A Socio-Cognitive Theory of Information Systems and Initial Applications

by

Christopher J. Hemingway

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy.

The University of Leeds
School of Computer Studies

August 1999

The candidate confirms that the work submitted is his own and the appropriate credit has been given where reference has been made to the work of others.

Abstract

Much has been written in the academic literature about designing information systems (IS) to satisfy organizational, rather than purely technical, objectives. The design of systems to address the requirements of end-users has also received considerable attention. Little has been said, however, about the relationship between these two facets of "best practice" and how they might be reconciled. This is of concern because the relationship is fundamental to the success of organizational systems, the value of which is ultimately realized through the activities of individuals and workgroups. The practical benefit of achieving an integrated approach is clear. Systems can be developed in light of the relationship between worker and organization, rather than as a result of a compromise between two 'competing' viewpoints. An integrated theory would also reduce the conceptual distance between current conceptions of individuals, which tend to downplay their status as social beings, and of social organizations, which often overestimate the influence of social organization on individuals' actions. A process of conceptual analysis and theory development that addresses this disjunction is presented in this thesis.

As the main contribution of this research, the socio-cognitive theory of information systems is a first attempt at providing an integrated treatment of IS phenomena. The theory is developed using a dialectic research method by drawing upon existing work in human-computer interaction, information systems, psychology and sociology. Following a consideration of dialectic as a research method, it is applied to existing conceptions of the individual and of social organization in these disciplines. The theory is then constructed to provide an explanation of information systems phenomena in socio-cognitive, rather than social and cognitive, terms.

Having presented the theory, its potential contribution to realizing the practical benefits of integrated approaches to IS development is illustrated through the development of a systems development lifecycle and an evaluation methodology. Recognizing that IS development is primarily concerned with the relationship between individuals and social organizations, the lifecycle model focuses attention on addressing skills issues during the development process. Extending the focus on skills and intersubjective communication, the evaluation methodology outlines a method, consistent with the socio-cognitive theory, for analysing working practices and assessing the impacts upon them of IS-related change.

Acknowledgements

First and foremost, I thank Tom Gough, my supervisor, for his guidance and support throughout my doctoral studies. He is a consummate teacher and a valued friend.

I am also very grateful to Simon and (soon-to-be) Sarah Rollinson for their help, support and good humour over the last couple of years and to my other friends, especially Robert Hernik, Graeme Tennyson and Jawahar Patankar.

I thank my fellow students in the School of Computer Studies, particularly those with whom I have shared the ORIS lab (and the ups and downs of research) since 1995: Abdullah Akbar, Will Barrett, Keramatollah Bassiri, Jackie Carter, Simon Horton, William Hsu, Abdulwahed Khalfan, Sharifullah Khan, Allah Nawaz, Bobby Sharma, Yindong Shen, Lisa Simpson, Valéria Times and Ryan Wang. I am also grateful for the help and support I have received from many of the staff in the School of Computer Studies, particularly those from the IS research group.

I have been fortunate in meeting many IS researchers who have provided me with ideas for further developing my research. Particular thanks are owed to Peter Mott of the School of Computer Studies, University of Leeds for his advice on the philosophical aspects of this research; Professor Niels Bjϕrn-Andersen of Copenhagen Business School for his helpful discussion of my paper at the IFIP WG8.2/8.6 conference; Dr Susan Gasson of SUNY Binghamton for advice on exploring the social psychology literature; and Professor John Ward of Cranfield School of Management for inviting me to attend his course on benefits management.

Finally, I would like to thank my parents, sisters Rachel and Chloë, step-brother Ben and the rest of my family for their patience, support and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Both commercial and non-profit making organizations now invest a significant and increasing proportion of their revenue into the development and use of computer-based information systems. Price Waterhouse, for example, found that companies spent an average of 1.5% of their turnover on IT in 1992/93 and 1.8% in 1995/96 (Price Waterhouse 1992, Price Waterhouse 1995). Computer Economics found that IT budgets increased by an average of 4.9% in 1998 for firms with revenues exceeding US$100 million (Computer Economics 1998b) and projected a global IT spend of almost US$1 trillion in 1999 (Computer Economics 1998a). Looking at the wider economic picture, the rapid expansion of the Internet (Butler, Barker and Levitin 1996) and developments in mobile information and communication technologies (ICTs) (Puuronen and Savolainen 1997) and electronic commerce (Edelheit and Miller 1997) suggest that the continued exploitation of ICTs is likely to have significant social and economic implications.

To gain a full appreciation of the societal impacts of ICTs, it is not sufficient merely to look at broad economic indicators. Even at the organizational level, it has been recognized that the economic implications are extremely difficult to evaluate. Financial analyses indicate efficiency gains that are often unremarkable and belie the radical impacts of new technologies on working practices and power relations within and between organizations. It is because the impacts of ICTs are frequently not measurable in terms of efficiency that accounting based techniques using the classical economic theory of the firm have proved largely ineffectual for corporate decision makers (Strassmann 1990, Willcocks, Lacity and Fitzgerald 1995). As indicated by the increasing volume of research addressing the social and organizational impacts of ICT-based systems, a plausible alternative would be to draw upon psychological and social theories to supplement the technical and economic analyses that underpin much human-computer interaction (HCI), information systems (IS) and software engineering (SE) evaluation research. Such an eclectic approach is adopted here in an attempt to establish a sound theoretical basis for explaining combined psychological, socio-economic and technical changes and to exploit such theory to propose an alternative approach to the evaluation of ICT-based information systems.

1.2 Background

1.2.1 Current Perspectives on the Information Systems and Software Development Lifecycles

In the 1970s, the need to develop increasingly complex software systems led to the development of various structured methods for systems design (Yourdon 1972, DeMarco 1979). These methods typically embodied the conventional 'waterfall' model (Royce 1970) and regarded analysis as the task of identifying the correct structure to encode the problem domain. Avison and Fitzgerald (1995) suggest that advocates of these models have come to recognise that user involvement is important but have left design decisions primarily under the control of technical experts. Since the 1980s, the empowerment of users during the development lifecycle has been taken more seriously, although the complexity of power and authority relations has yet to be fully appreciated. Mumford (1983) proposed the following typology, which summarizes much of the work on user-centred design since that time:

1. **Consultative participation** - Analysts discuss the system's requirements with users, but the technical experts perform design.

2. **Representative participation** - Representative users work on the design team and are, thereby, involved in decision making.

3. **Consensus participation** - Users drive the design process, making all the key decisions.

An obvious trend in this list is the transfer of authority and responsibility over design decisions from technical experts to users. It has rarely been noted, however, that the transfer of authority and responsibility to users does not empower them unless they are also given the *capacity* to act accordingly. Techniques for improving user capacities for contributing to software development have also received little attention in the literature. Consequently, software development has continued to be driven by technical experts, albeit with increased awareness of the need for user involvement.

Looking at information systems development more widely, a notable trend is the increase in end-user computing, particularly through the exploitation of desktop application packages. Such developments of an organization's IS capacity are typically small in scale and are usually developed in a relatively unplanned manner. The response of the Software Engineering community has been to regard such activity as distinct from more formal and large-scale software development. The IS community has shown that this kind of user IS development can be both a source of great benefit and harm to an organization's capability (enabling users to improve their own working practices, for example, but increasing the fragmentation and inconsistency of the organization's information), but has not presented an explicit lifecycle model that represents formal software development and user computing as interrelated parts of IS development. To address this omission, chapter 2 sets out a case for changes to information systems theory and practice through the adoption of user-centred information systems and chapter 6 develops such models using the socio-cognitive theory of information systems. Particular attention is paid in chapter 6 to demonstrating the effects of information systems and software development in terms of social transformation.

1.2.2 Current ICT Investment Evaluation Theory and Practice

Approaches to measuring the value of ICT investments can readily be divided into three categories: IS investment evaluation; usability evaluation; and software engineering metrics. Owing to the focus on the social impacts of systems and software development, noted above, information systems investment evaluation, which aims to establish the expected or actual contribution of ICT-based changes to the organization, is primarily considered here. Before discussing investment evaluation in detail, it is important to note that the three categories are in no way independent. The use of software engineering metrics to maintain the technical quality of systems has two key impacts on IS development. First, good technical quality is vital to user acceptance, as illustrated by notable failures, such as the London Ambulance Service (Beynon-Davies 1995). Second, the technical experts responsible for operating and maintaining software systems are key stakeholders whose preferences and interests should not be ignored. Similarly, usability evaluation derives its relevance from its attempts to accommodate user characteristics in the development process. Given that user acceptance has been shown to relate to productivity and other aspects of organizational performance, usability and software quality issues must clearly be accounted for. As illustrated in the following discussion of current perspectives on investment evaluation, progress has been limited.

Early approaches to ICT investment evaluation, through their derivation from accounting methodology, relied extensively upon monetary valuation and typically aimed to quantify the direct costs and benefits of an ICT investment. It was soon realised that these approaches provided a very limited analysis, even where efficiency and direct productivity gains were the main objectives (Couger 1987, Strassmann 1990). Consequently, a number of elaborations of standard accounting techniques were developed. The resulting methodologies, such as Information Economics (Parker, Benson and Trainor 1988) and Return on Management (Strassmann 1990), met with limited success, failing to supplant the less advanced accounting techniques. This was primarily because, in attempting to quantify intangible benefits, these methods diminished a number of familiar and important attributes of standard accounting practice (Hemingway 1997a), most notably by making analyses more abstract rather than being solidly based in fact. Nevertheless, interest in the so-called soft benefits continues to influence the direction of IS evaluation research (Farbey, Land and Targett 1995). This focus upon describing the 'true value' of an information system seems increasingly misguided, however, given that money derives its meaning from exchange within the market and, therefore, is of limited usefulness in describing the value of transformations that occur within a single firm and, moreover, need not relate directly to the firm's engagement in the market. Furthermore, there is a growing body of research aimed at providing an understanding of the evaluation process (Hirschheim and Smithson 1987, Symons 1990, Walsham 1993), which demonstrates the importance of considering the values of individuals and social groups affected by information systems changes.

The focus on the evaluation process, which has been driven by advocates of hermeneutic research, has provided some of the more interesting IS research findings of the past decade. This research has often focused on the political nature of the evaluation process (Symons 1990, Avgerou 1995, Bryson and Currie 1995), although much research has explored more far-reaching social and political issues relating to the introduction of technologies into organizations (Markus 1983, Zuboff 1988, Orlikowski 1992, Walsham 1993, Orlikowski and Gash 1994). This research has not yet provided an explicit framework or methodology for facilitating investment evaluation, although a general dialectic approach to building consensus regarding a system's value is typically advocated (see, for example, Avgerou 1995).

In light of the criticisms levelled against the accountant's focus on capital, multi-criteria methods (French 1986) offer an interesting alternative which: (a) do not demand the use of monetary values; (b) provide mechanisms for quantifying subjective values where criteria cannot be directly measured; and (c) deal explicitly with risk and attitudes to risk. The absence of monetary values is a mixed blessing: the analysis is able to represent a range of non-monetary values, but loses the familiarity of financial benchmarks that makes accounting techniques convenient aids to decision making. The quantification of subjective values, however, appears promising in terms of addressing the need to pay more attention to the informal evaluations of individuals. Several possible approaches are available for doing this, ranging from the negotiation of a single set of attributes and values for a group to the formulation of analyses for every individual involved (Ngwenyama and Bryson 1999). Whilst some reports suggest that users perceive some multi-criteria techniques as useful, the empirical evidence for and against these techniques is both sparse and inconclusive. Questions regarding the adequacy of multi-criteria methods concern the elicitation of subjective risk estimates and values (Angell and Smithson 1990). The reliance upon the individual to identify the relevant evaluation criteria, quantify them on arbitrary scales and estimate the various risks, suggests that considerable support ought to be provided at the elicitation stage. It is this stage, however, that receives least support, with methods tending to focus on different ways of manipulating the quantities once they have been derived. This deficiency seriously limits the effective application of multi-criteria methods.

As noted above, the lower level analysis of evaluation theory, as compared with social theory and economics, suggests the relevance of psychological and social-psychological issues. These have been explored in a limited way (see, for example, Clegg et al 1996) but seem to lose sight of the resource management issues that are basic to the investment decisions that IS evaluation aims to address. There has also been some documentation of the use of business analysis techniques (such as scenario analysis) for evaluating organizational change related to information systems (Clemons 1995). Whilst these techniques seem useful and are often highly acclaimed, their advocates offer little in terms of rigorous evidence of their social and psychological impacts. In other words, the techniques lack any formal theoretical underpinnings, typically being developed by practitioners or consultants and the claimed benefits are usually supported by only anecdotal evidence (Kuhn and Sniezek 1996). The limited academic research that has been conducted into the application of these tools has generally been inconclusive, although some reports suggest that the opposite psychological effects to those claimed have been observed (see Kuhn and Sniezek 1996, for example), thereby implying that business analysis techniques may potentially impede decision makers.

A number of strategic approaches to evaluation have been proposed in the IS literature. These methods provide higher level analyses than the above quantitative techniques, aiming to identify the general impacts of the investment on the organization's goals. The portfolio approach to evaluating information systems investments (Ward 1994), for example, supports strategic decision makers in prioritising the various IT related projects that have been proposed. Such a technique is clearly useful for deciding, say, whether to invest in computer-aided design facilities or a new administrative network. Strategic approaches make a much more limited contribution, however, to the choice between alternative implementations to address the same business problem - the focus of financial appraisal and some multi-criteria methods. Such strategic methods complement rather than substitute financial appraisal by supporting the selection of problems to be solved, rather than choices between alternative solutions. In doing so, they make a significant contribution to IT investment appraisal, but they do not address the majority of problems at the financial appraisal level and, moreover, typically do not serve to integrate these lower level analyses.

At the lower level stages of investment evaluation, analysis has been limited by the neglect of technical quality issues. Stakeholder preferences are taken into account by some methods and stakeholders frequently cite technical quality and usability issues as significant factors in their appraisal of a system's acceptability. As already noted, despite these interrelations, no attempt has been made to integrate investment evaluation methods with the findings of software quality or usability evaluation, which are largely within the domain of the software engineer, rather than the information systems or business analyst.

1.3 A Case for Information Systems Theories

Despite the proliferation of studies of information systems at all levels of analysis, there have been numerous comments by IS researchers about the discipline's lack of progress in addressing its substantive issues (Banville and Landry 1992). Even when the discipline's relative youth is taken into account, there seems to be some justification for these criticisms because information systems research does not seem to be cumulative. In considering why IS research does not seem to progress in a cumulative manner, two issues are considered here: (a) whether IS *should* be cumulative; and (b) whether IS *can* be cumulative. Whilst various arguments might be offered for and against the first point, the case presented here focuses on the relationship between theory and action, particularly the actions of practitioners and researchers in IS.

As an academic discipline, IS must attempt to demonstrate an improved understanding of its phenomena of interest. Like many disciplines in both the humanities and the sciences, IS can best demonstrate its understanding by collecting evidence and providing accounts and explanations consistent with that evidence. It must be determined, however, whether - like history - IS should focus on explaining particular events or - like physics - aim to provide generalised explanations in the form of theories.

