unformalisable people with necessarily formalised technology. There is no immediate reason to suppose the bridge is any less difficult to find in a SESL-based evaluation.

- Identifying (not to speak of measuring) the learning outcomes of a SESL evaluation is likely to be difficult in many situations. It has certainly been hard in the discussions of the studies in this thesis. This is especially so because of the fluidity of the concept of learning used here, which includes any kind of information, skills, fresh awareness, change of perspective or new ideas gained during any stage of the evaluation process, either by individuals or by the organisation as a whole. Practically speaking, this means that a manager wanting to ask the question, "what has SESL done for me?" is faced with a hard task. Can the benefits of learning through SESL be measured? Is it meaningful to try to do so? This is an open question; it may be possible to devise metrics to measure relevant learning (or to use existing ones, such as the Learning Company Questionnaire that derives from the research recounted in Pedler et al, 1991); but it is not clear whether such metrics have any meaning here. It is similarly difficult to quantify the benefits of organisational learning – although there are performance indicators (e.g. short-term absenteeism) which are closely linked to employee motivation, which in turn influences productivity (Clark and Matzdorf, 1998). Again, more research is necessary, perhaps by trialing and modifying metrics that measure organisational learning.

- Related to this is the whole question of explicit learning outcomes for managers, especially concerning the extent to which SESL is 'proactive'. Research has shown that the average large company has a life-span of forty years (de Geus, 1997) – how is it possible to ensure 'sustainable development' in the life of a company, to make it grow in a steady way that can be met in the future? It may be that SESL could provide part of the answer within its scope, given the focus on continuous learning that it encompasses; but the extent to which this is possible needs considerable further work.

- The links of the methodology with other approaches could usefully be extended. It bears some resemblance to Soft Systems Methodology (Checkland and Scholes, 1990), and this link could be further developed. In particular, there are subtleties contained in the extended form of the methodology (section 4.2.2) which could be developed further, especially with respect to the issues of feedback which are mentioned in chapter 4. There is also considerable influence from the work of certain evaluation researchers, especially those at the Tavistock Institute (Cullen et al, 1993), and Patton (1981); and again these links could be extended. Some of the work on evaluation within the field of information systems may also prove useful.

- The account here takes less account of current organisational practice in evaluation than might be appropriate. A highly interesting piece of comparative research would be to look at how a number of organisations currently go about the evaluation of their co-operative systems (whether using computers or not, and whether or not they think of it as evaluation); and to compare it to the concerns of SESL discussed here. Such research has generally been difficult to carry out (hence its absence here) due to the sensitivities of organisations to having their key secrets revealed, and telling them that the purpose of the evaluation was to see how far their approaches fell short of SESL would probably exacerbate such sensitivities. However, were it possible there is little doubt that the development of SESL would be considerably enhanced, through the incorporation of best practice from such studies and in reaction to less useful evaluation activities.

6.3. Lessons from the journey

This thesis represents the culmination of a long intellectual and personal journey. The material earlier has principally described the results of that journey. The following section, by contrast, recounts something of the process of that journey, of the personal learning which has gone on.

The work began fairly narrowly focused, on CSCW evaluation. It was stimulated initially by a term essay during my MSc, on the evaluation of the ShrEdit groupware tool (see section 3.2 and appendix C). I was clear what evaluation meant – it meant usability evaluation, deciding whether a computer system was easy to use. Many well-designed techniques exist for this from HCI. But it quickly became apparent that evaluation of multi-user systems would be different, because how one person used a system would alter greatly others' experiences with it (an idea which derived partly from the ShrEdit evaluation and partly from Grudin's (1988) paper on the use of shared calendars). The difference was strong enough that it seemed a topic for further research; but I still perceived CSCW evaluation as usability evaluation plus X, with X to be found.

During the time of this research, both the object of the evaluation and the approach to evaluation widened significantly. It became clear that the idea of just looking at CSCW technology was in itself inadequate – the technology was so bound up with other factors that this made little sense. Thus I adapted the term 'co-operative systems', originally simply a synonym for CSCW technology, to refer to socio-technical systems which involved computers mediating co-operation between people (generally within some organisational setting). It also became clear that given this widening, it made little difference whether the mediating technology was a groupware system, developed with the explicit intention of supporting co-operation, or rather any sort of computer technology which fortuitously supported co-operation[13]. Indeed, the concept of co-operative system can be seen as a paradigm shift from the widespread concept of an information system – while referring to exactly the same mixture of people and technology, by focusing on the relationships between those people, their co-operation, a different form of discussion is possible from that available if one focuses on the results of those relationships, the information they create and use.

My approach to evaluation also changed over this time, again becoming wider. Initially, I assumed evaluation was a matter of deciding which parts of a system were good and which were bad, according to some reasonably clear and objective criteria. (That these criteria should be as wide as possible tied in with the widening of the objective of the evaluation mentioned above.) From the work on PETRA, the importance of more than one perspective on the evaluation became clear, but still within the context of an overall judgement. Then, through my first studies and through reading the evaluation research literature, the concept that different groups and individuals might have different views on what was good or bad arose. So the importance of stakeholders became clear. But how to reconcile the different viewpoints of the stakeholders? At first I believed they were essentially irreconcilable by an outsider, and the evaluator's role was rather to highlight the existence of multiple stakeholders and perhaps facilitate a process of negotiation between them. This still seems a valuable activity. But during the RICS project, I became aware that an issue raised by some of the

---

[13] "Do you want to know how we co-operate using computers in my company?" said a friend, a senior manager in a large engineering company, around this time. "We save our spreadsheets on disk, send the disks through the internal mail, and the recipient then works on the spreadsheet further before sending it on." This clearly had to fit into any definition of co-operative system, yet it was not groupware.

evaluation research writers (especially from the Tavistock – see section 2.3.3.1) might be crucial to co-operative system evaluation – the notion that evaluation could be a process of facilitating multiple learning for the different stakeholders. And given the focus on organisational learning during the RICS evaluation, it was natural to realise that this learning was as much an ongoing *organisational* learning process as an individual one. Thus the shift from evaluation-as-judgement to evaluation-as-learning was complete in my mind.

It turns out, and especially from the evaluation of the research unit detailed in chapter 4, that this approach is particularly strong in coping with highly-politicised organisational environments, where making a simple evaluation judgement would be to favour one side or another. It may well be that a less subtle approach (such as business-process reengineering), which ignores the politics and looks to come to a single conclusion, would be so mistrusted by all sides that it would be ineffective. By being aware of the politics but concentrating on learning outcomes for all parties, the question "what's in it for me?" is answered many times over; and a richer evaluation results.

SESL now exists in its 'mature' form. Yet the journey has hardly begun for the ideas involved in it. Many people have to do evaluation of co-operative systems, for all the reasons mentioned earlier in this work; and it is not clear for many of them how to go about it. So the fact that someone should be working on a methodology for such evaluation is of great interest to many; and I have had much communication with people wanting to use it. In my own next piece of research, I am involved in a project on creating solutions, based in software engineering and organisational analysis, for the evolution of 'legacy systems' (old computer systems that no longer meet the needs of the organisation). The stakeholder and systems perspectives have proved useful in a paper on determining the maintainability of software systems (Ramage and Bennett, 1998); and we hope to make use of the evaluation guidelines (section 4.2.4) in developing later stages of our solutions. The tailorability of SESL means that it is as likely to turn up in the future in a modified form in other places. Other aspects of future work have been considered in the previous section.

At the start of this thesis, the main question of the research was why existing CSCW evaluation methodologies are so problematic, and how they could be better. I have argued, and demonstrated, that it is the lack of a wide focus – on viewpoints, on disciplines, on the object of the evaluation – that is the chief problem; and that the solution lies in building more responsive methodologies to the needs of all stakeholders, in examining the whole socio-technical system involved in co-operation, and in focusing the outcomes of the evaluation on learning rather than judgement. The SESL methodology is a solution to those problems.

Appendix A: Evaluating Management Information Systems[14]

A.0.   Introduction: What is this sheet for?

All organisations collect data and transform them in various ways into information. Some of this is done via computer-based systems, others rely on manual records. Some of the systems will have been developed specifically for the purpose, others will have grown up and been adapted and expanded over the years. Some will be working well, others less so. Some will no longer provide managers with the information they need to plan and run their operations when it is needed.

The term 'MIS' (Management Information Systems) is often used by computer experts to mean 'just' the hardware and software. We don't want to use this meaning here. We find that all organisations have management information systems, whatever they may be called. Not all of them are computerised or designed, planned and implemented – so we use the term in a broad sense to include any system that holds and gives out information that is necessary and/or helpful to manage an organisation. This includes card-files as well as sophisticated software, people and their knowledge as much as technology.

These guidelines are designed to help managers evaluate information systems within their organisations. They should help you to analyse and check the information systems in use to make sure they are achieving their purpose in the most efficient and effective ways possible. This evaluation should also assist you in making well-founded decisions on the way you deploy your information systems and the way you integrate them into your work processes.

There are many existing tools concerned with these kinds of issues (some examples from different sources are Balanced Scorecard, Goal Question Metric and ISO 9000-3). These 'evaluation guidelines' do not aim to add yet another tool to these, but to help managers to set an overall context, to determine the level of detail that they want to explore and to select the right tools.

We have attempted to cover a wide range of evaluation issues. The following lists of questions are not exhaustive, and their relevance will vary from one situation to another. However, our experiences taught us that a broader view is usually more beneficial than a narrower one – and many an organisation can tell a tale of their MIS not serving their organisational strategies and policies precisely because it is symptoms that are treated instead of their underlying causes. (Not to speak of those organisations who *don't* live to tell the tale...)

---

[14] Written with Fides Matzdorf, Sheffield Hallam University, based on ideas in this thesis. It was originally written for a workshop of the Higher Education Research and Application Forum run by the Facilities Management Graduate Centre, Sheffield Hallam University, that was held at Bradford University on 26 May 1998. It has been substantially rewritten since that workshop, taking into account the views of the delegates (see section 5.1).

A.1. *What do you want to get out of your evaluation?*

The purpose of an evaluation to some extent determines its outcome: it will make quite a difference whether it's done because organisational regulations demand one every 3 years, or whether it's done to find out how many new computers you have to buy. So the purpose of the exercise needs to be clear in order to make sure that the appropriate data/information is gathered at the right level. The evaluation will also be affected by the lifetime and the budget of the system under scrutiny. A common rule of thumb is that planning and evaluation should take 10% of time and money. Less money and time means less depth and detail. Another important point is to think about how you can make best use of the results of the evaluation, given the time and effort that go into it.

◆ Why are you doing the evaluation? What reasons have led up to it?

◆ What are you doing the evaluation for? What do you want to get out of it? Will any decisions be based on it? Is the main purpose of the evaluation for example...

  ➢ to make changes to the existing system?

  ➢ to make judgements about the effects of a long-established system?

  ➢ to decide which new software/hardware to buy?

◆ Are you looking at the way things are now (e.g. for assessment/appraisal or to measure past achievements), or their potential for the future (e.g. to set new targets)?

◆ If you have been asked to do the evaluation by someone else, what is their angle on it? Do they need to know certain things, expect certain outcomes or expect certain decisions to be made? What level of seniority in your organisation will the results reach?

◆ Will the main outcome be one of feedback (e.g. benchmarking, performance measurement, …) or decision-making (e.g. restructuring, purchasing new equipment, budget changes, strategic planning…)?

◆ Can the results be beneficial for some or for all stakeholders? How can the information be made available for different groups? How can the results be made available: as a report, presentation, part of a workshop, in the staff magazine, on the Web, …?

◆ How can the results of the evaluation be used to further best practice or be used as part of a benchmarking process?

A.2. What do you want to evaluate?

Do you actually know what precisely you want to look at? This might sound like nit-picking – surely it's obvious – but all too often the obvious answer is unhelpful. For example, if one was evaluating a multi-person calendar program, it might seem that interesting evaluation issues would include the 'user-friendliness' of the software, whether it did all the things one wanted, and whether it was efficient. However, if the level of mutual suspicion in the team using it was such that no-one wanted the boss to know where they'd be, then perfectly good technology might be under-used for reasons unrelated to the software itself.

♦ Are you mainly concerned with technology (hardware, software, networks etc.), mainly concerned with social processes (effects on people and groups) or with a mixture of the two?

♦ Is your chief interest general or specific, broad or narrow – are you looking at one aspect of an MIS system, or several?

♦ Which of the following aspects are relevant to your intended outcomes/purposes: technical performance, appropriate functionality, user-friendliness, meeting standards, effects on individuals, effects on groups/teams/departments, effects on organisations, effects on society…?

♦ In order to identify the scope of the system, it may be helpful to answer the following question. This also helps to identify problem issues, in preparation for section A.4 of these guidelines:

  ➢ Function: What does the system do? What is it supposed to do? Are these questions very different: does the system do more or less than its duty? If there is a discrepancy of this kind, why?

  ➢ Complexity: Is it 'high-tech' or 'low-tech'? Are special skills needed and if so at what level? What is the interaction with other systems (e.g. accounting) and with other media (e.g. post, fax, phone, email)?

A.3. *Who are the stakeholders and what are their interests?*

For an evaluation to be successful and to carry wide credibility it is important that it does not focus solely on **one** viewpoint (e.g. senior management, trade unions, or computer technicians) while disregarding others. If, for example, a board-level decision is made to buy new software without any consideration of the end users' needs or views, this can in the worst case result in a huge waste of money and staff time, with staff busy finding working solutions 'around' the new system.

♦ Who are the stakeholders?

   ➢ Who are the facilities service providers / department?

   ➢ Who are the users/customers/clients of the technology? Who has access to it?

   ➢ Who are the users/customers/clients of facilities supported by the technology?

   ➢ Who pays and for what?

   ➢ Who runs / services the technology?

   ➢ Who has an input?

♦ Who are the **key** stakeholders: whose views need most to be met? Whose involvement is so vital that the system will break down without it? What is the 'right' level of consultation? Who needs to be involved?

♦ What are the system's effects upon these key stakeholders?

♦ What are the interests of the various stakeholders involved? What stake do they have in the system? What would they lose if it no longer existed?

A.4. About the system – getting into the nitty-gritty

Here you need to give consideration to the details of the system. These will vary widely depending on the individual system as well as on the issues you need to consider. So the following questions are not exhaustive.

The following questions are concerned with who works with whom, and how:

♦ Who do the main users co-operate with, and how?

♦ What kinds of co-operation go on?

♦ What are the key MIS technologies, and what is their role?

♦ Do they help or hinder co-operation?

The following questions are concerned with organisational strategy:

♦ How do your MISs fit into the overall management of the organisation? How do they fit into the organisation's strategy? Do they serve it or obstruct it?

♦ What budgetary opportunities and restraints are there?

♦ Is there an overall IS strategy in your organisation? Within your department?

♦ How does this fit into your business plan?

♦ How could the technologies be better used? What obstacles need to be overcome to achieve this?

The answers to sections A.2 and A.4 should provide enough information to decide whether using any specific information or management tools (see examples in introductory paragraph) is appropriate and needed.

Appendix B: The BigBank Study[15]

B.1. Background

BigBank Financial Services (BFS) is a subsidiary of BigBank, dealing with life insurance, investment management, personal taxation, trusts, and executorship. Over the past two years, they have progressively restructured their business and introduced workflow, using the Staffware package.

The primary aim of this project, entitled Meridian and run from the IT department in London, was to cut costs. At the start of the project, in November 1992, BFS's cost-income ratio was 90p in the £ (that is, for every £1 of income, they were spending 90p in operating costs). Their aim was to reduce this to 50p in the £. At July 1994, when I performed this study, they have reduced costs to 55p in the £ – a saving of nearly 40% by their restructuring programme and the installation of workflow.

The restructuring began by centralising the organisation, previously based in 30 regional offices, into two offices based on function (one in Peterborough, dealing with investment management and personal taxation; the other in Manchester, dealing with the other services). This was used as a way to cut support stuff and middle management; but also to cut administrators – in total half of all administrators.

An important part of the restructuring was to stop the administrators performing menial tasks of data-entry, which with a workflow system will clearly be increased – BFS are not currently using document image-processing (DIP). Therefore there are data-entry groups called "Group Support Units" (GSUs) within each division. The aim of this is to "right-skill" administrators, rather than deskilling them (Alan Mabbs[16], pers. comm.), making them into professional administrators of the appropriate area rather than just clerks.

I visited the Peterborough site, and mainly spent my time in the Personal Tax division. This division is composed of eleven 'mainline' groups: eight based on geographical areas, one dealing with new business, one dealing with overseas business, and one dealing with specialist cases. There are also three Group Support Units in this division.

The central part of the workflow system is the 'work queue' – the jobs currently requiring to be done by a given person. Each item has some details of the owner of the piece of work, the group they're in, an indication of what it's about, and an expiry date (see section B.4.1). As work comes in in the morning, it is put on the queue, where it remains until completed. Managers can see the work queues of the staff below them.

An interesting feature of this case is the design method used in Project Meridian to establish the workflow processes for automation. These were reminiscent of participatory design: a wall in a conference room was covered in PostIt notes containing the existing processes. These were created and reviewed by people from all levels of the organisation (although not simultaneously, so more junior staff would not feel intimidated about talking frankly in front of their boss).

---

[15] This is an extract from my MSc dissertation (Ramage 1994) at the University of Sussex. In that work it appears as section 5.2. References to section numbers in that work have been omitted.

[16] My contact at BFS, involved in the workflow project.

B.2.   Method

Most of my time was spent in one group of the Personal Tax division. My main aim was to watch in as unobtrusive a manner as possible. Clearly, the participants were aware of my presence: several times they referred to me (for example, when making coffee they commented that the milk had run out sooner because "we have extra staff here today"). On occasion, they offered me information, or I asked questions; but in general, I tried not to interfere with their work.

I tried, of course, to impose my categories as little as possible on the situation, but this was difficult to achieve for two reasons: first, I worked in similar settings to the one studied for two years; and second the length of time available to the study. Because of the latter, I had a series of questions I wished to pursue (which I had deliberately not made explicit to myself, but were still present): the nature of the organisational structure; how well they'd used BPR and TQM; how work and documents flowed around the organisation; how artefacts were used. Answers to all these came up during the day.

Because of BFS's concerns about intrusion, I did not take any kind of recording equipment with me. However, I made fairly full notes, including several chunks of conversation. While these chunks were not written down to the level of accuracy necessary for conversation analysis, they still proved useful in the analysis of interactions within the group.

It is worth sounding a note of caution here: the group was not necessarily typical of those within BFS Personal Tax. I was told a couple of times about conflicts between the group's use of Staffware and the opinions of senior management about how it should be used (these are dealt with more fully later). Also, the day I visited was not typical of the group's normal actions: their manager, was away much of the day, which does not normally happen.

B.3.   Analysis of organisational structure

B.3.1 Hierarchy and Empowerment

BFS remains, even after the restructuring, a resolutely hierarchical organisation both in structure and in culture. Group G is composed of seven members in total: of these, one is designated Manager and another Assistant Manager. The group manager's role is to oversee the work of her group, and this is reinforced as well as supported by the workflow system: she (or her assistant) can view all the work on each group member's work queue. But the group manager is also subordinate to other managers. The manager in group G spoke at one point of the conflicts her group had with "senior management" over deadlines (see below).

It is difficult, then, to see to what extent the workers at BFS are empowered. An amusing example of this was the presence on the walls of two quality management documents: a 'quality mission' and a set of quality affirmations (e.g. "I shall plan my day"). The former was a fax signed by the Managing Director; the second was headed "Commandments". My impression in both cases was that, whatever the actual source of the documents (quality circles and the like – see below), they were impositions from on high.

I put the questions of hierarchies and flat structures to Alan Mabbs. His comment was "maybe that's historical, but in a business like this, I think it's inevitable", and that the likelihood was that this level of hierarchy would continue. Notwithstanding this, there seems to be a strong aim towards the professional bureaucracy in BFS – Mabbs describes workflow as "right-skilling ... turning people into professional tax administrators". Again, this shows the pull between the professional bureaucracy and the machine bureaucracy that is an aspect of the dialectical tension in workflow discussed in Ramage (1994).

B.3.2 Inter-team interactions

One thing workflow is classically used for is the interaction between geographically or functionally dispersed teams (especially in organisations which haven't gone through a total business process reengineering programme). In BFS, there is very little interaction between teams, with the exceptions of the mainline teams passing work to and from the GSUs, and the specialist mainline teams occasionally seeking assistance from the geographical mainline teams.

The first of these mainly involves the typing of letters, entering of data, filing and other secretarial and clerical functions. These are almost all paper-based, or use other non-computerised media, such as dictation cassettes – every administrator has a dictation machine. Work is passed to the GSUs using Staffware (so it remains on the work queues), but documents are not: these go to and come back from GSUs in entirely paper forms. Indeed, I saw one occasion (apparently quite common) where the GSU returned a note on Staffware saying that the work was complete, and that additional information was to be found on a note with the work.

The second form of interaction between teams occurs when a piece of work has been passed from a geographical team to a specialist team. The passing of the case uses Staffware (transferring it from one work queue to another). There is also the possibility of further interaction about the case between the original team and the specialists; this typically happens verbally.

The lack of inter-team interactions can be seen in two ways: as the result of a quite successful BPR programme (so that there is no need for work to move between teams, except clerical purposes); or as an instance of a rigidly divisionalised organisation.

B.3.3. BPR and TQM at BigBank

As mentioned earlier, an extensive restructuring programme led to the formation of the Peterborough site. This is felt to be only partly BPR by Alan Mabbs. Certainly the restructuring has ended up with a fairly traditional form within the divisions for a financial services company: several geographical teams with a few specialist teams. While these are not functional in the way Hammer and Champy attack, they are still quite strongly based on 'the way things have always been'. Also, a strongly divisionalised form has been retained in the company as a whole, with five divisions based on the type of product.

As to TQM, I saw more overt evidence of this: there was a quality mission on the wall (though signed by the managing director, as discussed above); there was a newsletter outlining the quality circles programme. The group manager said TQM was important to BFS: for example, a Quality Improvement Team (QIT), containing people from all parts of the company, had recently redesigned the standard letters used.

B.4   How does workflow fit in?

B.4.1 Workflow as co-ordinator

Mostly, Staffware seems to act as a co-ordinator of work within BFS. As work comes in, it is logged on Staffware, the type of work being set by the administrator. This is then used for the administrators to see what work needs to be done (and when); and for certain automatic processes, particularly the issuing of reminder notices. This fits with the 'process control' characterisation of workflow seen in several places in my MSc dissertation (Ramage 1994).

Deadlines are set in Staffware for each piece of work (typically four weeks). After this time the fact that the work has expired appears on manager's screen. Until the deadline, the administrator can organise their own work (which is fairly empowering). When deadlines are passed, it is up to the manager to find out from their staff member why this has happened; there is no automatic reminder mechanism. There has been conflict between this group and senior management over whether work should be completed as quickly as possible or close to the deadline. This is an important question, as the use of a date other than the deadline is a form of control by the management which supersedes the workflow system, and in a sense could render it pointless.

B.4.2. Mechanisms for in-group coordination

While Staffware was used as a coordinating mechanism within the group, other mechanisms were also used. The two main others I observed were conversation, and observation of others' actions. In distributed cognition terms, these are both forms of social-distributed coordination (Rogers and Ellis, 1994), but the former is explicit while the latter is implicit.

A surprising amount of coordination took place via conversation. Even when a case had been passed to the manager for approval, often the manager asked the administrator something about it before approving it. This was easy to do given that the whole group worked in one place: it is interested to wonder whether the same sort of coordination would have occurred had they been remotely located. Conversation was also often used for the transfer of knowledge and the tapping of the group's collective memory.

There was much coordination of activity by tacit means, based on observations by one administrator of another's use of a shared resource. These particularly included: someone taking a document from the printer; someone consulting files; someone consulting a book ("Tolley's Tax Tables"); and the use of a shared whiteboard on which completed cases were written at the end of the day.

B.4.3. The lost letters – what happens when the system goes wrong?

An interesting occasion of the system failing to work occurred when the Staffware link to the group's printer went down for most of the morning (cf. Rogers, 1992). By the time technical support had fixed this, many letters automatically issued as part of the completion of a piece of work had been lost. When the work was completed, it disappeared from the work queue, the usual way to access work in Staffware. No 'audit trail' of past work is available to users of the system (though it is available to the system administrator).

This is a real problem with the workflow set-up: it is possible to send a letter to the printer, have it lost in the system and not be recoverable by the user. Unfortunately, I was away from the group for the period of the afternoon when the re-sending of these letters happened, but the impression I got was that they intended to recover these from memory and paper records around their desks.

B.4.4. The reminders – all the interesting things happen without you noticing

One important part of the system is the issuing of automatic reminders to customers and others, to pay bills, give information and the like. These occur at regular intervals if appropriate responses haven't come back from the customer. Most administrators have a few automatic reminders issued every day, at 1.10pm. This kind of standardised, automatic process is exactly the kind of thing workflow does well, and it seems to fit in with the set-up at BFS.

These automatic reminders were mentioned to me several times during the morning, and I was told to watch out for them. Accordingly, I made sure I was in the group observing at 1.10pm. However, the time came and went without any notice at all – the printer whirred a bit but nobody jumped up to collect things from it. This shows partly how ethnographers manage to miss the most interesting things happening in front of them; but it also shows a positive side of workflow, that the reminders were produced without anyone needing to do anything at all, and that they were ready when needed.

B.4.5. Do they like it?

Most of the reactions I got from users seemed positive, though they made clear the extent to which it acts to coordinate work that mainly occurs elsewhere. The group manager, said, "I don't think we could manage this volume of work without it"; an administrator, said, "At least you know where most of the stuff is most of the time".

Less positive notes were sounded by another administrator: "On the downside they can see what you're doing and complain if you haven't done anything at the end of the day". This shows the extent to which workflow is used, within a very hierarchical environment, very much as a control mechanism. While this may be good for the business, it can also be challenging for the administrators who are at the 'bottom' of this hierarchy and can be watched by their managers.

Again here we see the tension between control and empowerment which in Ramage (1994) I have described as being at the heart of workflow.

Appendix C: PETRA – Participatory Evaluation Through Redesign and Analysis

Removed for copyright reasons in this version. Available as:

Ross, Susi, Magnus Ramage and Yvonne Rogers (1995). PETRA: Participatory Evaluation Through Redesign And Analysis. *Interacting With Computers*, **7** (4): 335-360.

The abstract of that paper follows:

Compared with single user-computer interactions, evaluating multi-user computer interactions is much more complex. We argue for multiplicity – of theory, method and perspective – in the evaluation of computer-supported collaborative work (CSCW). This allows us to address both theoretical concerns and practical design issues, and to incorporate the expertise and experiences of both researchers and participants. We propose the PETRA framework, incorporating a theoretically-driven evaluators' perspective to investigate the collaborative activity, and a design-based, user-focused participants' perspective to evaluate the supporting tool. Our study investigated collaborative writing, both in a face-to-face context, and supported by a computer-based group editor. In our instantiation of the PETRA framework, we used distributed cognition and a form of breakdown analysis to investigate the development of shared understanding in the two different mediating settings; and devised a rapid prototyping session (inspired by participatory design) to elicit participant reactions to and redesigns of the tool interface. Our findings show that computer-supported shared understanding develops technologically, using social coordination as a repair mechanism; and that the collaborative tool must be particularly sensitive to issues of awareness, communication, focus and ownership.

Keywords: CSCW; Evaluation; Multiplicity; User Involvement; Collaborative Writing; Distributed Cognition; Shared Understanding; Participatory Design; Usability Criteria.

Appendix D: Common MIS[17] Requirements Process – An Alternative Method

This document relates to the methods to be used by the "Common MIS User Interface Requirements Working Group" and has been prepared at John Smith's[18] request. It is mainly based on Bob Jones[19]' memo of 31 January 1995 and the attached notes, and the users' meeting held within the Engineering Faculty on Tuesday 7 February 1995. It is my belief from these sources that the methods currently intended for use for the study of user requirements for the Common MIS systems, both Cassandra and future systems, is not likely to provide an adequate solution. Accordingly, in this document I will outline an alternative methodology.

There have in the past been several high-profile failures of systems due to inappropriate requirements capture methods being used. Most notable were the Taurus system at the London Stock Exchange, and the routing system used by the London Ambulance Service.

D.1. Background

The method discussed here derives from work conducted within the field of computer-supported cooperative work (CSCW). This is an interdisciplinary research field which studies the nature of work that is conducted by groups of people, which currently does or might in the future use computer systems as part of this work. It first was defined as an explicit area in the mid-1980s in the USA, although its roots go back about a decade further. Within this university, a considerable amount of work has been conducted by the Computing and Sociology departments; this group has an international reputation in the integration between the study of work processes and the incorporation of the information gained into the software design and installation. Given this considerable amount of expertise 'on site', it seems worth reflecting on the requirements process before rushing into the design or redesign of particular Common MIS systems.

D.2. Why consultation is not enough

One of the recurring discoveries within CSCW, and the work which preceded it, has been that in the discovery of user requirements before the design or the installation of a system, the traditional method of systems analysts (or intermediaries) asking users what they want, producing a report and handing it to the implementers of the system, simply does not work. This is partly due to the information gained thus and partly due to the way it is passed up a form of hierarchy (and hierarchies are notorious among management researchers for throwing away information at each layer). Both of these tend to hide requirements rather than show them.

The problems with this method are two-fold: a lack of appropriateness of the system to user needs, and a lack of user acceptance. Firstly, they simply aren't fine-grained enough to find out what the user actually wants and needs to do with the system, partly due to lack of time in briefing meetings, partly because of the information filtering discussed above. Secondly,

---

[17] "Common MIS" is a pseudonym for the project conducted by a number of UK universities to produce common management information systems, under which the Cassandra system was produced.

[18] The leader of that working group

[19] A senior academic at Poppleton University

systems that are foisted upon users without consideration of their wishes will tend to lead to resentment, and hence to less efficient use of the system than is possible. This cannot be corrected simply by telling people to use the system, as they will in this case do so but less appropriately than they could. It is also worth noting here the comment made by several CSCW researchers, that the people who buy computer systems are not those who use it. It is also worth noting the paradoxical finding that the sure way not to obtain users' real requirements is to ask them directly.

D.3.   An alternative method

I suggest, then, that an alternative method is required, to get around the problems discussed above. There are two aspects to this, to reflect the different aspects of the process: studies of work in practice (known as 'ethnography' by sociologists), and user involvement in the design process.

The first of these can be used either to study work using an existing computer system that may require some modification (for example, Cassandra), or alternatively to study work before a new system is designed, or installed, to support it (the other Common MIS systems). The basic method is one of 'hanging around': sitting and watching what people are doing while they are working in their normal everyday work, taking lots of notes about it, maybe tape-recording some of what's happening. This is often supplemented by unstructured interviews with the various people involved. The key feature of this approach is that it aims to capture the aspects of work most significant to those doing the work rather than to fit what people are doing into theories or hypotheses. In the context of the Common MIS systems, this would involve those doing these studies in going around the university talking to people in the departments and in the central administration who use the system both day-to-day, and who receive reports from it, finding out their work methods, needs and problems. This should ideally be done by those with experience in such matters, such as the CSCW researchers mentioned above.