Information Systems is ultimately concerned with social and psychological phenomena. Whilst psychologists have mostly adopted a scientific approach to studying human behaviour, sociology has no methodological orthodoxy (Giddens 1995) and the ontological status of basic social phenomena are still the subject of considerable debate. A key difficulty in providing general explanations of classes of social phenomena is that there are no clear criteria for evaluating such explanations. Although IS relies heavily upon the study of social organization, the issue of evaluating generalised accounts is less problematic because IS has a substantial engineering component. Regardless of the efficacy of generalising across social phenomena, the IS community has largely rejected the alternative of regarding information systems development as a technology driven process that can only be appraised in a summative manner. Consequently, the aim of developing a general understanding of IS phenomena in order to improve information systems practice has fairly widespread acceptance. This motivation is significant because it sets a criterion for evaluating the community's attempts to explain its phenomena of interest: the explanations must be sufficient to serve as the basis for purposeful and effective intervention. This criterion provides a standard by which IS theory can be said to be cumulative in its explanatory power, whilst avoiding a commitment to a strict natural scientific approach to studying IS phenomena.

Although the above criterion may not be regarded as ideal, particularly by those who insist that all scientific activity should be a pursuit of objective truth, it is suggested here that it is a much stronger criterion than it at first appears. In fact, the requirement is very strong for a discipline in such an early stage of development. In many respects, the criterion parallels that proposed by Hacking (1983) for justifying ontological commitments made in physics, such as the existence of electrons and other subatomic particles. Whilst the evidence collected by the IS community cannot be as precise as that gathered by physical scientists, close attention to consistency in research method will enable the IS community to clarify the ontological basis of its theories and explanations.

Developing an ontology for IS theories is, at present, a difficult process because research results are highly fragmented. Many different research methodologies have been used and studies are often exploratory, suggesting new phenomena rather than testing the validity of existing claims. Furthermore, the many ontological commitments to postulated phenomena are rarely made explicit and, owing to the lack of an established research technique, cannot readily be inferred from published IS research. Nevertheless, the development of an IS theory with an explicit ontological basis is a necessary starting point for improving the precision and accuracy of current explanations of IS phenomena. For this reason, the main activity underpinning this research has been the development of an ontology through the analysis of the current literature in IS and its reference disciplines. Such an analysis is inevitably limited in what it can achieve, but represents a step towards providing testable information systems theories.

Given that IS deals with social phenomena and a theory-driven approach to its study is advanced by this thesis, a potential criticism must be addressed. The notion of theory driven research in the social sciences raises the prospect of the self-fulfilling prophecy: the possibility that observed phenomena may be artefacts of the research process. Such effects are not confined to the theory-based study of social phenomena, but also affect observations made in a purely exploratory manner, because observation is affected by the preconceptions and prior knowledge of the observer (see chapter 3). These impacts certainly complicate the research design process and make the analysis of results less certain, but this is an unavoidable complexity of studying human action. Besides, there is a further problem with exploratory research: the researcher is free to interpret his or her observations with few constraints. Theory-based empirical research distinguishes itself from other types of activity by its emphasis on intersubjective agreement underpinned by well-defined research methods. Sound theory bolsters intersubjective agreement by providing a common language for discussing the phenomena under investigation. Such shared meanings are less apparent in IS research than in more established research areas; something that cannot be addressed without initial attempts at developing substantive theories to underpin research design. It is this close relationship between theory and research method and its implications for research quality and progress that motivate theory development in this thesis.

The primary difference between an historical account and a theory is that a theory aims to explain a broad class of past and future events that share significant commonalities, whereas a history explains a specific class of previous instances. The generality of a theory draws out regularities from the past that aid prediction in similar future circumstances. The scope of a theoretical account is not absolute or objectively defined, but depends upon the judgement of the theorist as to which similarities and differences will define the class of phenomena to be explained. In the natural sciences, the subjects of theories can be precisely defined and are often subject to control. Consequently, theories can be formulated with precision. In contrast, social phenomena are very difficult to define and cannot be subject to control. In both cases, however, theories have proven to be of value both for furthering understanding and for providing practical guidance. Although not advocating instrumentalism, the practical implications of theory are of particular interest to the development of this thesis.

The practical value of theory stems from its generality. Assuming that the future will resemble the past, generalization facilitates prediction and, consequently, the future application of knowledge and skills. Each new experience corroborates or challenges theories, reaffirming their value, or leading to their modification or rejection. Considering generalization in its broadest sense, all technologies can be regarded as embodiments of theory, even those that have evolved through use, rather than being designed around scientific principles. Effective technologies, better than anything else, demonstrate the validity as well as the value of theory. As indicated in chapter two, computer-based technologies are not of as good a quality as might be hoped. Some of the reasons suggested for this lack of quality are:

1. Software engineering has focused on developing complete systems without the use of highly standardized components.
2. Software design has paid insufficient attention to the physical aspects of technologies.
3. ICT-based systems seem difficult to specify and adapt according to user needs.
4. The interaction between human and technological artefact is difficult to apprehend.
5. The effective use of ICTs is contingent upon the users' skills for using information to guide their actions as well as their ICT skills.
6. The social impacts of ICTs have proven hard to identify and open to broad interpretation.

The first point is not that software development problems are treated holistically and not sufficiently decomposed to make them manageable. The main obstacle to quality is that each component is constructed anew and time pressures lead to quality being compromised. Standardised, configurable components would provide a way by which systems could be constructed quickly and at low-cost, yet achieve high levels of reliability. Such an approach is being pursued by advocates of software reuse, component-based software and modular systems, such as enterprise-wide systems. However, many of the conceptual problems in determining how to identify and design solutions of general application have yet to be resolved. Similarly, some progress has been made in addressing the actual limitations of technology artefacts during design (point 2). A recent example is the unified software development methodology based upon UML (Jacobson, Booch and Rumbaugh 1999). The modelling of technical constraints is, however, primitive and informal.

Despite the extensive efforts of HCI and information systems researchers, establishing user needs, mapping these onto technical specifications and providing sufficient flexibility to address changes in user needs have proved difficult areas in which to make progress (point 3). Increasingly, SE, HCI and IS have looked towards ontology and epistemology for solutions to the problems of addressing diverse user needs and, more importantly, different user perceptions of the problem domain (Checkland 1981, Avison and Wood-Harper 1990, Walsham 1993). In so doing, the research community has increasingly come to realise that the notion of a single description of the world is an inadequate basis for designing an information system for many users working in complex social organizations. Users' models of the real world differ, often with good reason, and such differences cannot be decisively resolved.

The fourth point - understanding the interactions between human and technological artefact - has become increasingly important for three reasons: (a) ICTs have become available to non-expert users whose usage is (or should be) incidental to their activities; (b) ICT systems have become increasingly complex and used for a wider range of tasks; and (c) ICTs are increasingly used to mediate interpersonal communication. Reflecting the influences of cognitive science and ergonomics, HCI has two theoretical cores - models of perception and low-level decision making, and models of physical interactions - that are reasonably well integrated (for example, GOMS, Card et al. 1983). The relationship of this body of theory to theories of high-level cognition and action is much more limited. The main consequence of this is that HCI pays comparatively little attention to the interpretation of information and how this affects user action. In contrast, software engineering focuses on data processing issues and information systems focuses on organizational concerns, whilst neither consider the detailed cognitive issues of information design. It is, perhaps, for this reason that the combined effects of interface layout, data model, and organizational/social factors on users' interpretations of information have not been fully appreciated by any of the disciplines. This is a fundamental obstacle to designing quality systems because no body of theory can provide guidance on how changes to the information made available to users will affect the ways that they perform their tasks. As identified by point 5, the consideration of all three factors is also essential for establishing effective combinations of artefact designs and user training.

Finally, the longer term implications of ICTs at the organizational and societal levels cannot be appreciated or managed without some understanding of the nature of personal and inter-personal impacts of ICTs. Achieving such understanding has proved difficult for the IS community. Alternative philosophies and research approaches have shown some promise, but raise the problem of how to relate the findings to those of software engineering and HCI. Without such a relationship, the contribution of the theory to practice is limited. Indeed, the guidance offered so far is not obviously compatible with SE and HCI best practice.

The above discussion illustrates the need for sound integrated theory to serve as the basis for prescriptive methodologies for developing ICT-based information systems. It also shows that the pursuit of sound theory has led to the increased interest of all three disciplines in philosophical issues. These issues are clarified and given a new treatment in chapter 5.

1.4 Research Approach

The original aim of this research was to develop a more satisfactory methodology for evaluating IT investments. Research in this area had long been concerned with conceptual problems, such as intangible benefits, and issues of subjective interpretation, such as the divergence of stakeholder values and viewpoints. An analysis of these issues (Hemingway 1997a) suggested that these problems stemmed from deficiencies in the premises of the various methods. Although interpretive research had shed some light on these issues, progress seemed to be fundamentally limited by the ways in which the relationship between the individual and the organization was conceived. In the case of cost-benefit analysis, for example, attempts were made to accommodate individual differences in viewpoint to provide a 'complete' analysis of social/organizational impacts. Such attempts violate the assumption that accounting analyses should address specific audiences and the economic premise that organizations work towards organizational goals; both of which are relied upon by the techniques that cost-benefit and other methods extend.

Further investigation of the limitations of existing evaluation methods led to the analysis of the economic theory of the firm upon which most evaluation methods were, in effect, based. A major deficiency appeared to be that a simplistic model of intrafirm activity led to an overemphasis on efficiency gains derived by substituting components of a system (i.e. an organization or a business process) with functionally equivalent but technically superior alternatives. Whilst adequate for evaluating many types of automation, this simple model is implausible in the context of modern computer-based information systems. The reasons for this lie not in the technologies themselves but in the role of information in changing the ways in which an organization's contituent parts interrelate. Specifically, Hemingway (1997b) argues, ICTs encode rules, some of which would previously have been applied as part of routine working practices and norms. Such changes, particularly on a large scale, alter the interactions between workers and, thereby, affect the 'how' and 'why' of organizational activity as well as what is being done. Conventional economic theory is, therefore, ill-suited as a basis for evaluating such changes and a replacement theory is required.

Having establised the need for an alternative theoretical basis for IT investment evaluation, the next task was to identify a suitable alternative. Such a theory would need to provide a detailed treatment of the relationship between the individual and the organization so that problems like those raised above could be resolved. In an attempt to identify an appropriate theory, the literatures of IS and several reference disciplines were examined. During this search, particular attention was paid to examining the implications of the theory at the various levels of analysis. When looking at a theory of the individual, for example, its implications for the nature of social organization would be considered in detail. When looking at a social theory, however, particular attention was paid to the plausibility of the assumptions made about individuals' actions.

The IS literature offered few attempts at integrated theory development. A growing recognition of the limitations of conventional scientific methodology had led to a more pluralistic research culture. Whilst the resulting diversity of research had greatly enriched the evidence available to IS researchers, it had reduced consensus on the meaning of basic terms, such as information, system and organization. This situation continues to prevail and has not proved especially conducive to theory development (Baskerville and Lee 1999). There was, and continues to be, a growing awareness of the need to address epistemological issues when conducting and evaluating research. Less attention has been paid, however, to clarifying basic terms and their referents in the real world. A notable exception to this trend is a paper by Mingers (1996), which argues in favour of utilizing a model of the individual that attends to the physical aspects of human experience and action as well as the mental aspects addressed by most cognitive theories. At the time this research was conducted, Mingers had provided only the beginnings of an embodied theory of cognition, with the analysis of social organization limited to a literature review. In contrast with Mingers' theory, few information systems theories have developed a model of human cognition. Consequently, understanding of the relationship between the individual and the organization has been limited by the lack of explicit theorizing about cognition and action.

In contrast with information systems, HCI has retained a scientific research methodology and focused upon the individual user. A review of HCI theory suggested that, although the conventional approach had facilitated theory development - drawing quite heavily upon the cognitive architectures developed by the cognitive science community - the rationalistic approach resulted in theories focused on goal-directed behaviour rather than explaining the nature of human action as a form of social conduct. The evaluation of ICTs provides a good illustration of the limitations of HCI theory. It is generally assumed that users work towards well-defined, short-term goals. Little consideration is given to the relationships between an individual's various activities or to the influence of social context and social roles on the development and attainment of goals. Clearly, such an assumption is problematic if one is to balance an HCI-based usability analysis against an economic evaluation that assumes shared social goals. Thus, whilst HCI theory is valuable within the confines of its narrow assumptions, it would need to be considerably extended in scope to provide a useful basis for understanding the relationship between the individual and the organization.

Like information systems, sociology has no methodological orthodoxy and is open to a broad range of research methodologies and philosophical perspectives. The focus of theory development is upon social organizations and societies, although there is a long tradition of considering the relationship between the individual and society. In comparison with IS, however, sociology places less emphasis on research that supports the planning of organizational or social change and research tends to be more descriptive. Nevertheless, social theories have been used to inform previous IS research and represented a substantial body of relevant theory. Having being considered by other IS researchers (Orlikowski 1992, Walsham 1993), the theory of structuration (Giddens 1984) appeared to satisfy many of the requirements identified by the analysis of the IS evaluation literature. For this reason, structuration theory was studied in some detail. In studying structuration theory, particular attention was paid to Giddens's use of ego psychology to underpin assumptions about the individual actor, as compared with the cognitive psychology used in HCI. A particular concern was the extent to which ego psychology could be used as a basis for design and intervention in IS, rather than as a descriptive tool in sociology. Although the quality and depth of explanation offered by structuration theory was found to be valuable in understanding some aspects of information systems and their social contexts, it has serious limitations as a basis for HCI-type study. The additional emphasis given to the situated nature of cognition and the analysis of routinization, for example, was not supported by any justification in terms of physiology and basic cognitive processes, such as learning. Yet, without explanations at these lower cognitive levels being consonant with the higher level analyses offered by structuration theory, its use to support IS-related change would not ensure that design rationales based around the individual user would be consistent with the organizational/social objectives that justified and facilitated the change process.

Having failed to identify an appropriate theoretical basis for developing an IS evaluation methodology to solve the problems identified, two alternatives were evident: (a) develop a methodology using existing theory, with the intention of making its limitations explicit; and (b) developing an theory that integrates social and psychological explanations of IS phenomena. Given that the limitations of existing theory seemed fundamental, relating to the basic concepts of information, system and organization, the first alternative was considered neither to be feasible nor to offer significant improvements over existing evaluation methodologies. For this reason, the development of an integrated theory was attempted.

In developing an integrated theory, care was taken to draw upon existing theory in the various reference disciplines, whilst ensuring consistency through the identification of the assumptions of existing theories and their reinterpretation with respect to a single conceptual basis. Given that such a development process would be interpretive, a particular concern was ensuring that the study was sensitive to the historical context of theory development. Although an ideal approach would have been to study the histories of each academic discipline and of the relationship between the computer, the individual and society, this was clearly beyond the scope of the research and a more modest analysis was devised. Firstly, in order to gain an appreciation of the role of ICTs in society and the precursors to the IS discipline, telecommunications and software engineering were briefly studied. This study provided some insight into what are, currently, the most widely used ICTs and the most widely used methods for developing computer-based systems. Second, to provide a review of how the relationship between the individual and ICTs had been conceived, the discipline of HCI was examined. Finally, the IS discipline itself was reviewed to reflect the contribution of social and organizational analyses to the understanding of ICT-based systems.

Given the longer history of telecommunications and its increasing convergence with information technology, its study provided several insights into the theoretical and practical difficulties facing information systems. Having made the transition from being a subject of interest to the electrical engineer to a household (and, increasingly, a personal) commodity, the telephone serves as a useful reference point for making sense of developments in computer technologies and their application. Of particular value was the comparison of the notions of 'information,' 'communication' and 'system' in telecommunications and information systems, which led to the realization that the theory must be able to provide precise definitions of these terms that had clear referents. In the absence of such precision, explanations of IS phenomena would be ambiguous and subject to broad interpretation. The increased awareness of these basic conceptual issues was reflected to a degree in a first attempt at theory development (Hemingway 1999). Further analysis, some of which is presented in chapter 2, confirmed the initial impression that more attention needed to be paid to making explicit which phenomena are to be regarded as concrete objects, events involving concrete objects, mental phenomena or other kinds of entity. Consequently, a search of the philosophy of science and philosophy of social science literatures was undertaken.