It is possible for those doing the ethnographic studies to report their findings to the system designers, and for those findings to form the system requirements. To an extent, this will prevent the first failing of traditional methods above, the lack of appropriateness to user needs, and it certainly forms a good starting point for the requirements capture. The remaining problem, however, is one of user acceptance: the best system in the world will not do the task it was designed for if its users don't use it.

A way to get around this problem, to ensure a 'buy-in' to the system by the users, is to involve them directly in the design process (a method known as participatory design). A common way of doing this is to first conduct ethnographic studies as described above, give these to system designers to produce a prototype system in a form that is easily changed (perhaps as easy as a set of drawings on paper), then to discuss this with a group of representative users of the system, have them comment on it and suggest changes, then for the designers to produce a further prototype and users to comment on this, and so on. It should be noted that the users are those mentioned above who use the system day-to-day and also who get reports from it. The aim of this method is that it will, as well as more closely refining user requirements than the initial ethnography, increase the effective use of the system, as (some of) the users were involved in defining what it should do. This method has been very successfully used in a number of different situations, particularly in Scandinavia.

That these methods will take more time and effort than the traditional approach is undeniable. However, the expertise to conduct them is available within the university, and by using the above methods it may be possible to stop the redesign of Cassandra, and the design and

introduction of the other Common MIS systems, from having quite so many problems as the initial introduction of Cassandra has had. It is important to be seen to clearly separate the process of developing requirements from existing organisational structures. This issue is particularly acute in the case of 'gatekeepers' and 'filters' of information.

Finally, it is worth mentioning that I have a research interest in these systems, which fit well into my PhD work; and therefore I am able to spend a reasonable amount of time doing, facilitating and coordinating the tasks described.

Appendix E: Write-up of RICS case study

This paper does not especially discuss the use of SESL in the RICS study, except *inter alia* as it discusses the way I conducted various aspects of the research. A discussion of that relationship can be found elsewhere.

I discussed many questions with the members of the project at various stages during the study, but four came out as most important, and they will be discussed here: how was the project 'learningful', how the research has been better than the traditional model (these two are summarised in one section below), what has been the effect of the topic upon the research style, and how technology has facilitated their learning.

The focus of this study is principally on styles of co-operation and communication rather than the tools to support those. This arises from the issues that turned out to be interesting as the study progressed – the nature of learning in the project and how the research was conducted – and the fact that the project team's use of computerised technology was limited to email. I shall however discuss the use of email and its relation to other communications technologies (phone, fax and post) as appropriate.

This paper describes the background to the study – what I was doing there and what was my initial remit; then the way in which I conducted the research; and then explores the four key questions that arose, as given above, in turn.

E.1. Background to the study

The study was a one-year evaluation of the learning processes of a research team. This team was researching learning processes, and specifically organisational learning, within the surveying profession (funded by the Royal Institution of Chartered Surveyors, the RICS). It took place at Lakeside University[20], which is based in a large city in the north of England. The project was primarily carried out by four people: a full-time researcher (Fiona), and three other staff members acting as advisers to her (Mark, Dennis and Mary). Around two months into the course of the project, Fiona realised that although the project team were looking at surveyors' learning processes, they had made little consideration of their own. As she knew of my interest in evaluation, she therefore asked me to come in as an outsider to evaluate their work. I started this process in late November 1995, and conducted my final interviews in early November 1996, shortly after the end of the project.

E.2. The conduct of the research

My main methods for conducting the research were a standard set of ethnographic ones: interviews, attending meetings and hanging around. My role was enhanced slightly, as well as giving me some useful data, by additional activities I carried out, principally the programming of a database/spreadsheet combination to allow for easy entry of the data from a questionnaire; and providing feedback at various stages, but especially the summary I prepared of an exercise during one of the steering committee meetings. My role was also frequently to be a "human reminder" of their commitment to the process of change and of conducting their research in a more reflective manner.

---

[20] This name, and others, have been changed.

During each of these activities, just studying them was often enough: by conducting interviews, I acted as a transmitter of ideas from one person to another, for example; and I gave each individual space in which to reflect on the workings of the project with another person who knew what they meant.

E.3. How has the project been learningful?

The key question here has to be to what extent the project acted in a 'learningful manner', by which I mean both its willingness to innovate and its ability to learn from its mistakes. This is a development of Pedler et al's (1991:1) definition of the learning company as "an organisation that facilitates the learning of all its members and continually transforms itself". A discussion of the sources of this learning will form the following section; here I wish to examine instead the evidence for saying the project was innovative and reflective.

E.3.1. Conduct of Steering Committee meetings

The way the project conducted its meetings, especially those of the Steering Committee (a body consisting of representatives of the various stakeholder groups – especially from the RICS and surveying firms, but also people with knowledge of organisational learning) seemed to me to be rather innovative. This group met five times over the course of the project, and I attended three of those meetings. Although they used a fairly standard committee format – taking minutes, sitting around a table, serving lunch before the meeting etc. – several things were attempted to give space for more reflection. (However, see section E.3.3 for a discussion of a conflict which occurred because of a lack of reflection during one of the Steering Committee meetings.)

Most importantly, from the third meeting onwards, the data so far gathered in the project (variously, questionnaire responses, graphs, statistical analyses of quantitative data etc.) was put on the walls of the meeting room for Steering Committee members to read and comment on during the lunch held before the meeting began. This allowed more time for reflection than if this information had been handed out at the meeting, and also allowed members to reflect on the aspects that arose for them rather than just what seemed important to the project team during the meeting. It also served to signal to the committee members, many of whom were used to meetings being conducted in one style only, that these meetings were to be different. This approach was also taken in the "RICS roadshow" used in many presentations of the project later on (see section E.3.4).

Next, the minute-taking was conducted in a way that allowed space for reflection on, and ownership of, the items under discussion before they were concluded. Instead of someone taking notes during the meeting and then writing up minutes at the end of the meeting, the procedure followed was: At the end of each item, a pause occurred while Fiona drafted a minute. She then read this out and members could comment on it, ask for changes, or re-open the discussion according to what they'd heard. This seemed to give much more space to the members to consider what had been decided.

In one of the meetings, an exercise was conducted, at the project team's instigation to determine how the project was progressing so far. Eight sheets of flipchart paper were put up around the room – labelled "What I give", "What I gain", "What this Steering Committee is there for", "What I like about this Steering Committee", "What I don't like about it", "What this Steering Committee lacks", "How the Steering Committee could be improved" and "What I'd like to do here if I had a free wish". Those present went around the room writing comments under these headings, and sometimes adding to others' comments. As this seemed a particularly good form of self-evaluation, I undertook to summarise the contents of the

flipchart sheets. Below are parts of the document I wrote from this. Many of the views expressed are therefore the committee members' rather than my own.

I have divided the eight topics in three categories: input and purpose, feelings and 'how I would like it to be'[21].

1. Input and Purpose

a) *A focus on knowledge and ideas:* What members of the committee felt in large part they both gave to it and gained from it, and part of the purpose of the committee, concerned ideas and information. Many had lots of good ideas of their own and were able to put these in to the project. All gained in this area also – knowledge of and insights into surveyors, reflections on the nature of steering committees and projects, additions to their academic research, and lots of learning. Indeed, some felt this was a significant purpose of the committee: to exchange and encourage ideas.

b) *Experience in and experience out:* Many people said their input to the project consisted of energy and time, for some a great deal. There was also mention of members' experiences being important: their links to the surveying profession and the RICS, their feelings for the important parts of organisational learning, and their abilities in dealing with statistics, computers, writing and presentation. For some, one of the outcomes of the project had also been in terms of experience, particularly of what surveyors do and of how steering committees and project teams run.

c) *Support and encouragement:* A number of people stressed that, for them, the committee acted as a support mechanism. It gave encouragement, it supported the community that was the project (especially the project team) and it provided a forum for sharing. It also encouraged those who felt isolated about the ideas being discussed, that they were not alone in their enthusiasm and others were out there who also felt the importance of organisational learning.

d) *Keeping the project on track:* Finally, several people stressed in different ways that the Steering Committee was there to be the supreme body over the project: to make key decisions, to guide and monitor its progress, to act as an 'official frame', and to keep it going along the right tracks.

2. Feelings

a) *A learning committee?* Some people found the meetings refreshingly informal in their style (although one disliked that); others found them depressingly formal, especially in the lack of communication outside of meetings but also generally in the way the meetings are run. The learningful nature of the meetings was appreciated by many: there was comment on the openness of the meetings to challenge norms and assumptions; on the honesty and candour with which the meetings are conducted (no politics); and on the opportunities to share and listen. This meeting was also felt to be especially open – the exercises conducted to reflect on the interim report and the work of the committee were felt by many to be particularly valuable.

b) *Wonderful people, but not enough time to get to know them:* Many people expressed the view that the people on the committee were extremely valuable and enjoyable to work with, provided a wide variety of different perspectives, and gave a great deal of commitment to the project. However, there was little time to get to know them, as the agendas were too full and meetings were too short – this made for a degree of superficiality in the relationships developed within the committee. For some, this was also due to how long a journey they had to make to get to Lakeside University.

3. How I would like it to be

a) *Widen the membership of the committee:* Several comments were made on who should be on the committee, all wanting to increase its numbers. A number of people wanted more practising surveyors on the committee; one wanted more academics involved in organisational learning on it; and someone suggested that an RICS database contact should have been invited to meetings. Several people felt it would have been helpful to decide roughly how many people were needed at meetings, and then to invite double that number.

---

[21] This was a reference to the Learning Company Questionnaire used by the project, which poses 55 statements and asks someone filling it in to give a number from 1-7 as to 'how it is' and 'how I would like it to be'.

b) *More time and more involvement would have helped!* Given the stress placed elsewhere on there not being enough time (see 2b above), it is not surprising that several people wished there had been more time available in the meetings; one person suggested this might have made the committee higher on people's priorities. Closer involvement and more regular attendance by more of the committee's members would have been useful.

c) *More like today – variety, imagination and participation:* Many people had found this meeting very stimulating and helpful in its style, and wanted that to continue. In particular, words like challenge, imagination and variety were used as indicators of the kind of things that might happen. Exercises like the ones carried out at this meeting were worth continuing. Another helpful aspect of the current meeting had been a more participative style of running the meeting; this was highly beneficial and should continue.

d) *We want to do work in the meetings:* Comment was also made by several people on the content of the meetings, and further things that could occur during them. The key focus of this was working together – there was a general feeling that the committee hadn't done enough actual work as a committee, but rather had made decisions. One suggestion arising from this in the subsequent discussion was that the next meeting (if the project was extended) involve an analysis session of the questionnaire data, to try to put as many brains to bear on this as possible. Another possibility suggested on the flipchart sheets was that there be more discussion on the concept of a learning organisation, and on how to use the ideas in surveying firms and elsewhere (in particular, the faculty). Finally, someone said meetings should contain more fun!

E.3.2. Self-awareness

A key part of learning and change is to know oneself, and thus to gaining knowledge about what needs to be changed (cf. the second part of Pedler et al.'s definition of a learning company as one that "continually transforms itself"). Thus self-awareness becomes crucial to learning. How aware was the team of its own collective learning, and its own need for learning?

It was about two months into the project that Fiona realised there was an anomaly between its focus on organisational learning and the lack of focus of the team on its own learning processes. There were two immediate results of this: the "go-round" sessions held at each G4[22] meeting and my involvement as an evaluator of their learning.

The "go-rounds" were so called because they consisted of ten minutes or so at the start of each meeting going around the team with each member saying how they had been feeling about the project recently: what made them excited, what frustrated them, what ideas they had had. At the team meetings I attended, I observed a wide range of different styles, ranging from the placid, with each person just simply stating what they were up to and how they felt; to the more stormy, even to the point of someone bursting into tears on one occasion in frustration at the current situation. I remember little conflict over what people said: on the whole, these statements were accepted without question or debate (although this may have taken place at other meetings than the ones I attended). The extent to which the go-rounds took place waxed and waned: at first they were very important, then they became less significant, with occasional recurrences. This can be characterised either as the team becoming more and more comfortable with each other (and so more aware of the others' situations), or a general lack of interest in looking at themselves as the time pressures became more intense and various conflicts drove holes through their consensus. The latter phenomenon certainly seemed to be present on occasions.

---

[22] Because there were four main members of the project team, they often referred to themselves as the "Gang of Four", and in writing shortened this to G4.

When they did happen, the go-rounds seemed to me to be evidence of an attempt by the team to act in a way entirely appropriate to the learning organisation: to integrate their work and emotional lives, to conceive of themselves as whole people with many different kinds of needs, rather than robots who just worked. It also served a role in seeing the working style of the project as itself an experiment in new practice, so that they were not just offering a vision of the future to surveyors but actually living it out. Such consistency is admirable. As I've shown, it wasn't universal; but when it happened it was excellent.

The second part of their openness to considering their own interaction was my role as an external evaluator. I was asked to fulfil this role because of the need for some kind of facilitator to guide their own reflections, and because of my interest in evaluation. So I attended a number of G4 and Steering Committee meetings, conducted two series of interviews with the team (I intended a third set, but they never materialised), hung around in Fiona's office, and wrote them a database program to enter the questionnaire data.

Reflecting on my continuing presence in the project as it went on, it was clear to me that I wasn't doing anything much in particular except being there, and occasionally asking questions, and yet my presence seemed to have some kind of effect; so I characterised what role I had in the project as one of facilitating organisational learning, and that seemed to fit for them. A helpful metaphor for this that I have since formulated is to think of the evaluator not as a scientist in a white coat, finding out a clean and objective truth, but as a midwife, assisting at a messy and untidy birth: and this was very much the way it felt in this project.

Did my being there change them much? I'm not sure. I certainly gave them occasional feedback on paper (when I made a list of all the relevant documents I'd produced for them, it came to a surprisingly big one, about ten). I suspect that some of the questions I asked provided them with space to reflect on their experiences – in the stakeholder mapping exercise (documented elsewhere), for example, this certainly seemed to be the case. Yet my chief role seemed to be that of a 'human reminder' of their commitment to considering their own learning. I was around quite a bit, especially at formal meetings, and I was not really playing the unobtrusive ethnographer: at Steering Committee meetings, for example, I did not wear a suit like everyone else but rather my standard graduate students' uniform – casual trousers and a shirt. My actions, too, were different: I sat in the corner writing, rather than talking about the project like everyone else. So just by being there I perhaps had some effect, a further encouragement not to get stuck in the old ways of thinking about research projects but rather to move forward in new ways.

E.3.3. Open approach to team interactions: more aware of conflicts?

If one of the team's basic aims was to be open in their interactions, how do this mean they handled conflicts? One source of conflict is if various members of a group hide their feelings about what is going on, and "commitment to airing differences and working through conflicts" is therefore one of Pedler et al's (1991) features of the second characteristic of a learning company, "participative policy making". In the RICS project, the view of how much conflict was around, and how it was handled, seemed to depend on different individuals. As the full-time researcher, Fiona was most likely to come into conflict on a day-to-day basis with others, and two examples of how she participated in, and resolved, conflicts will be useful.

The first happened with Mark over a letter to go to the Steering Committee members, drafted by Fiona but signed by Mark. He was unhappy with the formatting of it – there was not enough space at the end for signatures to go on without appearing (to him) cramped, and so he blocked it from going out that day, a Friday. As Fiona had taken a good deal of effort to

ensure it was complete for sending out that day (as it was important to get the information in the letter to the committee members by a particular time), she was considerably angered by this. The following Monday, she went to see Mark and said she was very unhappy about this, and she didn't want to write letters for him again if that was going to happen. Mark commented to me much later in an interview: "We certainly aired our differences on that particular issue, amicably I think – but it was far better that she let me know that she was upset by what I had done, and it gave me an opportunity to explain why I had done it."

Another example was a conflict between Fiona and Dennis over the August 1996 meeting of the Steering Committee, which he chaired (as Mark, who would normally have chaired it) was on holiday. There were a number of occasions during that meeting where Dennis's chairing meant that Fiona's voice was not heard, and that decisions were thereby made over areas in which she felt considerable ownership (such as the title of the report) essentially without her. The chief contributors to this meeting were men, leading her to afterwards conclude that the conduct of the meeting had been in a male way that had thus excluded her. This soured relations between Dennis and Fiona for a little while (at a time when it was very important for the final stage of the research that they work together), until they had a meeting at which she raised this issue and they were able to clear their differences.

However, this discussion of the surfacing of some conflicts must not be felt to mean that *all* conflicts were resolved so openly. Mary commented in an interview after the end of the project that the lack of close day-to-day working by the whole project team meant that there were potential areas of conflict which didn't really arise.

In particular, gender was less of an issue than it might have been (given that two of the team were male and two were female). It certainly seems to have been the case that gender was one of the causes of the conflicts between Fiona, Mark and Dennis (although along with other factors: as Dennis put it, the fact that he was tenured, secure, middle-aged, as well as male, whereas Fiona was younger, and hired only for the project, as well as female). However, the issue only really surfaced as an explanatory factor after the August Steering Committee, as described above.

A further area of conflict which Mary identified as quiescent was between the 'academics' and the 'practitioners' on the project team – not so much their experience but their orientation towards management theory or everyday management (Dennis being a pure academic and Mary a pure practitioner, with Mark as a practitioner with an academic leaning and Fiona an academic with a practical leaning). There was certainly a tension between the two sides of the project, but this balance in the team's membership and a general preference to produce a report that would actually be read by surveyors rather than being left on a shelf may have stopped the development of personal conflicts.

E.3.4. Presentation style, talks and reports

Not wanting their report to be left on the shelf through being rejected out of hand by sceptical surveyors (hence having to be fairly conservative), and yet to raise enough issues of change (hence having to be radical) required a tricky kind of juggling by the project team. Their solution was an interesting one. The bulk of the report forms a reasonably standard academic project report in the management sciences, albeit a nicely presented one: it begins with the background and aims of the project, describes the literature (organisational learning) on which they based their work, goes through the data-collection process and analyses those data in appropriate statistical ways, and draws some conclusions for surveyors. And yet it is framed by two other items, which change the tone completely and speak significantly to other needs. There is an extensive set of appendices which contain graphs in abundance, codified data and

statistical analyses. There is also, beginning and end the report and interwoven into it, a set of mini-dialogues between an unidentified authorial voice ("I") and "three grey-suited people" who are the representative voices for surveyors: this is entitled, as if it were from Chaucer, "The Surveyors' Tale".

The effect of the appendices is a strange one. I was present at the Steering Committee where a small set of the graphs contained therein were first shown to the surveyors. This was at a time when the project, for various reasons (many beyond their control), was running short of time, and so was asking the RICS for money to continue their work for another four months. The very existence of these graphs (regardless of their content) seemed to convince Cliff Hardcastle, the RICS representative, that it was worth continuing the work: if they could produce real numbers, they must be doing something right. Fiona says the same phenomenon has happened in a number of the talks she's given to surveyors on the outcomes of the project: she talks about organisational learning, the questionnaire, the need for change in a sharply competitive sector, and so on; but it is when the graphs first appear that the surveyors take notice. And again, they don't feel the need to see the detail: just the graphs' existence is enough to convince these people (clearly positivists!) that there is some 'Real Science' going on here, and that it's worth listening to. The power of the (metaphorical) White Coat is mighty indeed.

So the graphs serve to justify their other work. What of the Surveyors' Tale? The point of this piece of writing is not just to break up the report and make it less dry, and easier to read (although it does serve that purpose *inter alia*). Its primary purpose is to challenge surveyors' *Weltanschauungen,* their basic world-views and assumptions about work and organisations. It is possible to do this through the standard writing-style, but one's hands are tied, because even if one makes statements that are attempting to challenge the existing world-view, if one writes in a way that is strongly associated with it (and shaped by it to some extent, as with the case of the graphs), then one is sending contradictory messages. There is an explicit message of "change your world-view", and yet there is a meta-message that assumes that world-view. Since surveyors tend to be fairly conservative, they will (indeed most of us would) tend to go with the meta-message, as that is the path of least resistance, that preserves the *status quo.*

It is for these reasons that the Surveyors' Tale makes sense. Written as a dialogue, it contains within its form the topic under discussion – change, new thinking, and reflection. It could thus be described, using Gregory Bateson's term, as a *metalogue*: "a conversation about some problematic subject … such that not only do the participants discuss the problem but the structure of the conversation as a whole is also relevant to the same subject" (Bateson 1972:1).

This attempt to challenge assumptions through multiple kinds of message was an important feature of the presentation style of the project team, and it can be seen in another form of presentation that Fiona introduced: the Roadshow. I have above (section E.3.1) briefly discussed the way that wall-displays were used at a number of Steering Committee meetings, showing graphs, numerical data, qualitative responses to questionnaires, discussions about the project team such as the stakeholder map, and so on. This first arose for the Learning Company Conference, in April 1996, with the aim of letting multiple voices be heard: Fiona and Dennis, who were giving the presentation, were concerned that only their two voices would be heard there, despite the large number of others who had things to say about the project. They were also concerned to get the maximum amount of information across in a short time, and this necessitates the use of more media than just the traditional talk-with-overheads. Therefore they used tapes – video and audio – of various other members of the project team and Steering Committee, and the wall-displays discussed above.

Innovation and experimentation in presentation styles were an important feature of this project, and they certainly seem to have contributed to getting its messages across in a way that questioned assumptions rather than playing along with them. This is therefore a way in which learning can be seen to be not just an end in itself, but effective for other action.

E.3.5. Openness to change in research style

Of course, this was not just any team but specifically a research team. It is therefore particularly instructive[23] to look at how the team conducted their research, and whether that was more learningful than the 'standard' approach.

Some examples of the way this might be held to be so included:

- They changed their mind fairly early about the definition of success. Originally they had planned to use a fairly objective measure – a listing of good surveying firms by region and type published by an independent firm, and known as the 'Green Book', which is widely quoted. However, this proved to be too rigid: in particular, it failed to list smaller companies or those operating in just one area, and it was important to the team to include such companies. So they moved their view of success to one of peer acceptance, sending out a set of preparative questionnaires to worthies within the surveying profession, asking them for lists of companies *they* found to be successful. These then formed the basis of the recipients of their later questionnaire.

- Partly as a result of the previous change, and partly due to the slow response of the RICS to their request for addresses of the firms they wanted to send questionnaires to (a large number), they encountered timing problems in the second half of the project. This required a reasonable amount of restructuring – the project plan was reconstructed several times (although that is fairly common in some research projects) and they later needed to ask for more money from the RICS to continue the project for a further four months.

The argument might also be put that these changes were fairly superficial. Mary expressed it thus: "If anything, we were fluid because we didn't sit down and construct a proper research methodology.", and described their fluidity as expedient rather than intentional .Thus the question is open; but there is some evidence that their approach was fairly open to change.

E.3.6. Development for all

A key feature of the learning company, according to Pedler et al's definition, is that it supports development opportunities for all its members. I raised the question with the project team members as to their awareness of what development it gave them.

Fiona, as the full-time researcher on the project (and so with no other sources of development opportunities) had the chance to go to a number of learning events, both academic conferences and consultant-focused networks. These were of personal benefit, but also significantly benefited the project – the kind of "win/win" outcome which is anticipated in Pedler et al's work. In Fiona's words (interview, 6/5/97): "wherever I went, I came back and brought something to the project; whichever outside or inside training events I went to … I brought back to the project"; Mark also remarked (interview, 4/11/96) that this learning for Fiona has also benefited the project.

---

[23] Cf. the argument given by Garfinkel (1967) of the importance of looking at participants in a study as people doing a real task as well as just fitting in to one's research aims.

For Mark, the developmental benefit was that he had not worked on a research project before. Although head of the surveying division at Lakeside, the traditional culture, like at many other new universities[24], at that university had principally rewarded teaching, administration and perhaps consultancy, rather than research, and that had been his path. In his words: "I don't have a particularly strong research background, I've tended to be more involved in administration and got promotion that way rather than by research" (interview, 12/12/95). He also reported that finding out about organisational learning was beneficial for him – he already was fairly self-aware and keen to develop new learning opportunities in his division (see section E.4), but the academic aspects also were new and exciting to him.

Dennis and Mary both have a more closely-developed academic background. They reported that the main benefits of the project for them were the chance to be involved in the survey and its analysis, and the subsequent writing of papers. Dennis also found that from a political point of view, bringing money into his department (the business school) was beneficial.

E.4. Influence of the research topic upon the research

A question that occurred to me early in the project was the reflexive one: "how much does the fact that they were looking at organisational learning affect their own learning processes?". We have seen above the ways in which they were learningful, and it is tempting to conclude a causal link here. Examining the question and talking to the project team about it has lead me to conclude that this link is inappropriately positivist in its attempt to establish causality. Instead, I suggest that a systemic relationship took place between the membership of the team, their backgrounds, and the fact that they were looking into organisational learning. This can be seen in the follow diagram:
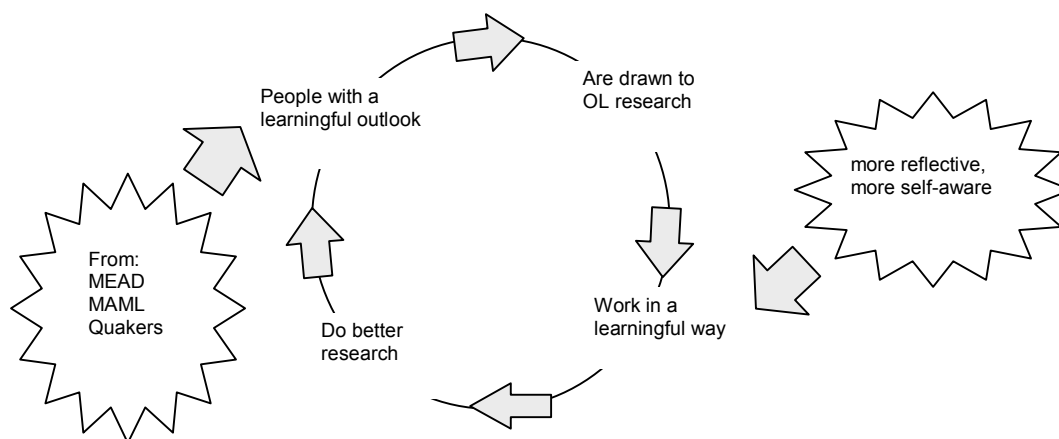
Figure 9: The learning loop of the RICS project

The first, and perhaps the key, point on the loop here is the item "people with a learningful outlook" at the top left. It seems that the team brought together were exceptionally well-focused on learning as a 'good thing' in itself (as having intrinsic value), in their lives and working processes as well as a research topic. This can especially be seen in the involvement of various members of the team in:

- The Department of Management Learning at Lancaster University – two had studied there, and two had served on the editorial board of the journal *Management Learning*, run

---

[24] i.e. those which were formed out of polytechnics around 1992

from that department. The department has a long history of focus on learning in all areas of life. Fiona remarked in an interview (January 1997) that "reflexivity was encouraged … what I learned in Lancaster was 'you're looking outside for an example of whatever – well aren't you actually doing it?'".

- The involvement of two of the team in the Religious Society of Friends (Quakers), a group that emphasises learning as a key part of life and as a lifelong process. A specific influence was the method of minute-taking during Steering Committee meetings, as described in section E.3.1: Quaker business meetings are based on the minute-writing being an integral part of the discussion, with the item continued until a minute has been agreed by everyone (votes are not taken, the decision being made by consensus).

- The practical application of concepts of organisational learning to management work within their own areas of Lakeside University, which two of the team had actively pursued in the past.

So we have seen something of the backgrounds of the members of the team, and how they are used to learning. It seems to be the case that the fact of being learners by nature has drawn each of them to organisational learning research, and this is the next item on the cycle.

That they worked in a learningful way has been discussed above in section E.3, and seems a reasonable conclusion from the experiences discussed there. In particular, they were more reflective and more receptive to self-study (as was shown by their constant self-assessment on both content and process issues throughout the project).

Was the research better as a result of this? The research was certainly strong. Dennis expresses the ways this was so clearly (interview, 4.11.96):

> "The project is good enough to sustain a number of interesting publications, it's probably the most significant bit of data-gathering, certainly on this side of the Atlantic, about the links between organisational learning and some kind of measure of success."

Informal feedback to members of the project team that has subsequently been passed to me also suggests this – the project seems to be well-regarded by those knowledgeable in the organisational learning field. Its effect on the surveyors is more questionable; I have little evidence on that subject (and it would be hard to measure), although it certainly has the *potential* to have been influential.

As to whether it was better through the effects of organisational learning, that is also hard to show. Mary raised the question (see section E.3.5) of whether their research style was as much to do with expediency as design; Mark, on the other hand, was quite clear (interview, 4/11/96) that the style of the research had a great deal to do with the people involved. For myself, I would say that the link is a complex one, but thinking more about what they were doing and why they were doing it certainly seems to have been of benefit.

E.5. Has technology helped them to build an LO?

Their main communications technologies were email (Microsoft Mail), plus the various 'old technologies': fax, phone, post. These were often integrated with each other: there were several instances where a number of emails went back and forth, then the discussion blossomed into a phone call, which then led to a document being faxed, which preceded a face-to-face meeting... But email was the chief computer technology they used.

There were three main uses: community building, by keeping the group alive as a group between meetings and discussing various subjects; passing-around of documents; and dissemination of interesting information.

Several people commented in interviews that passing around of documents as attachments more quickly than the quirky Lakeside internal mail (since the team was based on three sites in the city). This was hampered by the lack of a single email system at Lakeside, but was useful both for dissemination of finished documents and to facilitate editing of draft documents. One occasion where it did become more useful was in the selection of which companies to which to send a questionnaire: email (via Microsoft Mail attachments) was used to allow the team members to select companies and see each others' choices. This raised one technical issue, as two of the team were unable to receive such attachments, and so their input to the selection process had to be by faxes to Fiona…

The use of email for community maintenance was quite important, Fiona said (interview, 6/5/97):

> "it involved making appointments and setting dates/times for meetings: this was quite important because G4 members would often not be reachable by phone [and she didn't have an answer-phone]. But different email software actually led to misunderstanding between Dennis & me – for a long time, over half a year, Dennis put replies & comments into the messages he received from me. It wasn't clear where the misunderstanding was – in the colour-coding that makes the annotations visible on Dennis's screen, but not on mine."

Dennis, on the other hand, was less certain that it had made such a difference. He said that it helped in that he was seven miles away from the rest of the team, so that informal contact just wouldn't happen with them, and so a flurry of emails sometimes kept up that contact; but he did feel the difference was fairly small.

As for the use for disseminating information, this was felt to be of mixed benefit. Its main instance was the passing around of messages which Fiona received from the mailing list *learning-org* (moderated by a member of MIT's Centre for Organisational Learning and concerned with both practical and academic aspects of the learning organisation). Mark commented that it had "been useful to help develop the rest of the team's understanding, though without reading it all" (interview, 4/11/96). Dennis described the messages as "fun and interesting" (interview, 4/11/96), but on the whole hadn't read a large proportion of them, not having the time.

Some of the more negative aspects of email that were mentioned included (as well as the problem with attachments above):

- Between June and September 1996, no messages got through to Dennis from Fiona or Mark, which nobody realised until much later – Dennis just assumed no-one was sending anything, and Fiona assumed that he wasn't replying to them. This became known as 'the great hole in cyberspace', and was a good example of a technical problem exacerbating the social problems of mistrust and communication. Dennis described it (interview, 4/11/96): "That was a weird example of new communications not being safe and certain, but being exasperating". Given that Fiona kept talking about Dennis's unreliability in dealing with messages, and their past history of misunderstandings, this was a problem.