Over the past twenty years, the philosophical debates in information systems have focused on epistemological issues, such as how to account for the differences in actors' world views and the subjectivity of the research process. Although these debates have benefited the IS community, the relativism of some IS research has led to a degree of self-doubt about the validity of generalized claims made by IS researchers (Baskerville and Lee 1999). The questions being raised, combined with the concerns about the ambiguity of basic terms, suggested that a key task in theory development was the construction of an ontology of information systems. A review of several discussions of ontological issues found two perspectives - one natural sciences (Hacking 1983) and one social sciences (Keat and Urry 1982) oriented - that provided scope for addressing ontological issues whilst permitting the diversity of epistemological viewpoints suggested by IS research. In particular, both considered the justification of ontological commitments to kinds of phenomena, whilst recognizing that each kind could be described in different ways and that no description could be regarded as objectively true.

Adopting a philosophical standpoint derived from the work of Keat and Urry (1982) and Hacking (1983), the literature of the various reference disciplines was again reviewed with the purpose of developing ontological commitments to basic phenomena and, building upon these definitions, providing explanations that were as precise as established evidence would permit. As discussed in chapter 4, the analysis of the sociology literature led to the conclusion that the only safe ontological commitments that could be made were to the individual actor as a physical being; the spatiotemporal regions in which social activities took place; and events involving actors and physical objects. This analysis concurred with that of Giddens (1984) in two key respects: (a) it implied that the apparent structuring of social activity was inferred from perceptions of the similarities between events; and (b) similarities in social activity were perceived because individual actors tend to act similarly in similar circumstances and alter their behaviours in accordance with those of other actors. There was, however, a key point of departure from the ontology of the theory of structuration: the reliance upon ego psychology did not appear to be justifiable as an account of empirical evidence from cognitive psychology. Thus, ontological commitments to the individual must be developed by drawing upon the psychology and HCI literature and the ontology of social phenomena then reconsidered.

Despite concentrating on established results in cognitive psychology, it was clear from a brief review of the discipline that controversies surrounding basic cognitive phenomena could not be decisively resolved with the empirical evidence available. Consequently, given the intimate relationship between information and subjective meaning, it would prove impossible to resolve many ontological issues raised by the comparison of social and psychological analyses. As a result, the research strategy shifted from a constructive process of ontology development to a process of synthesis broadly comparable with dialectic. The application of this process (described in section 1.5) began with the most basic communicative act and developed an explanation of how a combination of physiological similarity and interaction between individuals in co-presence could result in the emergence of shared meanings. Although clear ontological commitments could not be justified in this part of the theory's ontology, the commitments made were considered from several levels of analysis, ranging from the physiological level to social cognition.

Having 'reconstructed' the notion of the individual from a socio-cognitive standpoint, theory development continued by reinterpreting the earlier ontology of social phenomena to ensure their consistency with the remainder of the theory. Encouragingly, this stage of theory development uncovered few inconsistencies, which were resolved through a similar 'dialectic' process to the one applied to the psychology literature. At this stage, however, it was recognized that the focus on cognition had implicitly maintained the assumption of well-defined goals, which, in the light of the literature review, did not seem defensible. Drawing upon the social cognition, sociology, economics and IS evaluation literature, this assumption was investigated and the theory modified to provide more detailed explanations of personal and social value concepts. It is at this stage that the theory is weakest. Again, the reason for this stems from the difficulties encountered by all disciplines in investigating the concepts related to personal value, such as the self-concept, either empirically or conceptually. The explicit inclusion of a more detailed account of value was, however, felt to be valuable, although it is expected that this part of the theory will change considerably as a result of future research.

Having developed a basic socio-cognitive theory, which addressed key psychological and social phenomena from a single consistent viewpoint, the theory was extended to account for IS-specific phenomena, principally the information system, information artefact and ICT. Given the ambiguity found in the information systems literature regarding the meanings of basic terms, some new terminology was introduced at this stage to make clear the distinctions between subjective phenomena, such as information, and objective phenomena, such as the artefacts used by individuals to inform one another. A consequence of this approach is that some of the basic terms used in the theory are quite different from those already familiar to the IS community. For this reason, it was felt necessary to 'test' these concepts and illustrate their value by applying them to information systems problems.

Given the potential contribution of the theory's detailed consideration of value and its novel IS-specific terms, IS evaluation was considered to be an appropriate application area. IS evaluation had the further benefit that it had already been studied in some detail and, consequently, the problems that needed to be addressed were well understood. As evaluation activity is a key component of the IS development lifecycle and existing lifecycle models did not utilize the concepts provided by the socio-cognitive theory of information systems, a prerequisite to exploring IS evaluation was the statement of the systems lifecycle in socio-cognitive terms. Taking the notion of social transformation as a focal point, such a lifecycle model was developed and is presented in Hemingway and Gough (1999a) and chapter 6. The evaluation methodology based upon this lifecycle model is presented in Gough and Hemingway (1999) and in the second part of chapter 6. The methodology draws upon the value concepts and other components of the socio-cognitive theory to provide an approach to evaluation that is quite different to those currently in widespread use. Although empirical testing of the methodology was beyond the scope of this research, some reasons why this represents a potential improvement over current evaluation methodologies are considered towards the end of chapter 6.

1.5 Dialectic as a Research Method

As indicated by the research approach, this research has relied upon conceptual analysis and the criticism of existing theory in order to develop an integrated theory of information systems. In order to claim that such a research process results in a valid contribution to knowledge, a systematic approach to analysis must be taken. As in the case of empirical research, this is best ensured by the use of an appropriate research method. Methods for analysing and criticizing theory are not widely addressed in the IS literature, however. For this reason, methods used in other academic disciplines were drawn upon throughout the research to develop an appropriate method. A significant influence in this regard was the process of dialectic, which has similar aims of criticizing and integrating/synthesising theory.

There is a long tradition of using dialectic, going back at least to Aristotle's analysis of the philosophy of Plato (Owen 1986). More recently, dialectic has been used by Hegel, who developed the method with the aim of making philosophy "an objective, demonstrated science." (Hegel 1976, p.53). Hegel's work has since had a significant influence on the use of dialectic in philosophy (Beiser 1993). Throughout this time, dialectic has been regarded as a method for developing more general and/or more valid knowledge claims, making it a potentially valuable tool for theory development.

Popper (1940) characterized Hegel's dialectic as a process of thesis, antithesis and synthesis. This reflect's Hegel's aim of showing that knowledge claims were self-contradictory and, subsequently, resolving these contradictions to attain a higher level of truth. Forster (1993) provides a more specific description of Hegel's method:

1. Show that category A contains category B.
2. Show that category B also contains category A and that the categories are, therefore, self-contradictory.
3. Derive a category C that subsumes A and B with their senses changed so that the

contradiction is resolved.

This account of Hegel's dialectic is sufficiently specific to be used as a method for developing knowledge claims. It reveals, however, some aspects of Hegel's philosophy that are incompatible with the philosophical basis of this research. Two points are of particular concern. First, the method focuses exclusively upon categories, or concepts, reflecting Hegel's idealist philosophy. Second, the method assumes that dialectic will lead to a single, objectively true body of knowledge.

Hegel's writings advocate both realism and the development of a single true body of knowledge (Sayers 1985), indicating that Hegel believed there exists an objective world and that true explanations of that reality can be developed. As an idealist, Hegel believed that such truth would be attained through a process of reasoning that would eliminate errors and inconsistencies in knowledge. This view that truth is ultimately reflected in ideas creates a serious problem for Hegel's philosophy: it is not possible to demonstrate that knowledge is true of the world if knowledge and reality are assumed to be distinct. This problem does not arise, however, if the relationship between knowledge and the world is considered in realist and materialist terms. Such a philosophical standpoint has been taken by Marx (1974) and, more recently, by Sayers (1985) as the basis for reinterpreting the dialectic method. It is also quite consistent with Hacking's position with respect to knowledge in the natural sciences (Hacking 1983) and the position adopted by Keat and Urry (1982) with respect to knowledge in the social sciences.

The second major weakness of Hegelian dialectic is the assumption that a single body of true knowledge can be developed. It is now well-established that experience of the real world is interpreted and that subjective interpretation is necessarily partial and often inaccurate (Kuhn 1962). Moreover, theories must be represented in some form and it is also evident that forms of representation both conceal and reveal different aspects of the knowledge they represent (Potts 1996). Replacing objective truth with a relativistic account of knowledge, however, threatens to undermine the dialectic process altogether. If

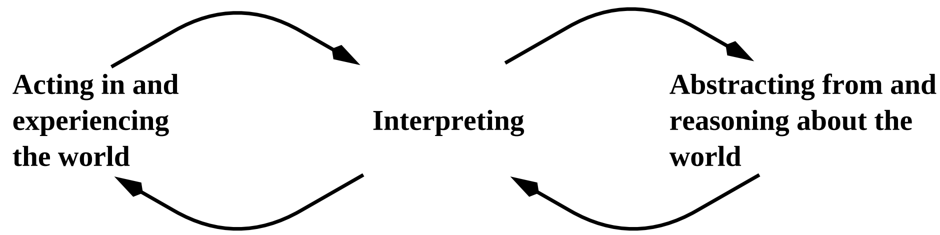


Figure 1.1: The Dialectical Relationship Between Acting In and Knowing About the World

no theory can be regarded as more true than another, then there is no basis for asserting that dialectic leads to better theories about the world. It should be noted, however, that, whilst the interpretive nature of subjective experience denies the possibility of objectively true knowledge, it does not rule out the possibility of evaluating the relative truth of knowledge claims, provided that experience and theory interrelate. Given that a close relationship between theory and experience/practice is consistent with materialism, this is an appropriate starting point for developing an alternative dialectic process.

Taking materialism, realism and subjective interpretation as starting points, a dialectic relationship between acting in and knowing about the world can be described (figure 1.1). The development of knowledge about the world is regarded as a process of: (a) actively engaging in the world, both to change it and to experience it; (b) interpreting one's experiences; (c) abstracting from those experiences and one's existing knowledge to develop knowledge about the world; and (d) interpreting that knowledge with respect to one's experiences of the world in order to guide one's actions. Clearly, a person is practically limited in what he or she can achieve in a given time and situation. Consequently, different kinds of activity emphasise different aspects of the dialectic process. Some kinds of activity are illustrated in figure 1.2.

Developing knowledge and evaluating its truth, like any other human activity, necessarily involves engagement in the dialectical process. The difference between everyday activity and the scientific pursuit of knowledge are, therefore, only matters of degree. They are social conventions and institutionalized practices that serve a variety of functions, the

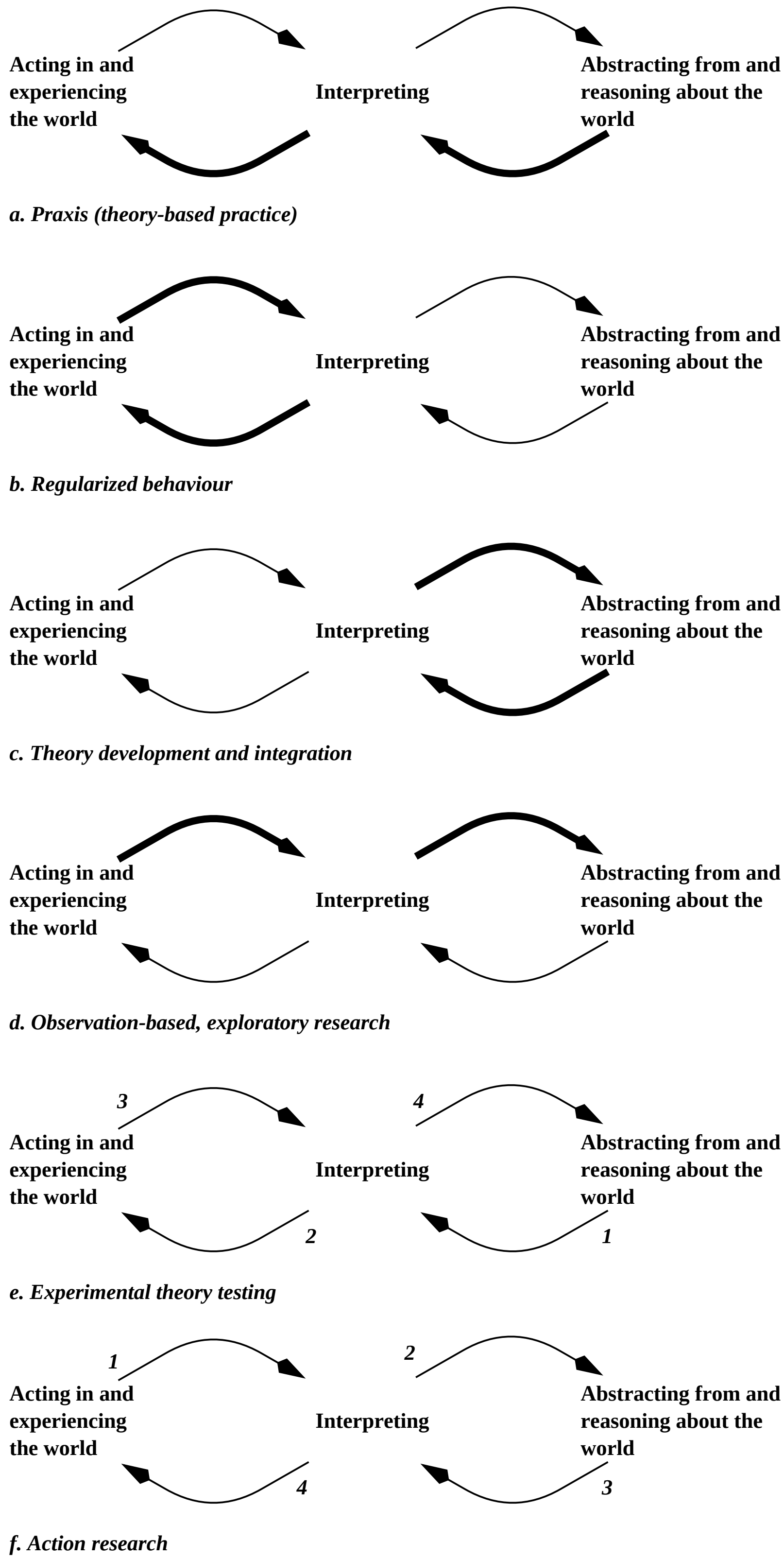


Figure 1.2: Illustrations of How Different Kinds of Activity Place Different Emphases on the Dialectic Process

most important of which are:

- specifying what constitutes valid observation and the reporting of experiences, and what constitutes acceptable scientific activity;
- determining whether the referents of a theory are real;
- appraising theories in terms of their scope, coherence, consistency and simplicity; and
- developing forms of representation that effectively communicate evidence and theory.

In order to establish a method for analysing, criticizing and developing theory, it is necessary to identify the contribution of theory and theory development to each of the activities above. It should be noted, however, that relatively little will be said about the final point because forms of representation are subjects of the research and are, therefore, addressed in the following chapters.