- Mary found that speed was an issue for her. If Fiona sent out a document for comment, Mary couldn't always comment immediately. Dennis, on the other hand, could, so that Fiona then sent out a revised copy before Mary could comment on the first. She said (interview, 4/11/96): "Well, why should I bother? Where am I in this?", and that she couldn't cope with the speed technology puts on her, so was less inclined to do the work.

- Both Mark and Mary reported that frequently it didn't save any paper, as they printed out most messages (and especially attachments) they received, to compare versions of text or just to read them in more detail.

In summary, Fiona remarked (interview, 6/5/97) that email use involved a mix of what one might say on an answer-phone and what one might send by post:

> You can see the role of email if you think about how life in the project team would have been if we hadn't had it. We would have had to send it all by mail. … The collaborative writing process – and there was a lot of it … e.g. the questionnaire changes, comments & annotations on the minutes – all that would only have been possible with days & days & days in between; and then it would have to be collated on paper. … It would have been a lot slower, and wouldn't have been so collaborative – which might have been an advantage at times! – but with this team, email was quite an enabling device. … If I look at all those things we did with this email stuff, it would have been far more time-consuming and a hassle to work together as we did. Certainly the chunky text bits, they would have been just less collaborative full stop.

Appendix F: Report on co-operation within the Unit[25]

F.1. Introduction

This is a report of the evaluation I carried out of cooperation within the Unit, specifically with reference to the use of information technology as an enabler of that cooperation, during June 1997. I present my findings as to the way things are within the Unit, and then some thoughts on how they might be better. Many of these points were presented at the team meeting on 1st July 1997, but this report substantiates them and puts them into context.

I must stress that none of what is said here should be seen as authoritative or objective: it is simply one view of one relative outsider as to what's going on. I must also stress that it *is* objective as regards the weight I have given to any person's views: I am not here presenting the views of anyone as having any more significance than anyone else's. The report has been written by me alone, without any further input from anyone than the interviews mentioned in it.

I conducted interviews with most of the members of the Unit (eleven in all), mostly during the week of the 9th-13th June. Most interviews were conducted in the relevant person's office, with others around – this made their conversation less personal, but gave more of a feeling of their public views. I also 'hung around' the various offices (especially the research office as the main public space), to get a feeling for how cooperation takes place. As a guide through the evaluation, I used the methodology I have developed for evaluating cooperative systems, SESL (Systemic Evaluation for Stakeholder Learning). This entails a constant awareness of the whole system of cooperation, including technology but also organisational and personal issues; a focus on the needs and perspectives of all stakeholders involved; and a view of evaluation as the facilitation of part of an ongoing learning process within an organisation.

The key questions I set out to ask in each interview were as follows:

- ♦ Who do you cooperate with, and how?

- ♦ What kinds of cooperation go on?

- ♦ Who are the key stakeholders of cooperation within the Unit? To what extent are their stakes being recognised by the current means of cooperation?

- ♦ What's the role of email, the common filestore (I: drive) and Web pages?

- ♦ Do they help or hinder cooperation?

- ♦ How could the technologies be better used?

The analysis below substantially reflects these questions. Although the initial focus of the study was on cooperation through technology, this is clearly tightly bound-up with cooperation in general, and many of the most interesting issues arising relate to the general case. I therefore first talk about cooperation in general, and then through technology, before presenting some recommendations. Wherever possible I have tried to use direct quotes from Unit members about particular issues, or my immediate notes of their words; I have left these anonymous for the sake of confidentiality, and as the views expressed, while they belong to

---

[25] Names in this report have been changed, but all other parts are as they were given to the Unit.

individuals, to a great extent reflect a more general (although mostly not unanimous) feeling in the Unit.

F.2. Cooperation in general

F.2.1. Who works with whom: reports from interviews

My starting question was to wonder who each person worked with. This is perhaps clear to everyone within the Unit, or perhaps it's confined to certain subsets; it wasn't to me, and so I've recorded here what I heard (plus a few surmises I made). The diagrams below shows what I saw. The first (figure 10) is in the form of a circle, with people roughly grouped by their working groups, but with not much judgement otherwise made about connections (in particular, position on the circle means nothing). The second diagram (figure 11) shows the offices in which people work, and the cooperation within and across those offices.
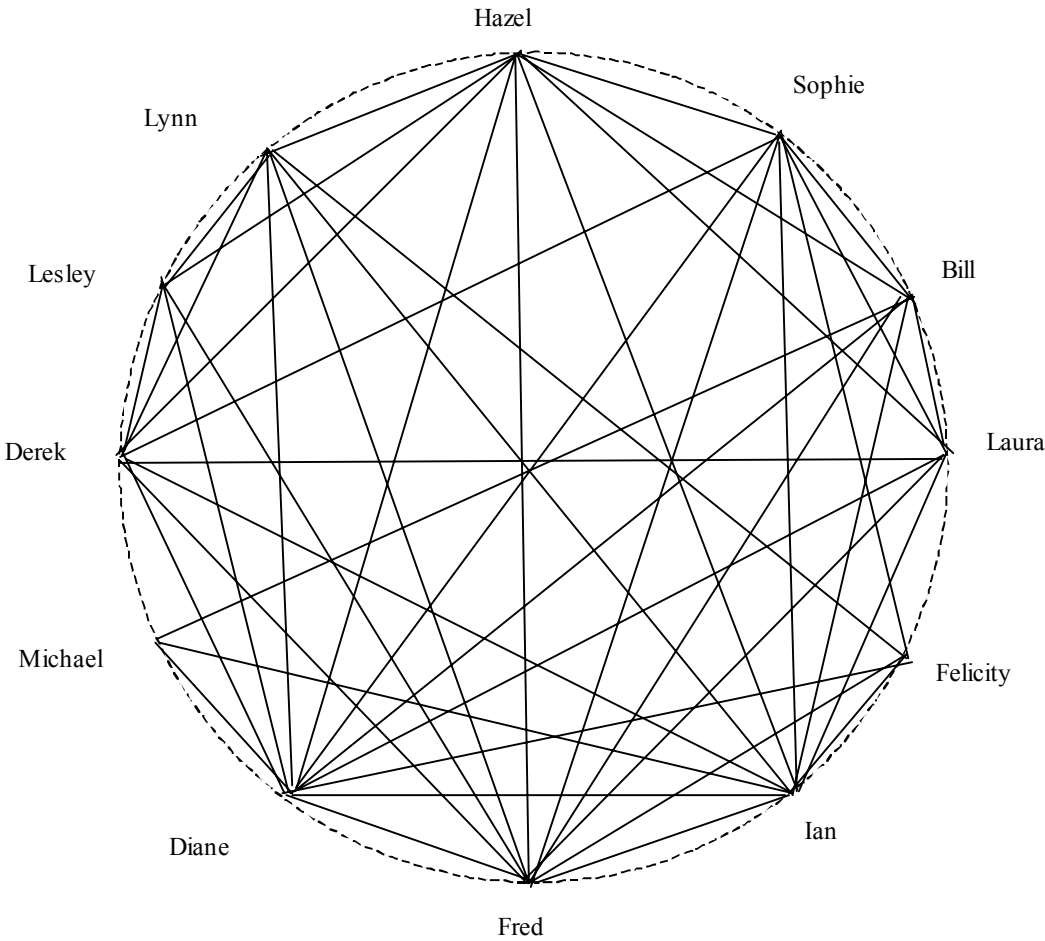
Figure 10: Who works with whom

The lines drawn here only show that there is some kind of working relationship between the two people. Clearly these take many forms, and in particular the amount of work that anyone does with each person with whom they're connected will differ. It would be interested to look at these diagrams with lines drawn in colour or different thickness to reflect these strengths (or by showing an arrow to indicate whether only one person reported working with the other, or both did); but I have not done that here.

Figure 11: Who works with whom (between offices)

The most obvious conclusions to me from the diagrams are:

- notwithstanding the discussion below on the two sides of the Unit, a large amount of interaction actually goes on between the two main offices;

- there are now *five* offices in which the Unit work takes place – for a Unit whose culture was largely established in one office, that's a big shift.
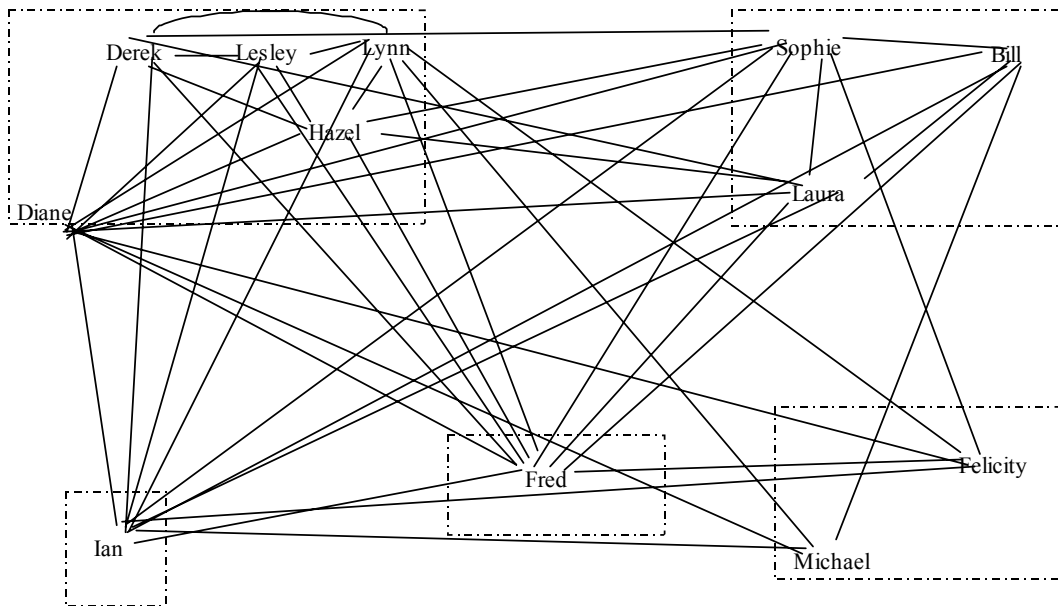
F.2.2.   Key stakeholders

I asked a number of people who they regarded as the key stakeholders of the cooperation in the Unit. This is a word which has many definitions now; I use it to mean those individuals or groups which affect or are affected by the process and outcomes of cooperation in the Unit.

The following groups were mentioned:

- Members of the Unit – one person said that if the Unit (and hence its cooperative processes) don't generate income, no-one has jobs, although a few people mentioned that the rapid growth of the Unit to date (and potential growth in the future) makes it increasingly difficult to respect all the stakes that Unit members have;

- Students on the postgraduate programme [as customers];

- Clients of the Research Fora and other research/consultancy work [as customers, and also as givers of input to various research projects];

- The rest of the School and the University [as suppliers and colleagues – also, as the Unit seems to be rather profitable, as financial recipients];

- Other higher-education institutions [as competitors but also as collaborators];

While nobody mentioned them (perhaps because they were on a list of who I thought were stakeholders that I showed to people), it also seems to me that in the context of information technology (IT), the IT support staff, and some other IT groups elsewhere in the University, are important stakeholders as suppliers.

Why should one wish to make such a list? Partly for the sake of comprehension of what goes on in the Unit, and awareness that there are a number of groups with different kinds of stakes linked to the Unit and its cooperation. Not all of these stakes are necessarily compatible with

123

each other (for example: students, or their companies, would presumably like to pay as low a course fee as possible; the University, with their financial interest, might be happier if higher fees were charged). Some of the groups can be expressed in more detail: I shall describe some of the sub-groups in the Unit, which might well have conflicting stakes. I have certainly observed that some individuals have stakes in the Unit which could (and do) come into conflict with other individuals' stakes (as I don't feel that naming any of those people or detailing their stakes are especially helpful or fair, I shan't pursue this question further).

It's also interesting to reflect on whether the stakes people feel relate to the outcomes of the cooperation from the Unit, or the processes involved in reaching those. Some people felt that the interest that many of the stakeholders had was more to do with what came out of the cooperation than the processes by which it took place, even though these processes may lead to less effective outcomes than otherwise available, and be less satisfying for those involved. "Looking at how one we work" is never comfortable, of course, but perhaps it may sometimes be of value: this report is, I hope, one small part of such a consideration.

F.2.3.   Relations between the two (main) offices

There is a common perception that the Unit is split quite decisively into two parts: the course side and the research side. As will be seen from figure 11 above, there are two core groups among which there is a good deal of cooperation, and between which there is much weaker cooperation. Each of these groups is based in a separate office, and communication between the offices is not very large in general.

When communication does take place, it does so because people from the course office go into the research office – to use the printer or fridge, to pick up stationery or to see the Unit administrator. When this happens, I observed, more informal interactions take place frequently, and this is one chief way that information is passed around and the community maintained. As there are few comparable general resources in the course office, this tends to be the direction in which most traffic occurs.

I had quite a strong feeling that the research office was generally seen as 'the main office' by many people, so that those going there were entering the heart of the Unit. This is presumably especially so because that was previously the only office of the Unit, out of which the course team moved. This is reinforced, of course, by the fact that the shared resources for the Unit are kept there – and also by Diane [the Unit administrator] having her office there, as she is the one person that everyone in the Unit communicates with regularly.

However, the Unit is now based in *five* offices, as I mentioned above: Fred [the director] has his own office, Felicity and Michael share the "broom cupboard" down the corridor, and Ian substantially works at home, and doesn't have a firm base when at the University. This makes the two-teams, two-offices model too simple, especially since some of the those people don't fit neatly into one team or another. Things are more complicated than they seem.

A number of people speculated on how things might be different when the Unit moves location in the summer, and will end up in one room again. I suspect this needs to be given some thought – it should help with informal interactions and passing-on of information (see the next section) and may also help the cohesion of the Unit as a single entity.

F.2.4.   Most cooperation is face-to-face

Almost all communication within the Unit, whether people are in the same office or not, is face-to-face. This was confirmed in the interviews by almost everyone. In particular, the communication tends to be informal, ad-hoc (as and when required) face-to-face conversation.

Someone will turn to their neighbour, shout across the room, or walk into another office, start a conversation about a specific topic or talk for a few minutes, and then leave. Sometimes these conversations are part of an ongoing cooperation on that subject, sometimes they are a request for information or assistance on a very specific issue that ties in with something the person asking is working on alone or with another group.

Of course, face-to-face communication is embedded in a rich set of other tools for cooperation: email, fax, phone, letters, etc. The use of email is detailed in a later section. Most people use all these other tools at various times of the day, but almost always to people outside the Unit. The only exceptions are specific uses of email, for example in relation to meetings to come or other formal occasions; and in communications with Ian when he's at home.

It also seemed to me from my observations, and from interviews, that most people work alone much of the time, and only with others for specific things. Thus the model of a team that applies here is one of individual work that fits together loosely. There are people with a close ongoing working relationship, of course, but much of their day-to-day work seems to be alone. If this is the best way for everyone, this is not necessarily a problem – but it's worth noting that it's so, and at least wondering whether there are other possible options. It does mean, though, that specific roles are not terribly well-defined: people seem to take on roles on a because-they're-there basis, and then they don't relinquish them quickly; but knowing who is responsible for what seems to be difficult for some.

F.2.5. Formal cooperation & information dissemination

There was a general feeling of disquiet among a number of people with the way that more formal cooperation in the Unit had taken place until now, especially in relation to the dissemination of information. A few people said that this was because no-one seemed to take responsibility for such formal meetings. As one person remarked, "the inherited information structure is not contributing to better collaborative working". This seems to have changed since I conducted the interviews (with two meetings taking place, and a third scheduled) – but it is worth making sure that these continue.