In considering the acquisition of valid scientific evidence, it should be noted that dialectic asserts a causal relationship between the world and the subject. A person comes to know about the world by experiencing its effects and by interpreting and reasoning from these experiences towards increasingly accurate descriptions of the things themselves. A critical component of empirical and theoretical research, therefore, is an appreciation of cause-effect relationships and causal attribution.

Ideally, causal relations could be described in terms of necessary and sufficient conditions:

- a cause is necessary if the effect is only present when the cause is present; and
- a cause is sufficient if the effect is always present when the cause is present.

A simple causal relation of this kind would, in principle, be falsifiable in the Popperian sense (Popper 1972). Unfortunately, the possibility of causally inhibiting factors complicates this situation because it cannot be the case that there exists a causal factor that satisfies the above criteria and a causally inhibiting factor that also satisfies them. There are, however, many kinds of events to which both kinds of causal factor are associated - as in the case of our ability both to start and to extinguish fires. Thus, a more complex account of causality is required, involving non-necessary and non-sufficient causes. It is also sometimes desirable to describe causes in quantitative or other relative or proportionate terms. Following Cartwright (1983), it is suggested that causal relations can be asserted on the basis of both:

- creating causally isolated conditions with respect to a particular effect and adding or removing causal factors in a controlled manner; and
- showing that the probability of the effect given the presence of the causal factor is greater (or smaller) than the probability of the effect.

The first requirement involves establishing minimally sufficient conditions to yield an effect. Where possible, variations in conditions and resultant variations in effects are recorded with precision. The resulting evidence is used to establish generalized claims about causal relations, using changes in probability of occurrence as evidence of causality.

The human subject complicates the analysis of causality in two ways: (a) objective causal relations only become evident through a dialectic process between the world and an individual; and (b) human subjects engage in a dialectical relationship with the world, which must be explained in causal terms. The use of dialectic to understand phenomena as objective phenomena has two implications. First, causal isolation is an experience of causal isolation and, therefore, the situation under study can only be regarded as causally isolated insofar as it appears to be isolated and insofar as knowledge of relevant causal factors has progressed. Second, where technologies support causal analysis the subject's interpretation of the form(s) of representation used will affect his or her attribution of causality.

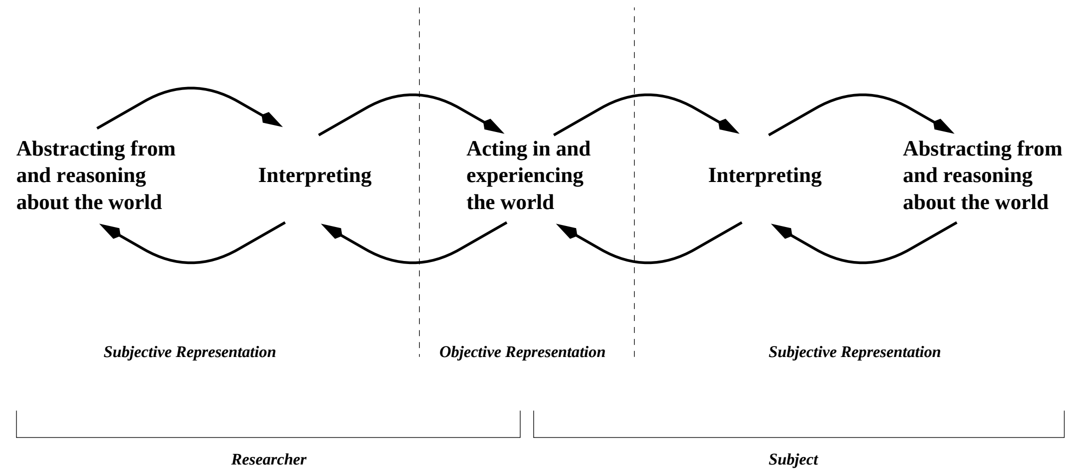


Figure 1.3: The Dialectical Relationship Between Researcher and Subject

Considering the causal analysis of human action, the situation is more complex. Causal attribution must attempt to predict the effects of the actions of others whilst recognizing that the person, as a causal agent, has acted freely (Hospers 1956). We must also recognize that any such causal attributions are made on the basis of the observer regarding the actor as an object. Such attributions do not represent the intentions of the subject that led to the effects observed. There is a fundamental difference between agency and action, and regarding the person as an object only provides insight into agency. To appreciate human action, a more complex dialectical relation must be understood (see figure 1.3).

By understanding the dialectic relationship between researcher and subject, it is possible to build upon an objective causal analysis by working out what a person believes he or she has done (as opposed to what evidence suggests were the effects of his or her behaviours) and why he or she acted in that way. In order to do this, it must be recognized that subjective processes involve personal knowledge and intentions. Moreover, actors consider the social interactions they observe to be the results of the intentional actions of others. Thus, social actions are based upon the interpretation of knowledge with respect to an individual's intentions and the intentions of other actors. The various interpretive research approaches can be effective in revealing a person's intentions and their perceptions of the intentions of others (Boland 1985, Walsham 1993). For the researcher to make sense of the views of others, he or she must also recognize that his or her interpretations also involve personal intentions relating to the conduct of the research.

When utilizing a text as evidence, rather than a particular social interaction, the dialectic process is somewhat different (as is the case with some forms of ICT, as emphasized by the distinction between interacting with and through ICTs presented in chapter 5). The most obvious difference is that the intentions of the author are not affected by his or her interpretations of readers' actions. Moreover, neither the text nor the intentions behind its construction will be affected by the intentions of the reader. These differences must be taken into account when devising methods for empirical study based upon the methods for analysing texts, as in the case of hermeneutics (Hirschheim 1992), and vice versa. A failure to do so increases the risk that the intentions of the subjects under study will be misinterpreted.

Having considered objective evidence relating to events and related this to evidence concerning the intentions and knowledge of actors, an attempt can be made to utilize the resulting evidence to develop one's own knowledge. A first step in this process is to identify the referents of the texts or actors under study. When doing so, the aim is to evaluate the evidence relating to the entities and events postulated. Where the evidence suggests that a particular kind is real, this knowledge can be checked against one's existing body of theory and one's own ontology, thereby, improved. Where the evidence suggests that a postulated kind is not real, the evidence must be used both to identify possible errors in one's own ontology and to reinterpret the text or empirical evidence. The purpose of reinterpretation is to establish whether: (a) the initial attempt had resulted in a misinterpretation; or (b) whether sense can be made of the subject's or text's explanations in terms of one's own theoretical viewpoint. At all times, the process is a dialectic one in which 'negotiations' between the evidence being studied and the theory being developed work towards a more satisfactory explanation.

It is inevitable when studying human activity that concepts will be used for which there is little evidence. In such cases, the dialectic process must often relax the materialistic requirement and make a trade-off between the value of the theory and its rigour. Other evaluation criteria, such as the utility, consistency and simplicity of the theory, although not as strong as the materialist requirement, can be used to support the dialectic process. In cases where sources of evidence are in conflict, however, those claims with the most concrete evidence must carry more weight in theory development.

1.6 Organization of the Thesis

Expanding upon the background and key issues outlined in this chapter, four perspectives on information and communication technologies (ICTs) are presented in chapter 2. The main aim of this review is to illustrate how current perspectives on information systems and software development and evaluation are singly too narrow in scope, yet too disparate to be applied in combination. Having considered the limitations of current information systems approaches, the contributions of reference disciplines are considered in chapters 3 and 4. Chapter 3 considers various psychological and sociological perspectives on the individual, paying particular attention to the different ontological commitments made by the theories, as well as their coverage of the key aspects of individual action and interaction that need to be addressed by the IS community. Chapter 4 presents a similar analysis at several social levels, again considering both psychological and sociological theories. Given the subject matter of information systems research, particular attention is paid in this chapter to the analysis of organizational forms.

Having analysed relevant theory and empirical studies at various levels of analysis, from the single person to supra-organizational forms, chapter 5 presents the socio-cognitive theory of information systems as an integrated theory for explaining the psychological and social issues relevant to the development of ICT-based information systems. Chapter 6 presents initial applications of the theory to propose new models of the information systems and software development lifecycles and to address some of the problems of ICT investment evaluation. Chapter 7 reflects upon the research process, comments on the potential contribution of the proposed theory to improving information systems theory and practice and considers options for future research to develop and test the theory and its applications.

Chapter 2

Current Perspectives on Information and Communication Systems

2.1 Introduction

This chapter reviews four disciplines that are primarily concerned with the design, development, evaluation and use of information and communication systems. The first - Telecommunications Engineering - is briefly considered in terms of theory and practice related to the transmission and receipt of signals. This discipline is distinguished from the remaining three by its technology-centred focus and lack of attention to the 'informational content' of the signal. This area is typified by Shannon and Weaver's theory of communication (Shannon and Weaver 1963).

The second discipline - Software Engineering - also has its origins in technology-centred design, although the emphasis is upon the efficient storage, maintenance and dissemination of structured data. The interest in structure is manifest in a shift in emphasis from physical to logical design considerations.

The third discipline - Human-Computer Interaction - represents a significant shift in theoretical underpinnings, drawing upon cognitive psychology to provide a more human-centred approach to the design of ICTs. This discipline pays relatively little attention to the logical aspects of design, such as data modelling, because the discipline regards itself as supplementing existing software engineering practice with insights into the processes of interaction between users and ICTs.

The fourth discipline - Information Systems - aims to cover the areas addressed by Software Engineering and, to some extent, Human-Computer Interaction, from a social/organizational perspective. Given that most computer-based systems are used within organizations, the approach to development emphasises the aims and needs of the organization or social group. Rather than attempting to develop reliable technological artefacts or match technical design with human physiology and cognitive abilities, the aim is to manage the social and economic impacts of technological change. Consequently, Information Systems pays less attention to the engineering aspects of design than the other three disciplines.

Having provided a broad review of the main research areas relating to the use of ICTs in social systems, the chapter concludes with a statement of the key issues relating to their improved design, development, evaluation and use.

2.2 Telecommunications Engineering

In general terms, Telecommunications Engineering is concerned with the use of cable networks and electromagnetic wave broadcasting for long-distance public and private communications. Effectively beginning with the invention of the telegraph by Samuel Morse in 1837, this branch of engineering has come to include telephony, radio, television, satellite and mobile communications. Increasingly, telecommunications networks are being used for data transmission and are being integrated with computer networks, leading to an increasing overlap between the studies of information and communication technologies.

In its present state, and for much of its history, telecommunications engineering can be considered in terms of three sub-areas: product engineering, manufacturing engineering and systems engineering. There is a broad analogy between these sub-areas and the stages of design, development and implementation of computer-based information systems at the organizational level. This analogy is exploited in the identification of key information and software systems development issues at the end of the chapter. Given the aims of this thesis, however, manufacturing engineering is not considered in any detail.

2.2.1 Product Engineering

As the design stage in telecommunications, product engineering refers to the development of prototype devices for use in a telecommunications network. The process leading from the initial concept (or a specific communications problem), through research and development, to a working product suitable for mass production is informed by a substantial body of theory, primarily consisting of: (a) representations of signals; (b) probability theory; and (c) information theory and coding. A brief consideration of the characteristics of these areas illustrates the nature of information and communication as seen by the telecommunications engineer. This perspective provides an insight into the development of the infrastructure required for multi-user and inter-organizational information systems and can be contrasted with the other disciplines considered in this chapter.

Relying extensively upon the theory of communication proposed by Shannon and Weaver (1963), the communication process is typically modelled as shown in figure 2.1. The communicated message is regarded as a patterned signal. No significance is attached to the interpreted meaning of the signal, which is assumed to match the meaning of the message at source, provided that the signal is only minimally disturbed during its transmission. Indeed, the individuals acting as source and user of the message are effectively ignored, as

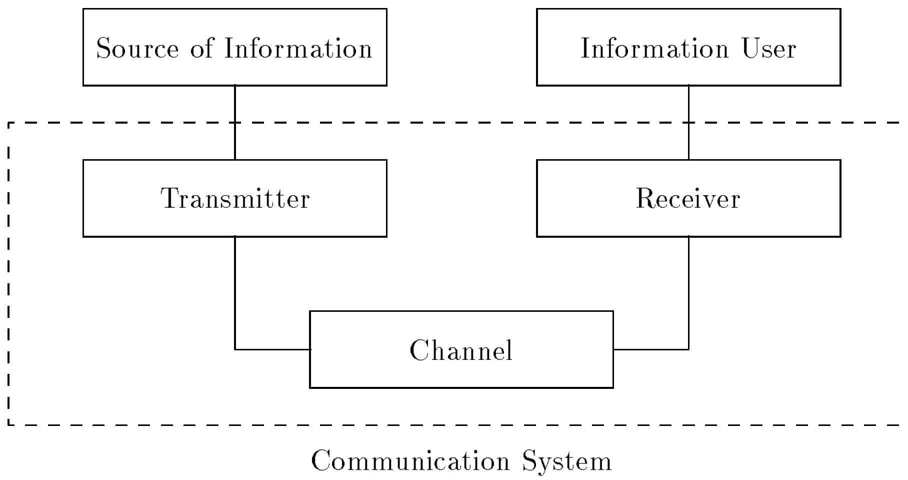


Figure 2.1: The Telecommunications Engineering Model of Communication

illustrated by Haykin (1994, p.3):

> "A source of information may be characterised in terms of the signal that carries the information. A *signal* is defined as a single-valued function of *time* that plays the role of the dependent variable ..."

Having identified the signal as the fundamental concern of the communications process, telcommunications engineering has developed numerous devices for representing and manipulating signals so that the impacts of the transmitter, channel (including interference) and receiver can be studied. In practice, Fourier series and Fourier transforms are the main methods of signal representation, providing a frequency-domain description of the signal. The manipulation of signals is known as modulation and involves the modification of a carrier wave in accordance with the message signal (Haykin 1994). The types of modulation available depend upon whether a continuous sinusoidal wave or a periodic sequence of rectangular pulses is used as the carrier. A speech signal, for example, may be used to modulate a (continuous) radio wave by modifying either its amplitude (AM), its frequency (FM) or its phase (PM). Alternatively, the speech signal could be periodically sampled and the sampled values encoded by modifying the amplitude, duration or position of the rectangular pulses on the carrier wave. Increasingly, however, pulse-code modulation is used, because its conversion of modulated pulses into binary symbols provides a number of advantages:

- the signal is more robust to noise in the environment because the 0s and 1s can be regenerated (i.e. each pulse need only be determined to be a 0 or 1 for perfect recovery of the message, whereas other modulation techniques require the whole range of modulated values corresponding to, for example, the voice, to be discerned);
- diverse sources of information are easily integrated because they are all converted into a common binary format; and
- the transmission can more easily be made secure.

Owing to the consideration of signals as patterns, rather than as intentional constructions, the communication process is regarded as probabilistic in nature. The main concern of the telecommunications engineer is to increase the probability that the message (of unknown appearance) will be received accurately. Given that transmission is imperfect, there are two factors - noise and interference - that can be characterised in terms of probability theory. A further application of probability is the construction of efficient mechanisms for coding messages. The Morse Code, for example, uses short codes for frequently used symbols and longer codes for infrequent symbols.

Information theory makes use of probability to provide a mathematical basis for describing the relationship between a message source and a communication channel. Two concepts serve as the basis for information theory: the *entropy* of an information source and the *capacity* of a communication channel. Entropy is a measure of the average information content per symbol for an information source with a given alphabet (i.e. set of symbols). The calculation reflects the fact that a symbol occurring with probability $p = 1$ is uninformative, whereas the less likely symbols are to occur, the more informative each occurrence of a symbol will be. Where the prior probabilities of the occurrence of each symbol in an

$$H(\varphi) = \sum_{k=0}^{K-1} p_k log_2 \left(\frac{1}{p_k}\right)$$

where φ is the alphabet $s_0 \dots s_{K-1}$ and $p_0 \dots p_{K-1}$ are the prior probabilities of occurrence of $s_0 \dots s_{K-1}$ respectively.

Figure 2.2: Entropy of a Message Source (Adapted from Haykin 1994)

alphabet are known, the entropy of a source using that alphabet is calculated as shown in figure 2.2. The entropy for a source ranges from H(φ) = 0, in the case where the probability for some symbol k is $p_k = 1$ and there is, therefore, no uncertainty, to H(φ) = log_2K in the case where all symbols are equiprobable and there is, therefore, maximum uncertainty.

An important special case is the binary source, where the probability of symbol 0 is p_0 and the probability of 1 is p_1 = 1 - p_0. The calculation for the entropy of binary source is:

$$H(\varphi) = -p_0 \; log_2 \; p_0 \text{ - } (1 \text{ - } p_0) \; log_2 \; (1 \text{ - } p_0) \text{ bits.}$$

The entropy calculations can be extended to consider blocks of n symbols:

$$H(\varphi^n) = nH(\varphi).$$

The use of binary encoding can be exploited to improve the efficiency of communication. This is achieved by assigning binary code-words to each symbol in the alphabet and minimising the average code-word length; the minimum average value being determined by the entropy of the information source, as stated by Shannon's source coding theorem (Shannon and Weaver 1963).

Information theory supports many other aspects of communication process design, including data compaction (the removal of redundancy from a signal) and the calculation of the critical rate of a communication channel, using the channel coding theorem. In doing so, it provides a sound mathematical basis for modelling communications systems and finding efficient solutions to communications problems. Whilst mathematical modelling does not translate directly into practical application, information theory provides a useful basis for identifying general solutions, which can inform the development of artefacts that have the approximate characteristics suggested by the mathematical model.

From a practical perspective, two factors have a significant impact upon the design of actual products: customer requirements and regulation. As a social factor, customer requirements have a reflexive interaction with telecommunications designs, with design being influenced by and influencing customer perceptions and demands. The uptake of mobile telephones, for example, has significantly increased because the initial customer base realised that mobile phones could have an impact on their lifestyle as well as their business conduct. As demand increased and the industry placed more emphasis on the lifestyle aspects of mobile telephony, economies of scale reduced unit costs and mobile phones became more commonplace. This cycle has recurred in the telecommunications sector. Companies selling telephones in the 1920s, for example, formed the Comité Consultatif Internationale des Communications Téléphoniques à Grande Distance to coordinate the advertising of the telephone as a necessary item and not a luxury good. Their campaign placed considerable emphasis upon the use of the telephone to improve the individual's social life (Young 1983).

In terms of product engineering, the primary concern in telecommunications is developing an appropriate interface between the customer (end-user) and the telecommunications network. Interface issues can be usefully considered in terms of two distinct categories of issues: products and services, roughly corresponding with ergonomics and usability, respectively. Service provision usually involves a great deal of hardware and software at the telephone exchange, but the emphasis on usability derives from the systems engineer's goal to make telecommunications networks transparent to their users.

Ergonomic issues include the design of handsets, the layout of keyboards, sound quality and, increasingly, display quality. With mobile communications, power consumption is also an essential consideration, with battery size, power and weight being critical to the acceptance of mobile technologies. Despite the importance of these issues to end-users, few methods exist for integrating ergonomic principles into the design process. One recent approach - High Touch (Lee et al. 1997) - has had some success in this respect. Evaluation also seems quite sparsely covered in the literature. In general, it seems that the typical user is catered for and only current and soon-to-appear services are considered at the design stage. Consequently, the focus is on standard, reliable and low-cost products that are replaced when new services require additional product functionality. Some flexibility is being introduced into telecommunications products, however, through their increased convergence with information technologies. Many telephones, for instance, now have 'soft keys' with programmable functions controlled by the service provider. The development of forward-looking telecommunications standards, such as GSM (Redl, Weber and Oliphant 1998), are also having a positive impact in terms of service quality and flexibility.

Whilst the attempt to cater for typical users has led to generally usable products that are highly robust, some human factors, particularly the needs of disabled users, remain poorly addressed. Whilst specially designed products are available, their benefits are typically confined to the disabled user's home and workplace. Personal telecommunications technologies, such as mobile telephones, can potentially make a significant contribution in terms of improving ergonomics for a wide range of user needs. Although the design constraints are considerably greater than for conventional telecommunications products, the combination of mobility and their personal nature mean that the benefits of such products are less confined to particular locations and the specific needs of user categories can be more precisely addressed through greater product differentiation.

2.2.2 Systems Engineering

Whereas product engineering is primarily concerned with technical quality and ergonomic issues, the development of telecommunications networks demands the consideration of four stakeholder groups:

- government and industry regulators and standards committees;
- service users;
- competitors, suppliers and business partners; and
- the shareholders of the network provider.

Each of these stakeholders imposes demands or constraints on the network provider. Systems engineering is concerned with providing telecommunications services to satisfy the needs of users within the technical, legal, economic and other constraints imposed by the other stakeholder groups. The range of issues that must be addressed includes: geographical coverage; mobility; transmission quality; spectrum/bandwidth efficiency; infrastructure cost; terminal price and complexity; timing factors (e.g. average call connection time); security; privacy of calls; reliability (e.g. number of lost calls); available service options; and future adaptability.

The telecommunications industry has, for almost a century, been subject to government regulation. The most significant aspect of this regulation, from the perspective of network providers, is the issuance of operating licences, which are specific to geographical areas and/or radiofrequencies. Whilst much of the technical standardisation is agreed by committees that have significant industry involvement, licencing decisions are ultimately political. The Europe-wide GSM standard for mobile telecommunications (Redl et al. 1998), for example, was developed by industry in response to the European authorities expressing the need for a Europe-wide standard. The development of this open standard was significantly influenced by network operators, who benefit from the flexibility in switching between equipment suppliers. In contrast, the North American standard was influenced by equipment vendors and, consequently, favours proprietary standards (Goodman 1997). In North America, roaming between areas covered by different network operators is technically more complex and, until recently, involved additional effort on the part of the mobile phone user or callers to mobile phones (for example, dialling different numbers dependent upon the probable location of the person whom one wants to call).

Whilst standards ensure that equipment is technically compatible, it does not guarantee that users will be permitted to communicate between different networks. Thus, an important issue to be addressed in improving quality of service is agreements for interconnections between network operating companies that operate in different geographical areas and/or provide competing services to the same customer group. In the case of different geographical areas, co-operation is rarely problematic from a business perspective (unless the operators intend to expand their present services) and the main problems to be addressed are technical interconnection and charging policies (Goodman 1997). In contrast, a service provider considering connecting to a competitor's network must consider the opportunities and threats in terms of market share.

Transmission quality and bandwidth efficiency are directly related, with quality improving with increased bandwidth or more efficient bandwidth utilization. In practice, the allocation of bandwidth by governments, coupled with the increase in required capacity resulting from increased customer demand, mean that the efficient use of bandwidth is a serious technical challenge. For this reason, various techniques have been developed for increasing the number of voice channels available in a given frequency band. Most significant is the use of signal multiplexing, where multiple voice channels are combined by dividing a channel by time, frequency, time and frequency, or by using a binary coding method (code division multiple access) (Goodman 1997). Whilst the various techniques all increase the capacity of network operators, they have two possible impacts upon end-users. Firstly, some multiplexing techniques cause signal quality to decrease as the number of multiplexed signals increases. Secondly, the complexity of the equipment used by end-users must sometimes increase. This can result in increased costs (particular in terms of upgrading existing devices) and can also affect factors such as power consumption.

2.2.3 Summary

The development of a mathematical basis for telecommunications engineering has made possible the construction of products that are of a high technical quality, yet low in cost. Consequently, although ergonomic principles are not fully incorporated into the design process, telecommunications devices mostly satisfy the needs of typical users. The requirements of users with special needs have been less well-addressed, although the trend towards personal telecommunications may ameliorate this problem.

Telecommunications networks continue to evolve, increasing in complexity and providing an ever broader range of services. The reliability and standardisation of system components means that the services provided by network operators are reliable and presented consistently to end-users, with the complexity of the technical infrastructure being largely transparent. These qualities have come at the expense of flexibility, however, because users typically have fixed functionality devices that sometimes need to be replaced to gain access to new services on a network. This situation is changing as a result of the convergence of information and communication technologies and a more forward-looking approach to developing technical standards. The interconnection of networks is not always satisfactory from the users' perspective because the policies and business interests of network operators sometimes limit the exploitation of available technologies. The influence of stakeholder groups over the formulation of national and international standards and regulations has significant implications in this respect, as illustrated by the comparison of European and US mobile telephony (see section 2.2.2).

2.3 Software Engineering

As noted in the previous section, software is increasingly being used to improve the flexibility of telecommunications products, which, in sharp contrast with modern ICT-based systems, usually have fixed functionality. Unfortunately, the enormous flexibility of ICT-based systems has not been well controlled, leading to low technical quality and frequent difficulties in gaining user acceptance. In attempting to understand these differences so that flexibility, technical quality and usability can be achieved, it is instructive to consider two basic differences between software and telecommunications engineering:

- software development methodologies usually regard the physical aspects of a problem situation as presenting peripheral constraints to logical design, rather than as the basic medium for development; and
- software systems have typically not been constructed from standard components of verified design.

The main consequence of focusing on the logical rather than physical aspects of design is that the limitations of computer and software technologies are not accommodated by the methods for exploiting them. Thus, whilst telecommunications engineering compares the abstract notion of entropy with the concrete limit of a communication channel's capacity, software development methodologies model problem situations in only abstract terms. Logical data models, for example, are usually finalised before considering the basic properties of the ICT artefacts being used to construct the information system. This characteristic of software engineering is considered here as a possible cause of low levels of software reliability and user acceptance for computer-based information systems.

A second consequence of ignoring the physical aspects of the problem situation is that little attention is paid by software engineers to how software systems are used (Gough and Hemingway 1999). This is also true of telecommunications engineering, as illustrated by the lack of integration of ergonomic factors into the design process. The complexity of modern computer systems, however, means that an acceptable level of usability is often difficult to achieve.

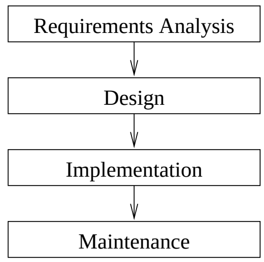
The tradition of developing software systems *de novo*, rather than using verified systems components further complicates the task faced by software engineers. Numerous techniques and technologies are being explored by the software engineering community for managing this complexity. Software reuse models (Boulange 1998) attempt to establish basic systems components; CASE tools (Parkinson 1991) and fourth generation languages (Wojtkowski 1990) focus on particular classes of problems; and software process improvement methods attempt to institutionalize development practices that reduce software errors (Fitzgerald and O'Kane 1999).

The following analysis examines the above three issues - ICT artefact characteristics, users' working environments and software/project complexity - from the perspective of improving the satisfaction of customers and end-users. Several lifecycle models (illustrated in figure 2.3), reflecting different approaches to addressing these issues, are considered below, followed by a summary, which identifies the main challenges to establishing technical quality and introducing user-centred development practices.

2.3.1 Conventional Software Engineering

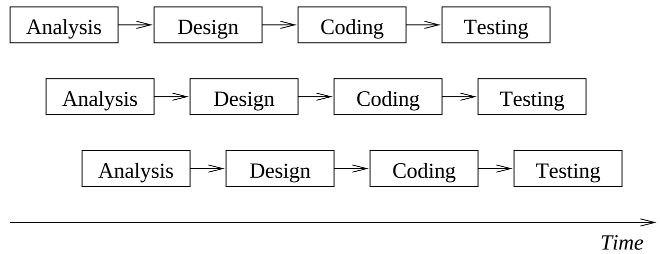
Despite the two fundamental differences between software development and other types of engineering, conventional engineering methods have inspired many attempts at systematizing the software development process, the most notable example being the conventional 'waterfall' model (Royce 1970). Although the limitations of this model are well recognized, its influence can still be discerned in many methodologies that are currently in widespread use (for example, SSADM, Goodland and Slater 1995). Figure 2.3(a) illustrates the waterfall model as described by Easteal and Davies (1989).

The waterfall model regards software development as an essentially linear process. Beginning with an informal statement of requirements by the customer, a detailed problem specification is constructed by an analyst. The specification is described in abstract terms and used to create a logical design for a computer system to satisfy the specification. This design is then translated into a detailed technical design, which is implemented using a specific combination of hardware and software. The implementation is finally tested to



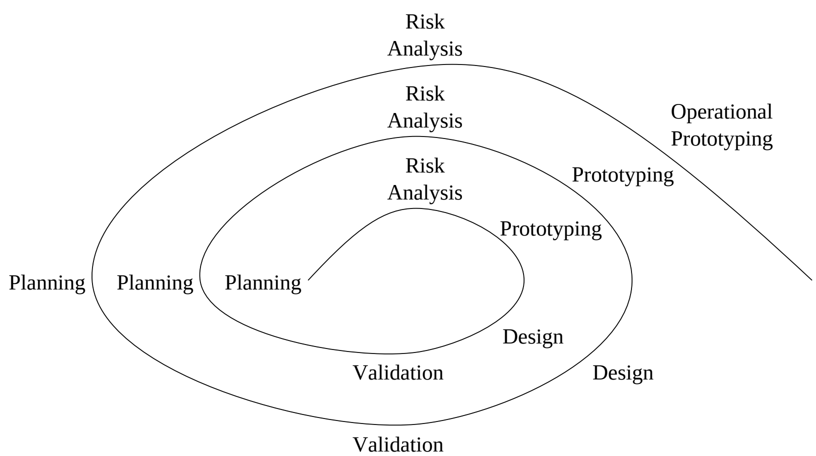
(a) The Waterfall Model

Adapted from Easteal and Davies (1989)



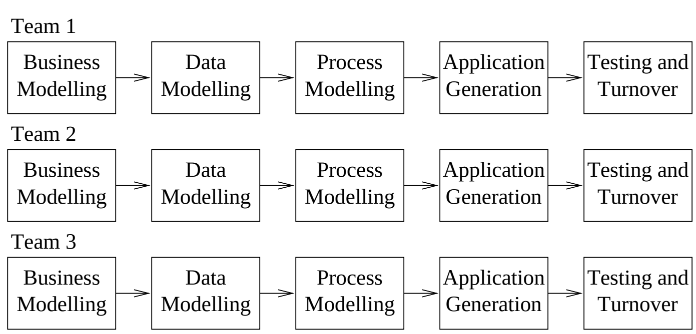
(b) The Incremental Model

Adapted from Pressman (1997)



(c) The Spiral Model

Adapted from Boehm (1988)



(d) The Rapid Application Development (RAD) Model

Adapted from Pressman (1997)

Figure 2.3: Four Lifecycle Models

ensure compliance with the specification before being released for operational use.

The basic lifecycle model provides no explicit mechanism for taking into account the characteristics of the available ICT artefacts until the logical design is finalised. Consequently, consideration of the technical skills required, the complexity of programming and testing, or of technical quality and efficiency is limited. Various tools and techniques - which often assume the waterfall model, but can be used in the context of different lifecycle approaches - have been developed to address the problems of identifying user skills, selecting hardware and software and estimating software complexity. These are discussed in detail in section 2.3.7.