This is especially so, as several people remarked, as the number of people in the Unit increases. Ad-hoc communication and information dissemination is an excellent tool for a small group of people all working together closely on a day-to-day basis, which was the starting position of the Unit and still the prevailing culture. However, as the number of people increase, not everyone can talk to everyone else nearly often enough, and this can lead to some people knowing more things than others, and hence gaining power over them, on a haphazard basis. Also, as the amount of *work* increases, this becomes more of a problem. One person remarked that the infrastructure for administering the research fora is a big drain on resources – if there were many more fora, it would just not be possible to maintain this method.

F.2.6. Organisational culture in the Unit and the University

The culture of any organisation will clearly determine the way in which it cooperates. Some of the culture of the Unit is detailed above – informal cooperation, a focus on outcomes rather than processes, lack of clear dissemination channels, the need to be 'in-the-know' and the power from being so; to these I would add the entrepreneurial model of the Unit as a lone group, centred around its founder, fighting to exist and grow in a hostile environment. These are fairly classic patterns for the early stages of small organisations like this one – although as the management theorist Henry Mintzberg makes clear in *The Structuring of Organisations*

(1979), as organisations grow the key to their survival is the ending of some of these features and the taking-up of a more self-aware, structured, cooperation pattern. (This does not mean bureaucracy or inflexibility – as many other management authors have written, a fluidity of structure and ability to change to meet the needs of customers is also vital, especially in a fast-changing world.)

It is also worth considering the culture within which the Unit is based. It seems to me, and this was confirmed by some of the interviews, that the predominant culture of Metropolitan University as a whole is a mixture of old-fashioned 1970s public sector thinking – highly adversarial between bosses and workers, "busy but not excited" (as one person commented); and the 1980s slash-and-burn management thinking, focused on profit and money, regarding people as costs rather than opportunities. This is a fairly deadly combination, which is self-reinforcing and rather potent – and prevalent in much of the University. The result is the sapping of energy, general apathy, and a constant struggle just to survive. No-one has any fun or is especially creative. Cost-cutting is (as in the School administrative staff review) conducted with little thought for the consequences, just as a short-term solution to a crisis. I am painting a highly negative picture here, and of course not every bit of the University has this culture. There are places where innovation and creativity are emphasised – but they seem to be despite, rather than because of, the fact they're in the University.

Aspects of this culture can be seen in the Unit too, inevitably because of its being part of the School. The Unit's comparative isolation from the rest of the School – and its self-perception as a new thing, "an island in a sea of crap against a system which doesn't support entrepreneurship" (as one person said) – definitely helps it to avoid much of this. After a short study like this, I am reluctant to identify too closely the ways in which the Unit falls into the University culture. The lack of open communication is one aspect that I've mentioned above; a couple of people commented on the unwillingness of the Unit at times to think in terms of processes rather than outcomes; someone said that "the Unit's intellectual capital is not being well-maintained": necessary continuing professional development of staff and investment in infrastructure are problems; and another comment was that "we don't have benchmarks of earning per head" (i.e. per total staff member, not member of professional staff). Care needs to be given to these areas, and no doubt there are others too. Work is needed, but the Unit is in a good position to do this.

F.3. Cooperation through technology

F.3.1. The common filestore (I: drive)

The key technological feature of cooperation in the Unit is the I: drive on the PC network, to which everyone has access. Here I shall examine how it is used and structured, how it relates to other drives, etc.

The purpose of the drive is clear – to save documents that happen during the work of the Unit and are needed (or might be needed) by more than one person. One description was: "the main things are general information for the general use of everyone". It is widely used by some people, less by others, but one person said: "I couldn't do without it: it's absolutely vital".

Its uses vary from person to person and task to task. Some people use it as a temporary store to pass files from one person to another (although these are often not deleted subsequently). Others puts things on it that they think are important to others, either actually or potentially. Some people put most of their files on it, with only a few personal files on personal (C: or H:) drives. Others try to avoid I: completely, keeping most files on their personal drives – which

allows them to organise them according to their own taste and know they won't be tampered with, but means other people don't have access to that information. As one person commented, "everyone seems to have their own private hoard".

The structuring of the drive is a bit of a problem. Some effort has gone into structuring the files of the course side into one "course" directory, and these are carefully maintained. The files of the research side, and those pertaining to the Unit in general, "could be structured better" (as one person acknowledged). There is a rough directory structure, but files are not always put in the appropriate place, or given comprehensible names, so that they can only be accessed by those who created them, or have been told about them, or who go searching through the whole drive. For some people, this leads to a great deal of frustration, and there is also a fair amount of 're-inventing the wheel' as a result. Since I conducted the interviews, Diane and Felicity have begun to organise these other files more precisely, and I hope this will continue.

I have mentioned elsewhere the importance in the Unit of being-in-the-know: the amount of secret information that is held on to by its creators and only revealed to a select band of insiders. Given this, the regularity with which the phrase "everybody knows that…" occurred in the interviews was of considerable interest to me. It was fairly clear that these things, which the people who used that phrase quite genuinely believe to be general knowledge, are no such thing. Examples include: files in the I:\ directory beginning with "con" are for contacts and with "let" for letters; the Health Forum members are stored in the file "members.doc"; the location of archived reports, papers, marketing material etc. It is certainly the case that everyone in the course team *does* know what's in the directory "course" and its sub-directories; but no-one else in Unit seems to. Raising this issue might seem petty: if someone doesn't know something, surely all they need to do is ask someone who does? The trouble is that one doesn't always know that one doesn't know the information, and that there may well be a host of useful files that one is therefore missing. It also concentrates power into the hands of those who do know things, which blocks creativity. Of course, this has been made worse by growth – in the early days of the Unit, when they were all in one room, John knew where everything was and could tell it to anyone (or so one person remembers) – but that's not possible now. Further growth can only make this problem worse.

Finally, a brief look at some processes by which the I: drive is used:

- Two people keep a log of all the files they put there, with names, creation dates and subjects.
- It is quite common that a chain of working goes on, where one person (Laura or Fred, in two examples I was given) types out some work roughly with the key ideas they want to be represented in place, then puts it on the I: drive, then another person (Sophie or Diane, in these cases) takes the file and modifies it – either still on I: or on their own C: drive – to add extra information, make it presentable for sending outside, etc.
- One risk of the drive, and its lack of two people being able to work on the same document at once, is that more than one copy can come into existence of the same file – these are then updated separately and inconsistently. This is cumbersome but manageable for the authors of the document, but, as one person said, "it doesn't make it easy in the unit for others to know what's going on".
- It is not uncommon that instead of using the I: drive for transferring files from one person to another, documents are sent by email instead, to avoid "cluttering it up" (the view of one person who's done this).

F.3.2.   Email

Email is little used within the Unit, but much used in communication with people in other parts of the University. It is also little used to communicate outside the University. There are two main ways in which it *is* used within the Unit: to disseminate general information, or for document distribution. A few people said they used email quite a bit to keep people informed about things going on in the Unit, especially dates of meetings and the like, or to remind others of things they need to do. A number of people prefer it as a way to send documents to others (via attachments) than the I: drive – typically, these are documents they are working on together which the sender has done some work on and wishes the receiver to do some more on. One person also reported that they had used email to avoid confrontations, and it seems likely that others will also use it in this way.

It was especially popular, though, for sending messages to people elsewhere in the University. It is used to contact tutors from different Schools, central and IT support staff, and a wide range of others in the School. One person said it formed their main source of what's going on in the rest of the School, and someone else uses the shared folders that are part of Microsoft Mail and contain a range of School-wide information and documents. Email seems to have replaced the telephone as the communication method of choice around much of the University, especially since if it's difficult to find people in the office.

This communication with the rest of the School has problems, though. A large number of the emails received are day-to-day messages sent around about things happening in the School. Comments like "you get so much irrelevant rubbish by email" were typical, and one person estimated that only 10% of these were relevant. This means it takes much longer to work through one's messages, and there's a much greater temptation to dismiss a large bunch as irrelevant.

Outside of the University, there seemed to be little use of email. The general perception seems to be that (with a few exceptions) most of the clients, student, forum members etc. that members of the Unit habitually communicate with are either not on email at all, or seem to have little respect for it. This may be the case, but it may also be worth examining whether there are useful ways that email can be used externally.

F.3.3.   The Web

The much-vaunted and much-hyped World-Wide Web is slowly taking root in the Unit. A few months ago, the Unit's own Web pages were set up, substantially by Lesley. These contain information on courses, research and Unit members. They seem to be a good start, but it is not clear just now how widely they are used or whether anyone from outside has made contact with the Unit through them. However, all the signs seem to be that the Web will increase in its importance, particularly for the kinds of "knowledge workers" that are the Unit's main customers.

A few the Unit seem to be regular users of the Web, and have obtained highly relevant information there; others have looked at it occasionally. However, one person's remark that "I could live without the Web at the moment" seemed typical of the attitudes to the Web. It may well be that it would be useful to increase the amount of use of the Web, as a tool for gaining new sources of information from unexpected areas (one of the key ways that creativity comes about); but I certainly recognise there are risks in this, as it often just as well feels like a set of tiny nuggets of useful information in a sea of rubbish, and so can waste a lot of time.

F.4. Conclusion: how things could be better

Stephen Covey, in his interesting book *The Seven Habits of Highly Effective People* (1989), talks about the importance of looking not just at things that are urgent and important, but also the all-too-often crowded-out items that are not urgent but are still highly important. He includes: "building relationships, writing a personal mission statement, long-range planning, exercising, preventive maintenance, preparation – all those things we know we need to do, but somehow seldom get around to doing, because they aren't urgent" (p.154). But effective people, he says, "are not problem minded; they're opportunity-minded. They feed opportunities and starve problems. They think preventatively. They have genuine … crises and emergencies that require their immediate attention, but the number is comparatively small." I wish to explore in this section some ways in which I think the Unit needs to work on these important-but-not-urgent items that seldom get done.

The recommendations here come from my own observations, and awareness of trends in similar organisations elsewhere; but also from the interviews. I asked each person, in some form or another, the question "how could the technologies be better used?"; and the question of how cooperation could be better in general was a theme that arose spontaneously from many people. This section is therefore a mixture of me and you. Many points here will be fairly clear from the discussion above. The recommendations should be seen as initial and tentative (they're only some ideas about the way things could go).

First, to look at cooperation in general. The shift to the new office brings real opportunities to change the way the Unit works. Firstly, given my above remarks about the University culture, it will separate the Unit from the rest of the University sufficiently to make attempts at developing a distinctive culture that is appropriate to the role of what is effectively a small training, research and consultancy company. It also gives space (literally!) to consider whether the Unit is to be one group or several, as was discussed at the team meeting I attended.

Related to this will be the shift from the Unit as a small startup to the Unit as an ongoing concern. As one person remarked in the interviews, "we need to become an organisation – previously we've not been an organisation". Developing clearer information structures and a more decentralised power-base will be important parts of this. Also important will be a clearer approach to the harnessing of the skills of the Unit's members, by encouraging them in professional development as knowledge workers, and in utilising all their brains to develop new and creative ways of growing the Unit's activities. This may require encouragement in learning; it may require a greater financial investment in the members of the Unit; it may even require a redistribution of some of the financial growth of the Unit to its members. What is crucial, though, is that the motivation of all people in the Unit to produce excellent work and to increase its activities be maintained and stimulated.

Information is also a key aspect to the growth of the Unit. Taking the Research Fora as an example: at the moment there are around twelve meetings a year; as discussed above, administering them is a big drain on resources. If these were to double in numbers, the kind of thing that surely must happen (or its equivalent) over the next few years, then I'm not sure whether they would be manageable. A lot more reuse of information, both content and organisational, would be needed to make it so – basically, a kind of "pro-forma forum" would need to be put together, a template for the organisation of fora into which individual aspects can be slotted. Many other examples of information reuse can be thought of, both in terms of administrative and content of research/teaching – harnessing the potential reuse of information is crucial.

One way of doing this is through technology, so let us turn to my recommendations for technology in the Unit. As I've said, it's an excellent way of spreading organisational knowledge and memory, storing past experiences for use in the future. At the moment, the I: drive partly fulfils that function, but it is limited both by minimal technology and inefficient use.

A lot could be done with an investment for the future in new technology. This might include faster PCs, as many people are currently limited in the usefulness of their computers by the speed of their PCs, but most possibilities will come via software.

Firstly, Windows 95, the successor operating system to Windows 3.11 that currently runs in the Unit, has a number of key benefits. It allows for longer filenames (up to 250 letters rather than the 8 of the present system), which makes the contents of documents much clearer. Next, the dialogue box brought up by the Open File menu option in various programs (Word etc.) is more detailed, containing information on when the document was last updated – this is a small addition but makes quite a difference. The other clear benefit is its stronger support for remote use of the network from other parts of the country or world. the Unit is not just confined to its home city any more, and making the organisational knowledge available to those who are working elsewhere would be very important.

Next, there are now a series of programs available now under the name of "groupware" that are explicitly designed to facilitate computer-supported cooperative working. Most of the major software manufacturers produce useful products, but the most widely-used, and the best, is Lotus Notes. The first function of this program is to be a document database, holding all documents written in the Unit in a similar way to the I: drive, but allowing for greater classification and for useful features like full-text search (to solve the "is there anything on this topic out there?"). This is also better integrated with the built-in email system than the current arrangement. Other facilities for enabling group working are also supported by Notes. Perhaps one of the most obviously useful is the group calendar facility: one can store everyone's appointments on one page, making it clearer when everyone will be around, and easier to book collective meetings. Notes, and similar products, offer a great deal of flexibility and support for a range of group activities.

One technological solution to the 'junk-email' problem is supported by Notes, but also by a number of specialised email programs: the use of filters. Many such programs allow one to specify a number of phrases which, if they come up in a message (or more often its 'headers' – to, from, subject etc.), will cause the message to be put into a separate folder for later reading. This means that all messages sent to "Everyone (School)", for example, could end up in a "School Messages" folder, rather than appearing in the general In folder with everything else. I receive a lot of email of various sorts and find this facility very useful for organising it.

I do recognise that all these suggestions about the technology are expensive: to use Notes, for example, requires quite a cost in paying for the software for each PC, but also in having an expert help you to determine how it is best used, and to set up the program in the most effective way – which is necessary to make it most useful – and it requires some administration to keep it working correctly. Some of the other options are cheaper (e.g. Windows 95 costs about £50 per machine, and setting it up is easier). However, given the projected growth of the Unit, and the unlocking of potential offered by technology, it might really be worth considering as an investment.

If this is not possibly financially, there are a number of ways in which the existing technology could be used better:

- The developing of a well-defined directory structure on the I: drive with a set of conventions for filenames that are known to everyone when they create files, and are then obvious when they want to access files. Setting up these conventions – and ensuring they are adhered to rigorously – is important. I would suggest that these conventions should be sufficiently obvious that they can fit on a single sheet of paper that will be pinned above each person's computer, to act as a reminder. In particular, I suggest that temporary files are kept in a directory specifically labelled thus, and deleted after a time.

- Once the structure and naming conventions are established, someone needs to be made responsible for the I: drive. That is, they need to look at the drive occasionally to make sure the conventions are being followed, and tell those who are flouting them to stop doing so. They should delete all temporary files older than (say) a week, and make sure no temporary files are anywhere but the correct directory for those. This responsibility could be delegated to someone else where there is a clear sub-group of files (like the way the "course" directory already works).

- Files should, by default, be regarded as public and kept on the public filestore. The private drives – C: and H: – should be only be used for personal or sensitive files that others must not have access to. This will allow information to be readily accessible to others when *they* need it, not when the author thinks they might.

- Likewise, making it clear what public, and potentially useful, information one has created – either on computer or on paper – is important. If it's of use to others, they need to know about it. It might be a good idea to keep a record – preferably on computer so it's searchable – of such files. Many organisations do this in two ways: if the document's creator finds it useful for others, they could put it in the file; but also if someone else reads a document created by another person, they can put it in, or add the fact that they found it useful to the file. This rather rapidly builds up a list of documents that were of interest to others, and hence not only identifies the key documents around, but also allows one to ask questions like "I know that person X is interested in such-and-such, and I would like to know a bit about that topic – so what documents has that person created or accessed recently?". All this could easily be done in a Word file or a simple database.