In terms of accounting for the context in which a software system will be used, the waterfall model has two serious and well-documented deficiencies. First, the customer - who defines the system's requirements - and the users are virtually excluded from the design process, which makes extensive use of technical design notations. Consequently, many design decisions that may affect the fit of software into the users' working practices are based upon the developer's interpretations of what the customer requires, supplemented by his or her understanding of what constitutes a user-friendly system. This approach has been cited as a reason for low levels of user acceptance (Willcocks and Mason 1987). Second, the one-off statement of user requirements, which is documented in a way that is difficult for users to understand, results in the analyst being faced with the onerous task of integrating the software system into modified business processes. Owing to the rigid nature of the software development process, it has been common for integration to proceed by making the workers fit around the computer system (Avison and Fitzgerald 1995). Such technology-centred changes can have negative impacts in terms of job satisfaction and business performance.

The low level of user participation achieved when using the waterfall model stems from its document driven nature. The criterion for progressing through the lifecycle is the completion of key documents at each phase (Boehm 1988). The document generation process begins in the customer's and/or users' language but becomes increasingly technical as analysis and design proceed. This process of translation demands that the user either becomes increasingly involved in technical detail or - more likely - allows the software developer to make design decisions. From the technical design stages onwards, the customer/users are completely excluded from the development process until the software is sufficiently complete to be exposed to end-users.

The limitations of the above process are well understood. Requirements analysis notations, such as Data Flow Diagrams (DeMarco 1979) and Entity-Relationship Models (Chen 1976), have been found to be less intuitive to non-technical persons than was previously thought (Moody 1996, Chan 1998). Indeed, empirical studies suggest that even software engineering practitioners have a limited appreciation of the semantic constructs used in Entity-Relationship Models (Hitchman 1995). Thus, although it has been accepted that the customer and users will be excluded from development at some stage, this may occur earlier in the development process than has been assumed, owing to the technical nature of conventional modelling tools.

Having stated the problem and made some contributions to its analysis, the customer and users are faced with a long delay before the software is ready for use. Evidence suggests that customer and user requirements change over time (Eason 1988). Consequently, the long delay incurred by following the waterfall lifecycle risks developing software that, when delivered, is solving the wrong problem. Another factor reducing software acceptance arises from the exclusion of the customer following requirements specification. Evaluation criteria are stated by the problem owner, but are then subject to interpretation by the software developers to define when the software requirements have been met. These metrics are then combined with the developer's technical criteria to determine whether the software is of an acceptable quality. As illustrated by decision theorists (see, for example, Alt 1971 and French 1986), preference elicitation is notoriously difficult and the effective aggregation of different stakeholder preferences presents futher difficulties. Thus, it seems unlikely that even the most user conscious developer can genuinely make choices on the customer's or users' behalf.

A final obstacle to gaining user acceptance stems from the separation of functional and non-functional requirements, which occurs at the analysis stage (Pressman 1997). Factors relevant to the system but not modelled in information terms are classified as non-functional and, consequently, regarded as design constraints. Given that many working norms, corporate standards and other issues fall into this category, there is a significant risk that adopting this perspective will result in the encoding and, thereby, reinforcement of old working practices. If, as evidence suggests, the greatest benefits from systems development stem from the adoption of more effective methods rather than efficiency gains (Strassmann 1990), the waterfall model may actually limit the benefits of any development activity.

As with software and hardware selection, the waterfall model is somewhat limited in terms of managing project complexity and risk. A great deal of progress has been made in addressing these issues through the adoption of process improvement standards, which have national and international recognition and through the developmeent of project management tools. A selection of these standards and tools is reviewed in section 2.3.8.

2.3.2 The Iterative Waterfall Model

As illustrated above, the waterfall model is too constraining because early decisions sometimes need to be revised in light of new information and/or changed circumstances. Before considering alternative lifecycle models, the impact of iteration on the waterfall model is examined.

A key criticism of the conventional model is that the logical design is finalised before technical issues are seriously considered. Where iteration is introduced, the logical design can be revised if it is found to be impracticable given actual technical constraints. Two points can be noted with regard to this approach to accommodating the characteristics of ICT artefacts. Firstly, it is likely to lengthen the already time-consuming development process. Secondly, it provides no explicit mechanisms for balancing the criteria for good logical design against the practicalities of implementation.

Few of the serious constraints on user participation are alleviated by the use of iteration because they arise from the technical, document-driven nature of the individual phases. Iteration has no significant impact upon the way in which each phase is conducted, but simply identifies elements of earlier documents that need to be revised to resolve difficulties at subsequent stages. The nature of the documentation itself, which is the key limitation to user participation, remains unchanged. A note of caution must also be raised regarding the process by which documents are revised. Objections would be expected if technical experts adjusted the documents developed by analysts and users to make implementation more convenient. Alteration of documents without the original authors' involvement may also deviate from the actual requirements of users. Where such revisions must be approved by the analysts and users, however, the problem of technical competency arises. Given that the later stages in the lifecycle require considerable technical skill, it is likely that the reasons for changing documentation will also be technical. Consequently, users are faced with requests to balance their requirements against technical issues that they do not fully understand. This pressure may be considerable if the software developers are from an external organization that competed for the development contract. In this respect, iteration may result in a more technology-driven development approach.

2.3.3 Incremental Development

A simple variation on the waterfall model is the incremental approach to systems development (figure 2.3b). This approach involves partitioning a software system into discrete components and ordering them in terms of their importance to the customer and/or users. The waterfall process is then applied to the various components, beginning with the core functionality and ending with the 'nice to have' features. Whilst the approach offers no advantages over the waterfall model in terms of addressing technical issues, it claims two key benefits in terms of user involvement:

- core functionality is available to the user more rapidly than via the conventional waterfall model; and
- feedback from initial phases can be accommodated by later increments to the system (Pressman 1997).

In the absence of systematic comparisons between the various lifecycle models, several reservations can be made regarding the benefits of this approach. Design decisions are still taken by the software engineer. Thus, the increase in user involvement is limited. Given that core functionality is likely to represent a significant proportion of the investment in software, the extent to which negative user feedback can be accommodated is unlikely to be considerable. Minor criticisms could be addressed in a timely manner through incremental development, but it is doubtful that this represents a significant advantage over the waterfall model, particularly in view of the need to develop new functionality at the same time as modifying the structure of existing code. Furthermore, releasing components to end users could conceivably constrain the designer's future technical decisions, which may affect the project's overall success. A likely example, which the addition of iteration to the waterfall model addressed, is that later design phases may call for the revision of earlier design decisions. The operational use of core functionality in the case of the incremental model may complicate, or even preclude, such revisions, particularly where the proposed changes affect interfaces that are rated highly by users or require the modification of data models that have been used to encode corporate data.

2.3.4 Prototyping

In comparison with the use of an incremental approach, the use of prototyping at the analysis and/or design stages can have a more substantial impact upon software development. A significant problem for many customers and users is that, although they can state their problem, they have insufficient experience of possible solutions to state their preferences between them. Consequently, their contribution to design decisions is very limited.

Prototyping can help to overcome this by providing the various stakeholders with some experience of alternative software solutions. Furthermore, experimentation with alternative work processes can help customers and users to clarify their requirements. The potential for improving dialogue between customer/user and analyst is significant. Whereas users find data models and technical specifications difficult to interpret and evaluate, they can usually state without difficulty their preferences between systems that they have used.

Whilst prototyping techniques can both prolong and raise the standard of user-analyst dialogue, even when simply introduced into the waterfall approach, they do have a number of limitations. Firstly, the development and use of prototypes must be carefully managed in order to avoid slowing down the systems development process or unduly increasing its cost. Both the cost and time for prototype development will vary with the method used, which can range from a paper mock-up to a 'rough and ready' version of the system (Olle et al. 1988) . Care must be taken to ensure that the prototype is only developed to a sufficient degree to allow customers and users to clarify their needs and preferences. Secondly, prototyping is only useful for analysing some aspects of systems and less useful for others. As noted by (Olle et al. 1988, p.203), prototypes "...greatly facilitate getting the visible aspects of a design specification correct" but it is "...much harder for the user acceptor to evaluate correctness, completeness and consistency of elements that show up only indirectly, such as complex computation and integrity rules." Thus, whilst prototyping can improve user participation, it will only address certain requirements and design issues. The discussion of other issues relies upon the use of traditional analysis techniques. It is, therefore, possible that the unmanaged use of prototyping could be detrimental to the development process, encouraging users to focus on what is easy to prototype and visualize, rather than what is critical to the success of the project. A related risk is that users may focus their efforts on refining the prototype rather than evaluating it as one of many possible solutions. In other words, prototyping might lead to the search for a 'local maximum' rather than placing an emphasis on describing the general characteristics of a 'globally optimal' system.

Having proceeded, with the aid of prototypes, to develop the requirements specification and design documents, there are two options for progression: to throw away the prototype or develop it into a working system (Pressman 1997). The throw-away option regards the prototype as a means by which to improve the quality and relevance of the analysis and design documents. Having completed these documentation stages, the systems lifecycle proceeds as normal, with the software engineer developing a completely new system based upon the technical design. This approach is time consuming (as with the waterfall model) and may be frustratingly slow for users who, having direct experience of what the system will be like, are eager to use it. Pressman (1997) suggests that it can be difficult to convince users that it is worth waiting for the properly engineered system and comments that some developers are pressurised into providing the throw-away prototype for operational use. Clearly, this is likely to fail the user in the longer term because integrity, reliability and other quality factors are compromised. In terms of addressing technical implementation issues, the only benefit of prototyping arises when technologies that might be used in the operational system are used at the prototype stage.

The second alternative, to purposefully develop the prototype into a workable system, has been questioned on similar grounds to the release of a throw-away model. It is maintained by many software quality specialists (Sommerville 1992, p.108) that quality cannot be engineered into intrinsically flawed and unstructured software. This argument is reinforced by evidence that continued maintenance causes deterioration in the structure of software and consequent performance degradation. In the event that it is possible to transform a prototype into a well-engineered system, it is arguable as to whether this will offer any time, cost or user acceptance benefits in comparison with throw-away prototyping. Given the poor structure of the prototype code, it is quite plausible that developing the final system from scratch will be the quicker, cheaper and less risky option, particularly for projects of some complexity. In terms of attending to ICT characteristics during design, this option is most unsatisfactory. Instead of making an informed choice of ICT artefacts, the selection is made by default. Thus, this option ought really to be confined to those cases where the prototype only requires minor (especially cosmetic) adjustment to become operationally useful. A final point to note is that prototyping encourages a focus on the requirements specified by users. In addition to these, the software engineer must also develop test plans and technical quality metrics so that the system can be properly engineered before release to the users. Evolutionary prototyping blurs the boundaries of analysis, design and development stages so that it becomes unclear when and how such quality issues can be addressed. Thus, it is reasonable to suggest that the use of prototyping requires a different view of the lifecycle than that offered by conventional engineering models. One such alternative model is the spiral model (Boehm 1988, Boehm et al. 1998).

2.3.5 The Spiral Model

The spiral model (figure 2.3c), as its name suggests, involves a number of cycles, each leading further towards an operational system. As a risk-based, rather than document driven, process, each cycle consists of:

- identifying objectives, constraints and alternative actions;
- the evaluation of possible actions and the management of risk;
- the elaboration of process and product definitions; and
- planning of the next cycle.

Central to the effective management of risks is the use of prototyping and other techniques for clarifying stakeholder requirements and for enabling a more informed and inclusive approach to making design decisions. Recently, negotiation support has been added to the spiral model to further assist in reconciling diverse stakeholder objectives (Boehm et al. 1998) . In many respects, the spiral model is almost a meta-model, accommodating the waterfall and other approaches to development. The aim of the spiral model is to enable project teams to choose an approach suited to their project type and to allow them to retain continuity in their project if their development approach needs to be adapted.

The introduction of support for stakeholder negotiations makes the spiral model much more user centric at the analysis and design stages, whilst ensuring that project boundaries and issues of technical quality are also addressed. It is less obvious whether the spiral model can make a significant contribution to improving participation throughout the development process. One possible obstacle is that the stakeholders and risks to be considered by each cycle are formulated by the project team without any detailed support by the spiral method. Thus, which stakeholders are included in the cycle and which risks are identified as most important is determined by the analysis skills of the project membership at a given time.

The risk-based approach clearly demands that software developers are skilled at stakeholder and organizational analysis as well as at assessing risks. These are not traditional skills of software engineers and, when prototyping is extensively used, the systems analyst is often given a diminished role because user and programmer can discuss the software face-to-face. The need for these skills is emphasised by the lack of support provided by the model for stakeholder, organization and risk analysis. No explicit guidelines are provided for identification or classification, nor has an attempt been made to present a comprehensive taxonomy of software risks.

A further weakness of the spiral model is its lack of attention to longer-term management and planning. For the software project to be successful, it must achieve long-term objectives agreed by the group. The spiral model places a good deal of emphasis on establishing objectives for the current cycle, but provides no mechanism for ensuring that objectives and decision criteria are consistent between cycles (Wolff 1989).

2.3.6 Rapid Application Development

The notion of rapid prototype development using computer-based tools has been extended, in the context of a waterfall-type model, resulting in the rapid application development (RAD) approach to software engineering (see figure 2.3d). This approach focuses on information systems problems (particularly those encountered in business environments) with the following characteristics:

- processes are based around transactions;
- inputs are well-defined and mostly expressible in alphanumeric form (whether user or machine input); and
- outputs either list attributes of a single entity or of a few related entities, or provide summary information about classes of entities.

The first constraint excludes the use of RAD for real-time systems where the time frame for data processing is dictated by the environment. Such systems require very different analysis procedures and technical architectures (Sommerville 1992). The other constraints relate to the visual approach to development that is utilized by RAD to achieve developer productivity gains.

Fourth generation languages and other RAD tools typically support the development of object or relational databases with menu-navigated forms-based interfaces and tabular reporting (sometimes limited types of graphs are also available). Thus, the solution is confined to a fairly narrow range of options and can be regarded as an explicit 'local maximum' approach where, for typical information systems problems, this will provide a satisfactory solution. Requirements for complex data processing must be elicited from the user by conventional analysis methods and encoded by the developer. Thus, RAD is not well-suited to systems with low levels of user interactions and complex data processing algorithms (Jones and King 1998). Given the right type of problem, however, RAD offers similar benefits to users as prototyping methods, with the additional benefit of providing an engineered product without significant delay.

From the user/customer perspective, RAD has a number of potential drawbacks that require careful management. Of the various difficulties identified by Ljubic and Stefancic (1994), for example, several indicate that the pace of change may be too rapid for some users. More generally, the emphasis on rapid software provision may encourage a short-term perspective, producing software that quickly becomes out of date. If this occurs, RAD simply quickens the software lifecyle and the claims of cost savings and improved alignment of software with business goals are less convincing over the long-term.

The RAD approach also has some technical limitations. Where large problems are encountered, for example, the project must be partitioned and components developed in parallel (as shown in figure 2.3d). Thus, problems of system integration may arise. Similarly, the focus on rapidly developing software within the confines of a narrow system boundary may create problems of integrating the RAD engineered product with existing software (Jones and King 1998). For this reason, Pressman (1997) suggests that RAD should be avoided where technical novelty or systems integration requirements are high.

2.3.7 Technical Skills Assessment and ICT Artefact Selection

A key aspect of developing ICT-based information systems is the *ex ante* evaluation of project costs, resource requirements and the demands of a project on staff time and skills. Early texts on systems analysis, such as Daniels and Yeates (1971), paid considerable attention to these issues. This is in contrast with the modern software engineering literature, which does not even provide comparable discussion when discussing systems development and implementation as a whole (see, for example, Sommerville 1992). As suggested above, a possible reason for this is that the characteristics of concrete ICT artefacts have typically been regarded as constraints. Consequently, the continual improvement in hardware performance have lessened the perceived need to consider such issues. It is argued here that issues of hardware and software selection are equally, if not more, important now than they were in the early 1970s because the range of hardware and software options is considerably greater and the need for integrated technical systems is increasingly reflected in user and business requirements.

The first occurrence of evaluation in a software development project involves collecting information to clarify the scope of the problem situation. A key element in such an evaluation is cost estimation. Given that user requirements are often unclear during the early stages of development, initial estimates are always very approximate. Nevertheless, the information they provide makes an important contribution to 'strategic' decisions, such as the choice between bespoke and packaged software. As information is collected and used to clarify requirements and guide project planning, the broad characteristics of the software project begin to take shape and will themselves inform the estimation of project costs. Thus, in considering the nature of project evaluations, it is useful to distinguish between a number of project types, as shown in figure 2.1.

	Bespoke	Component	Package	Service
In-House	Yes	Yes	No	Yes
Contract	Yes	Yes	Yes	No
Vendor	Yes	Yes	Yes	No
Provider	No	No	No	Yes

Table 2.1: Types of Computer-Based Systems Development

During the later stages of analysis and throughout the design process, technical issues must be considered. These range from application-oriented requirements, such as guaranteed response times, to system-oriented requirements, such as compatibility with existing hardware. The following list is adapted from Bennatan (1992).

- interfaces to existing equipment;
- database requirements;
- communications and network architecture;
- adopted standards;
- system reliability;
- timing constraints;
- programming languages;

- computer hardware availability;
- reuse of existing components;
- reusability of components to be developed;
- data integrity; and
- data security.

The use of software packages is unique among the options for providing software to end-users, its main advantages and disadvantages being summarized in table 2.2. The provision of information services, either by a specialist internal department or an external vendor raises a separate set of issues that are best considered in terms of the wider information system. These options are, therefore, considered in section 2.5.4, in terms of insourcing and outsourcing respectively.

Advantages	**Disadvantages**
Less programmer time	Suitable packages can be difficult to locate
Faster implementation	Possible inefficiency of use
Superior product	Need for tailoring
Lower Cost	Inflexibility
Low maintenance cost	High capital outlay
Staff freed for other work	Lack of acceptance of installation

Table 2.2: The Advantages and Disadvantages of Packaged Software Systems

Considering the *ex ante* evaluation of bespoke and component systems, various technical, cost and skills metrics can be used in the analysis. The Constructive Cost Model (COCOMO) is reviewed here to illustrate a typical software engineering approach. COCOMO is an algorithmic model that uses lines of code (LOC) estimates to make cost projections (Bennatan 1992). There are numerous techniques for providing LOC estimates, which can provide widely differing results. This variation is, in itslf, a limitation on the validity of COCOMO and derivative models.

One claimed advantage of COCOMO is that it enables the project evaluator to factor the number and quality of personnel into a cost estimate using multipliers derived from empirical data. A project that will generate many lines of code and involves only inexperienced engineers, for example, will use a large multiplier to adjust a typical cost estimate.

Another characteristic of software projects that can be addressed by COCOMO-derived models is the type of application being developed. Bennatan (1992), for example, revises COCOMO to distinguish between system, algorithmic, service and data processing software. Again, multipliers are used to reflect the tendency for system software to be more complex and time-consuming per LOC than data processing software. The reliability expected of the software system is accommodated by yet another multiplier. Boehm (1981) used the following classification to develop the original COCOMO model, although variations have again been proposed:

- slight inconvenience;
- losses easily recovered;
- moderately difficult to recover loss;
- high financial loss; and
- risk to human life.

Other factors, such as the software development environment can be accounted for by further adjusting the cost model.

It should be noted that the COCOMO multipliers are highly subjective in terms of classification and the recommended values are only very approximate. Whether adjusting an already rough financial estimate with such multipliers is informative or misleading is open to question. The various risks involved in using quite arbitrary values of this sort are illustrated in more detail in the review of similar approaches to information systems evaluation (see section 2.5.4).

In addition to project level metrics such as those used by COCOMO, there is an enormous range of lower-level software metrics, addressing factors such as correctness, reliability, efficiency, integrity, usability, maintainability, flexibility, testability, portability, reusability and interoperability. Although many of these factors can only be assessed subjectively, software engineering approaches have typically insisted upon the quantification of all factors, using arbitrary scales where direct measurements cannot be made. Agreement on the way in which subjective metrics should be constructed has not been achieved within the software engineering community. Consequently, the number of metrics available to address each of the above factors is also considerable (Bennatan 1992). It is, therefore, useful to consider how an informed choice might be made between the metrics currently available. Ejiogu (1991) suggests the following evaluation criteria:

- the metric should be easy to learn and to compute;
- the metric should be empirically and intuitively persuasive;
- the results of applying the metric should be consistent and objective;
- units and dimensions should be applied consistently;
- the metric should not be programming language dependent; and
- the metric should provide information that can be acted upon to improve the quality of the software product.

The extent to which such metrics are used by project teams depends upon a number of factors, including the team's (and organization's) expertise and maturity, the complexity of the project and the risks involved.

2.3.8 Managing Complex Software Projects

For software to function well over the long term, it must satisfy high technical standards in terms of reliability, integrity, security and maintainability. Given that quality cannot be engineered into intrinsically poor software, guaranteeing technical and usability standards can only be achieved by ensuring that the software development process is itself well-managed. Whilst standard project management tools and techniques, such as programme evaluation and review technique (PERT) and the critical path method (CPM), can be used, the standards and methods that directly address the complexity of systems development and the management of software risks are of particular interest. Standards for improving software development processes can be divided into two categories: (a) statements of minimum requirements and (b) methods for continuous process improvement. Key standards of each type are considered below, followed by a brief review of tools and techniques for estimating and managing project risk.

ISO9000 is a general quality standard that is intended to be applicable to any business process. The standard has been interpreted specifically for software processes, resulting in ISO9001. ISO9000-3 further contributes to the standard's application to software systems by mapping ISO9001 onto the ISO12207 standard, which also addresses quality issues in the software lifecycle. Like many other standards, the ISO9000 family lay out minimum requirements for an acceptable quality system. Whilst such standards may be reassuring for customers, they are less valuable for the software developers themselves, providing limited guidance on *how* to change existing practices when a quality-related deficiency is identified. Furthermore, Paulk (1995) argues that ISO9001 only addresses continuous process improvement - the cornerstone of total quality management - implicitly and, consequently, surmises that the standard is best suited to addressing quality issues in software contracts, rather than for providing guidance to software developers wanting to improve their working practices. The most serious limitation of the ISO9000 family, however, is that its generality leaves too wide a scope for interpretation. This, coupled with the lack of a requirement for auditors to have software engineering expertise as well as skills in quality assurance have led to instances where organizations have received accreditation despite having an *ad hoc* approach to developing and testing software (Paulk 1995).

In contrast to the generic ISO9000 standards, the Capability Maturity Model (CMM)

(Paulk 1995) developed by the Software Engineering Institute at Carnegie Mellon University was originally intended to improve the development of defence avionics software but, largely as a result of the absence of comparable commercial standards, has become a *de facto* standard in the USA. Owing to its origins in complex defence systems, CMM addresses only the software development lifecycle and, unlike ISO9001, ignores issues of software support and other systems issues. CMM does, however, place the onus on continuous process improvement and, because the model is specific to software engineering, provides a considerable amount of guidance on *how* to improve software quality and institutionalise software quality management practices.

The basis of the CMM approach is a classification of software development capability according to the stages shown in table 2.3. Rather than state a minimum level of conformance, CMM enables organizations to assess their strengths and weaknesses using a software process appraisal. The result is a set of recommended 'Key Process Areas' that, if successfully acted upon, will raise the development capability to the next level of the CMM and, it is claimed, thereby improve the quality of software products and reduce the costs of rework/debugging. Diaz and Sligo (1997), for example, suggest that the software defect rate can fall by as much as half each time a project advances by one CMM level. Central to gaining such benefits are detailed documentation of quality related procedures, the institutionalization of the documented practices and the use of peer reviews as an effective means of locating software defects.

Whilst CMM is claimed to benefit small and large organizations alike (Herbsleb et al. 1997) , several limitations have been identified. Firstly, CMM implementation is time consuming and costly, yet the benefits mainly accrue at levels 4 and 5, which may take several years to attain. Furthermore, Diaz and Sligo (1997) also present limited evidence to suggest that productivity may decline in early stages of process improvement owing to the effort required to learn and implement new process activities. It seems plausible to suggest that, whilst smaller companies can benefit from CMM in the long run, many of them may not be able to tolerate the necessary short and medium term losses. For this

Table 2.3: The Capability Maturity Model (Adapted from Herbsleb, et al. 1997)

reason, the argument made by Saiedian and Kuzara (1995) that CMM could be made more viable by adapting it to provide some payoff at earlier stages is credible. As noted by Herbsleb et al. (1997) , earlier payback would also help to retain management and software developer commitment to software process improvement by providing them with early evidence of tangible benefits.

Other limitations of CMM include the risk that its internal focus may reduce customer satisfaction (Herbsleb 1997); its lack of attention to issues of organizational culture and change management (Saiedian and Kuzara 1995); and the risk of placing too much emphasis on the, essentially arbitrary, maturity level numbers (Saiedian and Kuzara 1995). Perhaps the greatest limitation is that CMM only addresses software development and not the entire systems lifecycle. Consequently, technical quality may be improved, but software provision, from the customer's perspective, might remain unacceptable. Relating CMM to information systems evaluation methods, such as those discussed in section

2.5.4, might help to overcome this limitation. Nevertheless, CMM will continue to be widely used and is being revised to address some of the above criticisms. Furthermore, the high level of guidance that it provides makes it a preferred basis for software process improvement for many organizations (Paulk 1995). This strong position is reaffirmed by the efforts to adapt CMM guidelines to achieve ISO9001 conformance (Paulk 1995) and ISO12207 conformance (Ferguson and Sheard 1998).

Whilst standards for process improvement have had a significant impact on the technical quality of many organizations' software, the use of specific tools and techniques for managing software complexity and risk also plays a major role. An emerging use of metrics at the analysis phase, for example, is the use of simple linguistic metrics to estimate the quality of a requirements specification. A notable example is the Automated Requirements Measurement tool developed by NASA's Software Assurance Technology Center (SATC) (Rosenberg, Hammer and Shaw 1998). This tool uses simple metrics, such as lines of text, number of continuance phrases, number of weak phrases and number of option words or phrases in a natural language requirements specification to present a broad assessment of its quality. The SATC has had some success in demonstrating correlations between, for instance, a high number of weak phrases (e.g. "the user completes the field as appropriate") and the number of errors resulting from software engineers misinterpreting the requirements document. Similarly, at the design stage, metrics for estimating software design complexity can be used to ensure that the implementation of software and its future maintenance will be manageable (see, for example, Henry and Kafura 1984).

2.3.9 Summary

The discussion of software development lifecycles has identified three issues that are central to raising the quality and user acceptance of software systems: (a) paying more attention to the physical aspects of ICT artefacts; (b) improving user participation to ensure a fit between the software system and working practices; and (c) managing the complexity and ensuring the quality of software through improvement of the development process.

Several approaches to development, ranging from conventional bespoke software to rapid application development and software package selection were considered.

In reviewing the various lifecycle models, it was shown that they do not explicitly support consideration of balancing good logical design against the practicalities of technical implementation. User participation issues were also poorly addressed, leading to difficulties in terms of ensuring that the software contributes to an effective information system that supports working practices. Changes in the lifecycle can make a limited contribution to improving participation, but this is limited if conventional modelling tools continue to be used. Prototyping has made a significant contribution to supporting the requirements analysis process, but requires a radically different lifecycle approach. The spiral model has been proposed, but the transition from document to risk focus raises other difficulties. The most significant improvements in terms of the three key issues relate to the improvement of software development processes. The use of standards, particularly the Capability Maturity Model, has had a substantial impact on the management of complex software projects and a consequent impact on software quality (Saiedian and Kuzara 1995). The use of evaluation techniques to support decision making has been less successful, particularly at the early stages of development when the project as a whole needs to be considered. It was suggested that these high-level issues are best considered in terms of the broader information system. Another issue raised in the above discussion was the significant impact of domain-specific knowledge on the power of various stakeholder groups and that power relations generally have a considerable bearing upon the development process.

2.4 Human-Computer Interaction

Over the past thirty years, the impressive rate of technological change has been reflected in a fundamental shift in how and why ICTs are used. The computer has been exposed to non-technical users and is now expected to satisfy needs that are not readily expressible in a formal manner and that will be subject to frequent change. The convergence of computing and telecommunications has also led to a radical change in the nature of human-computer interaction. Not only has the manner in which users interact *with* ICTs changed, but users are increasingly interacting *through* ICT artefacts. The field of Human-Computer Interaction has emerged in recognition of these changing needs and the challenges that they present to the systems and software developer.

Most HCI researchers regard the integration of user-centred techniques into software engineering as their primary goal (see, for example, Dix et al. 1993 and Sutcliffe 1995). It is not surprising to find that there is a strong tendency for HCI research to focus on augmenting various parts of the software engineering process with user-centred tools and techniques. The more radical alternative of reconsidering the very nature of the software lifecycle from the perspective of the end-user has received little attention. The aim of this section, therefore, is to consider:

- whether a user-centred approach to systems development can be achieved by adding human factors tools and techniques into the previously discussed lifecycles; and
- what the metrics developed by the HCI community actually measure and what such measurements can be expected to achieve in terms of influencing the nature of the systems lifecycle.

The tools and techniques developed by the HCI community tend to assume either a waterfall approach to software development or an approach that supports the extensive use of prototyping. They are applied to the problem domain at a specific phase of the lifecycle and feed into the output of that development phase. Considering the waterfall model, techniques have been developed to contribute to almost every phase of the lifecycle (table 2.4).

The extensive use of prototyping for designing and evaluating interfaces is also evident. Despite this coverage, few holistic approaches to user-centred systems development have been proposed. Two methods that are comparatively broad in scope are considered here:

Requirements Analysis	Task Analysis
Design (Architecture)	-
Design (Detailed)	Dialogue Design
Implementation	Windowing Architectures Windowing Toolkits User Interface Management Systems
Testing	Usability Evaluation

Table 2.4: Correspondence of HCI Techniques with the Waterfall Model

the "psychological and organizational tools" developed by Clegg et al. (1996) and Usability Engineering (Nielsen 1993).

2.4.1 Clegg's Psychological and Organizational Tools

Clegg et al. (1996) propose a collection of five interconnected tools for incorporating psychological and organizational issues into the development of computer-based systems (see figure 2.4). The tools are aimed at encouraging a user-centred approach to organizational process and job redesign. Some of the tools can be used to develop organizational processes that do not involve any computer use. The task allocation and usability tools, however, are intended to address technological issues. Task allocation is the process of dividing activities into distinct tasks and assigning them to the human, the computer or both, based upon an analysis of four factors: technology, work environment, task and people. The process of allocation interacts with the job design activity to ensure that the tasks assigned to humans can be combined to provide better jobs than those in the current work system. The usablity tool "evaluates the usability of a prototype or operational computer system" Clegg et al. (1996, p.497). Workers are required to use the computer system to perform tasks representative of their workload and their effectiveness and satisfaction with the computer system are assessed.