- Finally, training is important. It seemed from the team meeting I attended that some people just didn't know about the full usefulness of the software they had on their computers. Whether new software is introduced or not, better training in specific uses of software might be helpful.

To finish from an image in another management book: Tom Peters and Robert Waterman in *In Search of Excellence* (1982) talk about the phenomenon seen in some of the most successful companies of the "skunkworks". These are small, innovative, groups of people on the fringes of big organisations with lots of flexibility and space for experiment. The Unit, as it becomes the [New Unit], has a real chance of being such a group – *if* there is investment in cooperation, both in the people doing the cooperation and in the tools and culture that supports that cooperation.

F.5.   Summary (flipchart notes from 1st July presentation)

F.5.1.   Findings – general

- Five offices, two groups
    - ⇒ What will happen after move?

- Most cooperation face-to-face
- Little formal cooperation
- Problems with scaling up?
- Dissemination of information: 1) organisational memory; 2) "top-down"
- Most people seem to work alone

F.5.2. Findings – technology

- I: drive
  - Used extensively by some, less so by others
  - General feeling of lack of structure
  - Importance of being-in-the-know
  - Scaling-up?
- Email
  - Commonest use is around the University
    - $\Rightarrow$ Most external communication by phone/letter
  - Big problem with 'junk' mail – limits use

F.5.3. Initial, tentative, recommendations – general

- Shift to new office brings real opportunities
  - $\Rightarrow$ Is the University organisational culture most appropriate?
  - $\Rightarrow$ Bringing the two sides together
  - $\Rightarrow$ "Why not change the way we work?" (BT)
- Information dissemination needs to be clearer

F.5.4. Initial, tentative, recommendations – technology

- Great way of spreading organisational knowledge
- If possible, look at new technology:
  - New PCs; Windows 95 (longer filenames, dates in "open file" etc.)
  - External dialup possibilities
  - Lotus Notes (text search, calendars, classify documents, workflow etc.)
  - Email filtering
  - What possibilities come from the Unit having its own fileserver?
  - Could this be paid for by someone else?
- Whether or not new technology:
  - Need to develop clear filenames, locations
  - Make sure file names & structures are adhered to
  - Publicise what information is available and useful
  - Model of all information being public by default, unless good reason
  - Someone needs to take responsibility for the I: drive filestore
  - Look at training options

References

Abbott, Kenneth and Sunil Sarin (1994). Experiences with Workflow Management: Issues for the Next Generation. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW '94).*

Ackerman, Mark (1994). "Augmenting the Organisational Memory: A Field Study of Answer Garden". *Proceedings of CSCW 94*, pp. 243-252.

Ackoff, Russell (1974). *Redesigning the Future.* London: John Wiley.

Argyle, Michael (1991). *Cooperation: The Basis of Sociability.* London: Routledge.

Argyris, Chris and Donald Schön (1996). *Organisational Learning II: Theory, Method, and Practice.* Reading, MA, USA: Addison-Wesley.

Axelrod, Robert (1984). *The Evolution of Cooperation.* New York: Basic Books.

Bannon, Liam (1991). "From Human Factors to Human Actors: The Role of Psychology and HCI Studies in System Design". In Joan Greenbaum and Morten Kyng (eds.), *Design at Work.* Hillsdale, New Jersey: Lawrence Erlbaum, pp. 25-44.

Bannon, Liam (1994). Use, Design, and Evaluation: Steps towards an integration. In D. Shapiro, M. Tauber & R. Traunmueller (eds.), *The Design of Computer-Supported Cooperative Work and Groupware Systems.* Amsterdam: North-Holland.

Bannon, Liam and Susanne Bødker (1997). Constructing Common Information Spaces. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work, ECSCW 97 (Lancaster, UK, 7-11 September 1997).*

Bateson, Gregory (1972). *Steps to an Ecology of Mind*, San Francisco: Chandler.

Beck, E. (1994). *Practices of collaboration in writing and their support.* CSRP No. 340, School of Cognitive and Computing Sciences, University of Sussex. (DPhil thesis).

Benford, Steve, Chris Greenhalgh, Dave Snowdon and Adrian Bullock (1997). Staging a Public Poetry Performance in a Collaborative Virtual Environment. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work, ECSCW 97 (Lancaster, UK, 7-11 September 1997).*

Bertalanffy, Ludwig von (1969). *General system theory: foundations, development, applications.* New York: Braziller.

Bikson, Tora K. and J. D. Eveland (1996). Groupware Implementation: Reinventing the Socio-technical Frame. *Proceedings of the ACM Conference on Computer Supported Co-operative Work, CSCW 96* (Boston, MA, USA, 16-20 November 1996).

Blandford, A.E. and Duke, D.J. Integrating user and computer system concerns in the design of interactive systems. *International Journal of Human-Computer Studies*, 1997, Vol. **46**, No. 5, pp. 653-679.

Blomberg, Jeanette (1986). "The Variable Impact of Computer Technologies on the Organisation of Work Activities". *Proceedings of CSCW 86* (Austin, Texas), pp. 35-42.

Bono, Edward de (1991). *I am right – you are wrong : from this to the new renaissance : from rock logic to water logic.* Harmondsworth, UK : Penguin.

Bowers, John, Graham Button and Wes Sharrock (1995). Workflow from Within and Without: Technology and Cooperative Work on the Print Industry Shopfloor.

Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work, ECSCW 95 (Stockholm, 10-14 September 1995), pp. 51-66.

Brown, John Seeley and Paul Duguid (1991). Organisational learning and communities-of-practice: Toward a unified view of working, learning and innovation. *Organisation Science* 2:40-57.

Burrell, Gibson and Gareth Morgan (1979). *Sociological Paradigms and Organisational Analysis*. London: Heinemann.

Button, Graham and Wes Sharrock (1997). The Production of Order and the Order of Production. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work*, ECSCW 97 (Lancaster, UK, 7-11 September 1997).

Card, Stuart, Thomas Moran and Allen Newell (1983). *The Psychology of Human-Computer Interaction.* Hillsdale, New Jersey: Lawrence Erlbaum.

Carroll, J. M. (1995a). Introduction: The scenario perspective on system development. In Carroll (1995b), op.cit.

Carroll, J. M. (Ed.). (1995b). *Scenario-based Design: Envisioning work and Technology in System Development*. New York: John Wiley.

Checkland, Peter (1981). *Systems Thinking, Systems Practice*. Chichester: John Wiley

Checkland, Peter and Jim Scholes (1990). *Soft Systems Methodology in Action.* Chichester: John Wiley.

Checkland, Peter and Sue Holwell (1998). *Information, systems, and information systems : making sense of the field*. Chichester: John Wiley.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.), *Perspectives on Socially Shared Cognition*, pp. 127--149. Washington DC, USA: American Psychological Association.

Clark, Liz and Fides Matzdorf (1998). *Motivation and Multi-skilling of Service Staff in NHS Trusts.* Research Report, Facilities Management Graduate Centre, Sheffield Hallam University.

Clement, Andrew (1990). Cooperative Support for Computer Work: A Social Perspective on the Empowering of End Users. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW'90)*, pp. 223-236.

Covey, Stephen (1989). *The Seven Habits of Highly Effective People*. London: Simon & Schuster.

CSMIL (1992). *ShrEdit: A Shared Editor for the Macintosh*. Computing Science and Machine Intelligence Laboratory, University of Michigan.

Cullen, J., J. Kelleher & E. Stern (1993). Evaluation in DELTA. *Journal of Computer Assisted Learning* 9:115-126.

Cusumano, M and R Selby (1995). *Microsoft Secrets*. New York: Free Press.

Dobson, J.E., A.J.C. Blyth, J. Chudge and R. Strens (1994). The ORDIT approach to organisational requirements. In M. Jirotka & J. Goguen (eds.), *Requirements Engineering: Social and Technical Issues.* London: Academic Press.

Donaldson, Thomas and Lee Preston (1995). The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. *Academy of Management Review* 20(1):65-91.

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. *Proceedings of CSCW '92 Conference*, Toronto.

Eason, Ken (1988). *Information Technology and Organisational Change.* London: Taylor & Francis.

Easterbrook, S. (1994). *Coordination breakdowns: Why groupware is so difficult to design.* CSRP no. 343, School of Cognitive and Computing Sciences, University of Sussex.

Easterby-Smith, Mark (1994). *Evaluating Management Development, Training and Education (2nd edition).* Aldershot, UK: Gower.

Ehn, Pelle and Morten Kyng (1987). The Collective Resource Approach to Systems Design, in Gro Bjerknes, Pelle Ehn and Morten Kyng (eds.), *Computers and Democracy*, Aldershot, UK: Avebury, pp. 17-58.

Erickson, T. (1990). Working with interface metaphors. In Laurel, B. (Ed.), *The Art of Human-Computer Interface Design*. Reading, MA, USA: Addison-Wesley:.

Evan, William and R. Edward Freeman (1993). A Stakeholder Theory of the Modern Corporation: Kantian Capitalism. In Tom Beauchamp & Norman Bowie (eds.), *Ethical Theory and Business.* Englewood Cliffs, NJ, USA: Prentice Hall.

Fanning, Tony and Bert Raphael (1986). Computer Teleconferencing: Experience at Hewlett-Packard. In Irene Greif (Ed.), *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 86)*, pp. 291-306.

Finkelstein, Anthony & Ian Sommerville (1996). The Viewpoints FAQ. *Software Engineering Journal* 11(1): 2-4.

Fisher, Roger and William Ury (1991). *Getting to Yes: Negotiating agreement without giving in* (2nd ed.). London : Century Business.

Freeman, R. Edward and David Reed (1983). Stockholders and Stakeholders: A New Perspective on Corporate Governance. *California Management Review*, **25**(3): 88-106.

Garfinkel, Harold (1967). *Studies in Ethnomethodology*. Englewood Cliffs, New Jersey: Prentice-Hall.

Geus, Arie de (1997). *The Living Company: Growth, Learning and Longevity in Business*. London: Nicholas Brealey.

Giddens, Anthony (1979). *Central problems in social theory: action, structure and contradiction in social analysis*. London: Macmillan.

Goodman, G. & M. Abel (1986). Collaboration research in SCL, in *Proceedings of CSCW'86*, pp. 246-251.

Green, Eileen, Jenny Owen and Den Pain (1991). Office Systems Development and Gender: Implications for Computer-Supported Cooperative Work. *Proceedings of the Second European Conference on Computer-Supported Cooperative Work (ECSCW 91)*, pp. 33-48.

Green, T. (1991). Describing information artefacts with cognitive dimensions and structure maps. In Diaper, D., & Hammond, N. (Eds.), *People and Computers VI, Proceedings of the 6th BCS HCI Specialist Group*. CUP: Cambridge.

Greenbaum, Joan and Morten Kyng (1991). Introduction: Situated Design. In Joan Greenbaum and Morten Kyng (eds.), *Design at Work*. Hillsdale, NJ, USA: Lawrence Erlbaum, pp. 1-24.

Grinter, Rebecca E. (1997). Doing Software Development: Occasions for Automation and Formalisation. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work, ECSCW 97* (Lancaster, UK, 7-11 September 1997).

Grudin, Jonathan (1988). Why CSCW Applications Fail: Problems in the Design and Evaluation of Organisational Interfaces. *Procs. of the Conference on Computer-Supported Cooperative Work (CSCW 88)*.

Grudin, J. (1989). The case against user interface consistency. *Communications of the ACM*, 32 (10), 1164 – 1173.

Grudin, Jonathan (1993). Interface. *Communications of the ACM*, **36** (4): 112-119.

Grudin, Jonathan and Leysia Palen (1995). Why Groupware Succeeds: Discretion or Mandate? *Proceedings of the Fourth European Conference on Computer-Supported Co-operative Work, ECSCW 95 (Stockholm, 10-14 September 1995)*.

Gutwin, Carl, Mark Roseman, and Saul Greenberg (1996). A Usability Study of Awareness Widgets in a Shared Workspace Groupware System. *Proceedings of the ACM Conference on Computer Supported Co-operative Work, CSCW 96* (Boston, MA, USA, 16-20 November 1996).

Guy, Ken, Paul Quintas, Michael Hobday, Luke Georghiou, Hugh Cameron and Tim Ray (1991). *Evaluation of the Alvey Programme for Advanced Information Technology.* London: HMSO.

Habermas, Jürgen (1971). *Knowledge and Human Interests.* Boston: Beacon Press.

Halverson, C. (1994). *Distributed cognition as a theoretical framework for HCI: Don't throw the baby out with the bath water – the importance of the cursor in air traffic control.* Technical report No. 94-03, Department of Cognitive Science, University of California, San Diego.

Hammer, Michael and James Champy (1993). *Reengineering the Corporation: A Manifesto for Business Revolution*. London: Nicholas Brealey.

Harper, R. & K. Carter (1994). 'Keeping People Apart.' *Journal of Computer Supported Cooperative Work* 2 (3) pp. 199-207.

Harper, Richard (1992). 'Looking at ourselves: an examination of the social organisation of two research laboratories', in *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 92)*, pp. 330-337.

Harper, Richard H.R. (1997). Gatherers of Information: Some Remarks on the Mission Process at the International Monetary Fund. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work*, ECSCW 97 (Lancaster, UK, 7-11 September 1997).

Harper, Richard, John Hughes and Dan Shapiro (1991). "Harmonious Working and CSCW: Computer technology and air traffic control". In John Bowers and Steve Benford (eds.), *Studies in Computer Supported Cooperative Work: Theory, Practice and Design*, Amsterdam: North Holland, pp. 225-234. (Edited Proceedings of EC-CSCW 89.)

Heath, Christian and Paul Luff (1991). Collaborative Activity and Technological Design: Task Coordination in London Underground Control Rooms. *Procs. of the Second European Conference on Computer-Supported Cooperative Work (ECSCW 91).*

Heath, C. & P. Luff (1992). 'Collaboration and control: crisis management and multimedia technology in London underground control rooms.' *Journal of Computer Supported Cooperative Work* 1 (1/2) pp. 69-94.

Heath, C., M. Jirotka, P. Luff & J. Hindmarsh (1993). 'Unpacking collaboration: the interactional organisation of trading in a city dealing room', in *Proceedings of ECSCW'93,* pp. 155-170.

Heath, Christian, Paul Luff & Abigail Sellen (1995). Reconsidering the Virtual Workplace: Flexible Support for Collaborative Activity. *Proceedings of the Fourth European Conference on Computer-Supported Co-operative Work, ECSCW 95 (Stockholm, 10-14 September 1995).*

Howard, Robert (1987). Systems design and social responsibility: the political implications of "computer-supported cooperative work". *Office: Technology and People*, **3**:175-187.

Hughes, J., D. Randall & D. Shapiro (1993). 'From ethnographic record to system design. Some experiences from the field.' *Journal of Computer Supported Cooperative Work* 1 (3) pp. 123-141.

Hughes, J., King, V., Rodden, T., & Andersen, H. (1994). Moving out from the control room: Ethnography in system design. In *Proceedings of CSCW '94*, North Carolina, pp. 429--439.

Hutchins, Edwin (1991). Organising work by adaptation. *Organisational Science*, **2**(1): 14-39.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA, USA: MIT Press.

Hutchins, E., and Klausen, T. (1996). Distributed cognition in an airline cockpit. In Engeström, Yrjö and David Middleton (eds.), *Cognition and communication at work.* Cambridge: Cambridge University Press.

Hutton, Will (1995). *The State We're In*. London: Jonathan Cape.

Ichikawa, Yusuke, Ken-ichi Okada, Giseok Jeong, Shunsuke Tanaka & Yutaka Matsushita (1995). MAJIC Video conferencing System: Experiments, Evaluation and Improvement. *Proceedings of the Fourth European Conference on Computer-Supported Co-operative Work, ECSCW 95 (Stockholm, 10-14 September 1995).*

Ishii, Hiroshi, Kazuho Arita and Takashi Yagi (1993). "Beyond Videophones: Team Workstation-2 for Narrowband ISDN". In *Proceedings of ECSCW '93*, pp. 325-340.

Jefferson, G. (1984). *Caricature versus detail: On capturing the particulars of pronunciation in transcripts of conversational data*. Tilburg papers on language and literature No. 31, University of Tilburg, the Netherlands.

Jungk, Robert and Norbert Müllert (1996). *Future Workshops: How to Create Desirable Futures*. London: Institute for Social Inventions.

Jurison, Jaak (1994). Measurement and evaluation of IT benefits: a stakeholder-based approach. In *Proceedings of the First European Conference on IT Investment Evaluation.* Birmingham: Operational Research Society.

Kant, Immanuel (1990). *What is Enlightenment?* Published with *Foundations of the Metaphysics of Morals*. London: Macmillan. Original publication 1784.

Kensing, F., & Madsen, K. (1991). Generating Visions: Future workshops and metaphorical design. In Greenbaum, J., & Kyng, M. (Eds.), *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Klöckner, Konrad, Peter Mambrey, Markus Sohlenkamp, Wolfgang Prinz, Ludwin Fuchs, Sabine Kolvenbach, Uta Pankoke-Babatz & Anja Syri (1995). POLITeam: Bridging the Gap between Bonn and Berlin for and with the Users. *Proceedings of the Fourth European Conference on Computer-Supported Co-operative Work, ECSCW 95 (Stockholm, 10-14 September 1995).*

Kotonya, G. (1994). *A Viewpoint-Oriented Method for Requirements Definition*. PhD thesis, Department of Computing, University of Lancaster.

Krauss, R. M., & Fussell, S. R. (1991). Constructing shared communicative environments. In Resnick, L. B., Levine, J. M., & Teasley, S. D. (Eds.), *Perspectives on Socially Shared Cognition*, pp. 172--200. Washington DC, USA: American Psychological Association.

Kuutti, Kari and Tuula Arvonen (1992). "Identifying Potential CSCW Applications by Means of Activity Theory Concepts: A Case Example", in *Proceedings of CSCW 92*, Toronto.

Latour, B. (1993). *We have never been modern* . Cambridge, Mass. : Harvard University Press.

Lave, Jean (1993). Introduction. In Chaiklin, Seth and Jean Lave (eds.), *Understanding Practice*. Cambridge: Cambridge University Press.

Lewis, Paul (1994). *Information-systems development: systems thinking in the field of information-systems*. London: Pitman.

Lotus Development Corporation (1998). *Compare Notes and Domino.* http://www.lotus.com/home.nsf/welcome/compare.

Macaulay, Linda (1994). Cooperative requirements capture: Control Room 2000. In M. Jirotka & J. Goguen (eds.), *Requirements Engineering: Social and Technical Issues.* London: Academic Press.

Machell, Joan and Murray Saunders (eds.) (n.d.) *MEDA: an evaluation tool for training software*. CSET, University of Lancaster.

Mark, Gloria, Ludwin Fuchs and Markus Sohlenkamp (1997). Supporting Groupware Conventions through Contextual Awareness. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work, ECSCW 97* (Lancaster, UK, 7-11 September 1997).

Mason, Robert and Mitroff, Ian (1981). *Challenging Strategic Planning Assumptions*, Chichester: Wiley.

Mintzberg, Henry (1979). *The Structuring of Organisations*. Englewood Cliffs, NJ, USA: Prentice-Hall.

Moran, T. P., & Anderson, R. J. (1990). The workaday world as a paradigm for CSCW design. In *Proceedings of CSCW '90*, Los Angeles.

Morgan, Gareth (1986). *Images of Organisation*. Thousand Oaks, CA, USA: Sage.

Muller, M. (1993). PICTIVE: Democratising the dynamics of the design session. In Schuler, D., & Namioka, A. (Eds.), *Participatory Design: Principles and Practices*. New Jersey: Lawrence Erlbaum Associates.

Mumford, E. (1994). Born again consultants?. *British HCI group newsletter* No. 25, British HCI Group.

Nardi, Bonnie A. (1996). *Context and Consciousness: Activity Theory and Human-Computer Interaction*. Cambridge, MA, USA: MIT Press.

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In *Proceedings of CHI'92*, Monterey, pp. 373--380. ACM Press.

Nielsen, Jakob (1993). *Usability Engineering*. London: Academic Press.

Nonaka, Ikujiro and Hirotaka Takeuchi (1995). *The Knowledge-Creating Company*. Oxford: Oxford University Press.

Norman, D. (1986). Cognitive engineering. In Norman, D., & Draper, S. (Eds.), *User-centred Systems Design: New Perspectives on HCI*. Hillsdale, NJ, USA: Lawrence Erlbaum.

O'Day, Vicki L., Daniel G. Bobrow and Mark Shirley (1996). The Social-Technical Design Circle. *Proceedings of the ACM Conference on Computer Supported Co-operative Work, CSCW 96* (Boston, MA, USA, 16-20 November 1996).

Olson, Gary and Judith Olson (1991). "User-Centred Design of Collaboration Technology", in *Journal of Organisational Computing*, 1(1): 61-83.

Orlikowski, Wanda (1992). Learning from Notes: organisational issues in groupware implementation. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 92)*, pp. 197-370.

Orlikowski, Wanda and Debra Gash (1994). "Technological Frames: Making Sense of Information Technology in Organisations". *ACM Transactions on Information Systems* **12**(2):174-207 (April 1994).

Pagani, D. S., & Mackay, W. E. (1993). Bringing media spaces into the real world. In *Proceedings of ECSCW '93*, Milano, pp. 341--356.

Patton, Michael Quinn (1980). *Qualitative Evaluation Methods*. Beverley Hills, CA, USA: Sage.

Patton, Michael Quinn (1981). *Creative Evaluation*. Beverley Hills, CA, USA: Sage.

Pedler, Mike, John Burgoyne and Tom Boydell. *The Learning Company: a strategy for sustainable development*. London: McGraw-Hill, 1991.

Peters, Tom and Robert Waterman (1982). *In Search of Excellence*. New York: Harper and Row.

Plowman, L. (1992). *Talking and writing in a group writing task: A sociocognitive perspective*. CWRG report No. 5, School of Cognitive and Computing Sciences, University of Sussex.

Plowman, L. (1994). The interfunctionality of talk and text. *Computer Supported Cooperative Work: An International Journal*, **3**(3/4): 229-246.

Plowman, Lydia, Yvonne Rogers and Magnus Ramage (1995). What are workplace studies for? In *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work, ECSCW 95 (Stockholm, 10-14 September 1995)*, pp. 309-324.

Plowman, Lydia, Richard Harper and Yvonne Rogers (eds.) (1996), *The 'Professional Stranger'*. CSRP 428, University of Sussex.

Polson, P., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: a methodology for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36, 741--773.

Price, If and Ray Shaw (1998). *Shifting the Patterns: Breaking the memetic codes of corporate performance*. Chalford, Gloucestershire: Management Books 2000.

Prinz, Wolfgang and Sabine Kolvenbach (1996). Support for Workflows in a Ministerial Environment. *Proceedings of the ACM Conference on Computer Supported Co-operative Work, CSCW 96* (Boston, MA, USA, 16-20 November 1996).

Prinz, W. and Syri, A. (1997). Two complementary tools for the cooperation in a ministerial environment. *Journal of Universal Computing Science*, August 1997, pp 843-864

Ramage, M. (1994). *Engineering a smooth flow? A study of workflow software and its connections with business process reengineering.* MSc dissertation, University of Sussex.

Ramage, Magnus (1996). Hanging around doesn't mean sitting on the fence. In Lydia Plowman, Richard Harper and Yvonne Rogers (eds.), *The 'Professional Stranger'.* CSRP 428, University of Sussex.

Ramage, Magnus and Keith Bennett (1998). Maintaining Maintainability. *Proceedings of the International Conference on Software Maintenance* (Washington DC, 16-20 November 1998).

Randall, D., M. Twidale and R. Bentley (1996). Dealing with Uncertainty: Perspectives on the evaluation process. In P. Thomas (ed.), *CSCW Requirements and Evaluation.* London: Springer-Verlag.

Rhenman, Eric (1965). *Industrial Democracy and Industrial Management.* London: Tavistock.

Roberts, Bill (1996). Groupwar Strategies: Six key technologies will tell you if you need Notes OR the Web or Notes AND the Web. *Byte*, July 1996.

Rogers, Y. (1992). Ghosts in the network: Distributed troubleshooting in a shared working environment. In *Proceedings of CSCW '92*, Toronto, pp. 346--355.

Rogers, Yvonne (1994). Exploring obstacles: integrating CSCW in evolving organisations. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 94)*, pp. 67-78.

Rogers, Y., Bannon, L., & Button, G. (1994). Rethinking theoretical frameworks for HCI. *SIGCHI Bulletin*, 26 (1), 28--30.

Rogers, Yvonne and Judi Ellis (1994). "Distributed Cognition: an alternative framework for analysing and explaining collaborative working". *Journal of Information Technology* **9**(2).

Ross, Susi, Magnus Ramage and Yvonne Rogers (1995). PETRA: Participatory Evaluation Through Redesign And Analysis. *Interacting With Computers*, **7** (4): 335-360.

Rowley, D. E., & Rhodes, D. G. (1992). The cognitive jogthrough: a fast-paces user interface evaluation procedure. In *Proceedings of CHI '92*, Monterey, pp. 389--396.

RSA (1995). *Tomorrow's Company*. Report of an Inquiry by the Royal Society for the encouragement of Arts, Manufactures and Commerce, London.

Ruhleder, Karen and Brigitte Jordan (1997). Capturing Complex, Distributed Activities: Video-Based Interaction Analysis as a Component of Workplace Ethnography. Pp. 246-275 in Lee, Allen S., Jonathan Liebenau and Janice I. De Gross, *Information Systems and Qualitative Research*. London, UK: Chapman and Hall.

Schein, Edgar (1996). *Organisational Learning: What is New?* Working paper 12, Society for Organisational Learning, Massachusetts Institute of Technology, USA. http://learning.mit.edu/res/wp/10012.html.

Schmidt, K., & Simone, K. (1993). Computational mechanisms of interaction for CSCW. *Comic Deliverable 3.1*, Esprit project 6225.

Schuler, D., & Namioka, A. (Eds.). (1993). *Participatory Design: Principles and Practices*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Scriven, Michael (1981). *Evaluation Thesaurus* (3rd edition). Inverness, CA, USA: Edgepress.

Scriven, Michael (1996). Types of evaluation and types of evaluator. *Evaluation Practice*. Vol. 17, No. 2, pp. 151-161.

Senge, Peter (1990). *The fifth discipline: the art and practice of the learning organisation.* New York: Doubleday.

Sharples, M., Goodlet, J., Beck, E., Wood, C., Easterbrook, S., & Plowman, L. (1993). Research issues in the study of computer supported collaborative writing. In Sharples, M. (Ed.), *Computer Supported Collaborative Writing*. London: Springer-Verlag.

Sherry, John (1995). Cooperation and Power. *Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work, ECSCW 95 (Stockholm, 10-14 September 1995)*, pp. 67-82.

Sommerlad, Elizabeth et al. (1992). *A guide to local evaluation*. Evaluation Development and Review Unit, Tavistock Institute of Human Relations, London.

Sommerville, Ian, Richard Bentley, Tom Rodden and Peter Sawyer (1993). *Cooperative Systems Design*. Report CSCW/10/93, Computing Department, Lancaster University, UK.

Sommerville, Ian, Pete Sawyer, and Stephen Viller (1998), Viewpoints for requirements elicitation: a practical approach. *Proceedings of the IEEE International Conference on Requirements Engineering* – ICRE'98, April 6-10, 1998, Colorado Springs, USA.

Sproull, L., & Keisler, S. (1991). *Connections: New Ways of Working in the Networked Organisation*. MIT Press: Cambridge, MA.

Stake, Robert E. (1995). *The art of case study research*. London: Sage.

Star, Susan Leigh and Karen Ruhleder (1994). Steps towards an ecology of infrastructure: complex problems in design and access for large-scale collaborative systems. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 94)*. pp. 253-264.

Stefik, M., Bobrow, D. G., Foster, G., Lanning, S., & Tatar, D. (1987). WYSIWYS revised: Early experiences with multi-user interfaces. *ACM Transactions on Office Information Systems*, 5 (2), 147--167.

Stern, Eliot (1991). "The Evaluation of Policy and the Politics of Evaluation". In *Tavistock Institute Annual Review 1990*, pp. 28-30.

Stern, Eliot (1993). "The challenge of 'real-time' evaluation". In *Tavistock Institute Annual Review 1992/3*, pp. 38-41.

Suchman, L. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge.

Suchman, Lucy and Randall Trigg (1991). "Understanding Practice: Video as a medium for reflection and design". In Joan Greenbaum and Morten Kyng (eds.) (1991), *Design at Work*. Hillsdale, New Jersey: Lawrence Erlbaum.

Tang, John, Ellen Isaacs and Monica Rua (1994). "Supporting Distributed Groups with a Montage of Lightweight Interactions". *Proceedings of CSCW 94* (Chapel Hill, North Carolina, USA), pp. 23-34.

Tognazzini, Bruce (1992). *Tog on Interface*. Reading, MA, USA: Addison-Wesley.

Turner, Phil and Susan Turner (1997). Supporting cooperative working using shared notebooks. *Proceedings of the Fifth European Conference on Computer Supported Co-operative Work, ECSCW 97* (Lancaster, UK, 7-11 September 1997).

Twidale, Michael, David Randall and Richard Bentley (1994). Situated evaluation for cooperative systems. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 94)* , pp. 441-452.

Urquijo, Silvia Ponguta, Stephen Scrivener and Hilary Palmen (1993). "The Use of Breakdown Analysis in Synchronous CSCW System Design". In *Proceedings of ECSCW '93*, pp. 281-294.

Vidgen, Richard and Tom McMaster (1995). Black Boxes, Non-Human Stakeholders and the Translation of IT Through Mediation. In Wanda Orlikowski, Geoff Walsham, Matthew Jones and Janice DeGross (eds.), *Information Technology and Changes in Organisational Work*, London: Chapman & Hall, pp. 250-271.

Wadsworth, Yoland (1991). *Everyday Evaluation on the Run.* Melbourne: Action Research Issues Association.

Walsham, Geoff (1993). *Interpreting Information Systems in Organisations.* Chichester: John Wiley.

Wan, Dadong and Philip Johnson (1994). Computer Supported Collaborative Learning Using CLARE: The Approach and Experimental Findings. *Proceedings of the Conference on Computer-Supported Cooperative Work (CSCW 94).*

Wann, T.W. (ed.) (1964). *Behaviourism and Phenomenology: Contrasting Bases for Modern Psychology*. Chicago: Phoenix Books.

Winograd, Terry and Fernando Flores (1986). *Understanding Computers and Cognition.* Reading, MA: Addison-Wesley.

Woofitt, Robin (1991). "On the Analysis of Interaction: an Introduction to Conversational Analysis". In Paul Luff, Nigel Gilbert and Dave Frohlich (eds.), *Computers and Conversation*. Academic Press: London.

Young, R. (1983). Surrogates and mappings: two kinds of conceptual models for interactive devices. In Gentner, D., & Stevens, A. (Eds.), *Mental Models*, pp. 35 -52. Hillsdale, NJ: LEA.

Zuboff, S. (1988) *In the Age of the Smart Machine: The Future of Work and Power*. Oxford: Heinemann.