Whilst it is clear that some benefits are likely to arise from task allocation and usability evaluation, it is instructive to consider how the five tools relate to the software development

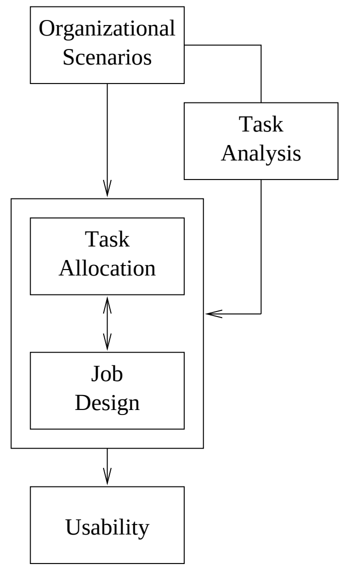


Figure 2.4: Psychological and Organizational Tools for Systems Development (From Clegg, et al. 1996)

lifecycle. Figure 2.5 illustrates how the tools combine with the waterfall model. Task allocation effectively defines the boundary of the computer system and, thereby, serves as an input to requirements analysis. The software development process, however, is relatively untouched by the various tools. It is not until either a prototype or operational system is ready that the usability tool is applied, effectively serving as a filter to prevent unusable systems being implemented as part of an organizational change or job redesign.

The lack of interaction between the software development and organizational change processes is limiting and potentially costly. It is quite plausible, for example, for software to be developed to satisfy the boundaries specified by the task allocation tool and still be rejected by the usability evaluation. Indeed, without a considerable number of user-related design parameters as inputs to a traditional software development lifecycle, it is difficult to see how the satisfaction of usability criteria could be guaranteed. The use of an evolutionary prototyping methodology would resolve some of these difficulties by allowing the usability evaluation stages to provide feedback to the job design and task allocation stages (as illustrated by the dashed line on figure 2.5). This feedback could then be combined with a revised task allocation to provide guidance for the further development of

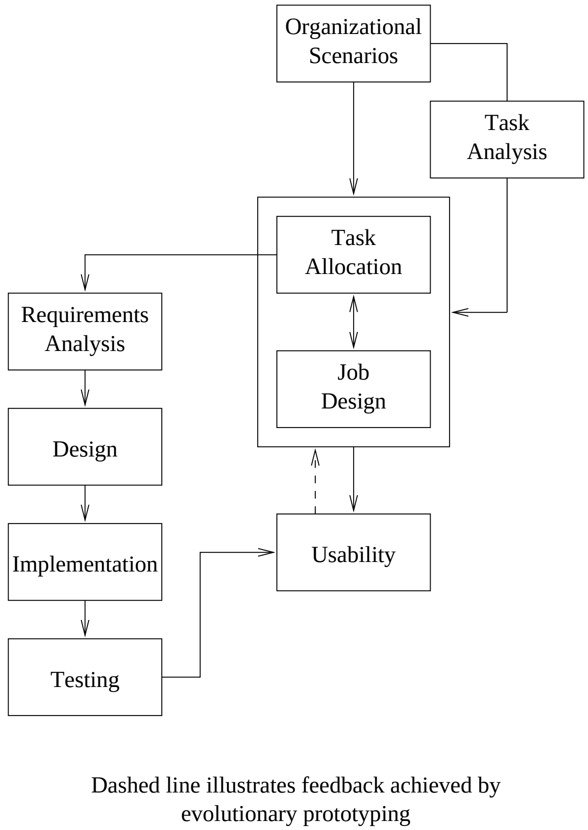


Figure 2.5: How the Tools of Clegg, et al. (1996) Relate to the Software Development Lifecycle

the prototype.

Like many user-centred techniques, the extent to which the tools proposed by Clegg et al. (1996) raise awareness of user needs during software development is limited. Even in the case of evolutionary prototyping, only a route for communicating user information is established. To ensure user involvement, the tools would have to be modified and incorporated into a framework that would stimulate interaction between software development and job redesign methodologies sufficient to ensure their alignment. A lifecycle model that may contribute to increasing the effectiveness of the above approach is developed in chapter 6.

2.4.2 Usability Engineering

Whereas the tools described above place most emphasis upon the redesign of user activities, the Usability Engineering approach put forward by Nielsen (1993) proposes a usability lifecycle (see table 2.5) focused around the development of the interface component of the software. The lifecycle places considerable emphasis on narrowing down the possible 'design space' as easily as possible so that the minimum of effort is required to identify a usable interface.

1.	Know the user
	a. Individual user characteristics
	b. The user's current and desired tasks
	c. Functional analysis
	d. The evolution of the user and the job
2.	Competitive analysis
3.	Setting usability goals
	a. Financial impact analysis
4.	Parallel design
5.	Participatory design
6.	Coordinated design of the total interface
7.	Apply guidelines and heuristic analysis
8.	Prototyping
9.	Empirical testing
10.	Iterative design
	a. Capture design rationale
11.	Collect feedback from field use

Table 2.5: The Usability Engineering Lifecycle (From Nielsen, 1993)

As with the tools proposed by Clegg et al. (1996), the usability lifecycle benefits most from the use of iterative or evolutionary prototyping. The parallel design stage encourages the development of numerous unrefined prototypes in the early stages of design, each produced by a separate design team. The various prototypes are then compared and a single prototype constructed using the best ideas from each of the parallel designs. The participatory design stage that follows exposes the interfaces to users who provide feedback for the next stage of interface prototyping. Although Nielsen suggests that vertical prototyping can be used at this stage to evaluate the combination of interface and functionality, he does not discuss how the evaluation results are fed back into the development of the system's functionality. Thus, many functionality issues that may have serious impacts upon usability have no explicit way of being resolved.

Whilst the usability engineering lifecycle is clearly software oriented, it does not describe the development of the entire software system. In fact, the entire lifecycle proposed by Nielsen is based on the assumption that the interface can be treated as distinct from the underlying functionality of the system. This assumption is the most significant factor limiting Nielsen's approach. Response time, for example, is a significant usability factor, yet this is usually determined by the efficiency of underlying algorithms and software architecture rather than by the visible components of the interface. Similarly, the user does not distinguish data validation from the interface through which they interact, yet the range of acceptable inputs for a given interaction is defined by the system function (e.g. when inputting a customer type, acceptable inputs would be determined by the customer types encoded in a lookup table). The boundaries of the usability lifecycle cannot be extended to accommodate such functionality, however, because it does not provide adequate support for the whole software engineering process. Thus, the only alternative is to integrate the usability lifecycle with a software development lifecycle.

Given that Nielsen's approach centres upon specific aspects of the software, rather than upon non-software issues (as in the case of Clegg et al. 1996), integration cannot be achieved by simply associating the usability techniques with corresponding phases of the software development lifecycle. It is not sufficient, for example, for the various design steps listed in figure 2.5 simply to co-occur with the architecture and detailed design stages for the whole software system (if only because the 'interface' components would be designed twice). Instead, it is necessary to integrate Nielson's approach and a systems development approach at the methodological level. Such integration would require the substantial extension and modification of the usability engineering approach described above.

2.4.3 Usability Evaluation

When considering how HCI metrics contribute to improved software engineering, two issues must be considered: (a) what is measured; and (b) what impacts these measures have on the software development process. The evaluation of software and systems from an HCI perspective has primarily been studied under the guise of usability. Whilst the value of this term has been challenged (Baber 1993), it is widely used to refer to the gamut of metrics developed by the HCI community. A widely cited categorization of usability metrics (Shackel 1986) identifies the following four measurement aspects:

- **Learnability** - The time taken to learn to use an interface to a specified level of competence with training and user support available;
- **Effectiveness** - The extent to which user performance meets or exceeds a specified level for a given proportion of users;
- **Attitude** - Users' perceptions of an interface's capability for (a) reducing adverse effectives on users and (b) encouraging user interaction; and
- **Flexibility** - The degree to which end-users can adapt their interaction with the computer in response to task variations.

Measures developed in each of the above categories can be either behavioural or subjective. Behavioural measures are quantifiable behaviours of the user, such as task completion times and the number of accesses to on-line help. Subjective measures are opinions of end-users described in terms of dimensions defined by the usability team (which may be the systems development team). The ratings are often comparative (e.g. the user is offered a choice between two menu layouts on a prototype) and may be holistic (e.g. a ranking of wordprocessor packages in order of personal preference). Measures are combined in various ways to yield usability metrics. The methods for acquiring and combining measures are diverse but are mostly task-centred and range from interviews and questionnaires for subjective assessment to observation and (electronic) monitoring of user actions for objective assessment (Lindgaard 1994, Sutcliffe 1995)

An early model for usability evaluation, based upon quantitative studies in applied psychology, is the GOMS model (Card, Moran and Newell 1983). GOMS models the user's cognitive structure in terms of Goals, Operators, Methods and Selection rules. Goals are hierarchical in form, with the primary goal representing a task to be performed by the user, such as editing a document using a wordprocessor. The lower-level goals represent perceptual, motor and low-level cognitive acts. GOMS can be applied to a specific application to estimate task times. In doing so, different software packages can be evaluated in terms of the user's task efficiency. Although the model can cope with quite complex behaviours, it is only accurate when used to model skilled behaviour, rather than activities involving complex reasoning. Consequently, its value for improving the usability of software is limited. Efficiency gains and error reductions might be achieved, but issues such as ease of learning and the ease with which the user's task goals can be translated into goals for human-computer interaction cannot be addressed.

MUSiC (Macleod, Bowden and Bevan 1996) is a well-known and commercially successful usability evaluation method, which was developed as part of the European ESPRIT programme. The method was developed to address three deficiencies of earlier approaches to usability evaluation. First, systems development is constrained in terms of time and resources. MUSiC was developed in a tool-based manner and constructed so that tools could be selected as appropriate to the complexity and resource constraints of a project. Thus, rigour of analysis can be traded-off against cost and resource demands. Second, some methods assume well-defined work activities, which can be decomposed into tasks and then optimised. The GOMS and keyboard level model (Card et al. 1983) are paradigm cases of such methods in HCI. Efficiency-oriented approaches of this sort are not appropriate where - as is often the case with computerised systems - novel working practices are to be introduced. Such methods are also limited in terms of the evaluation of tasks that involve considerable cognitive ability and user discretion, even if the tasks are highly computerised. Third, the MUSiC developers recognised that usability is a key quality attribute and, consequently, can best be improved through the institutionalization of good usability practices and the processual application of usability tools and methods.

The MUSiC method consists of seven stages, as shown in figure 2.6. As with many HCI methods, MUSiC is based around an incremental prototyping approach to systems development, with steps four to seven being iterated as often as is necessary to achieve a usable system (within time and resource constraints). Low fidelity prototypes are used initially, allowing serious problems to be identified and broad changes to be made. Increasingly refined prototypes are used in each iteration to permit fine-tuning of the system's interface.

Usability evaluations conducted according to MUSiC take place in a usability laboratory or, if this is not possible, in the end-user's workplace, using portable usability equipment. For this reason, a high level of usability analysis skill is required and system developers must be trained or usability consultants employed. The measures used are sensitive to user expertise and it is, therefore, essential that the user sample is selected with great care. Depending upon the chosen scale of the evaluation exercise, some or all of the following products may be used in the assessment (Macleod et al. 1996):

- **Usability Context Analysis Guide** - A guide to support the specification of usability requirements and to select the tasks and contexts for evaluation.
- **Performance Measurement Handbook** - A reference source to support trained users, which supports the design and conduct of evaluations.
- **The DRUM Toolkit** - Video analysis software to support the study of observed human-computer interactions.
- **The SUMI questionnaire** - A questionnaire for eliciting users' perceptions of software quality.

The measures collected may include:

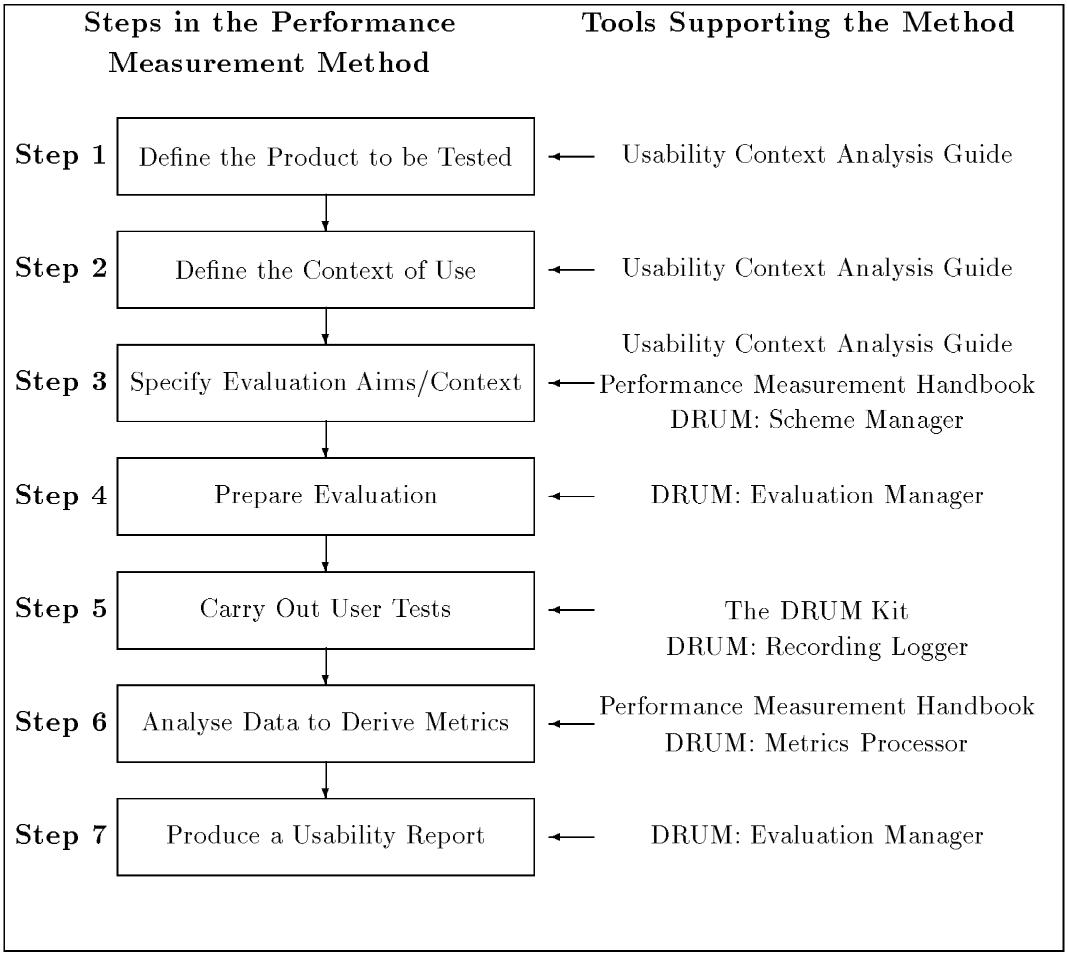


Figure 2.6: The MUSiC Usability Evaluation Methodology (From Macleod, Bowden and Bevan 1996)

- **User Effectiveness** - Calculated according to the proportion of task goals attempted by the users in the sample and the degree to which the task goals were achieved;

- **User Efficiency** - User Effectiveness divided by time;

- **Productive Time** - Total Time minus unproductive time (i.e. time spent using help and documentation, searching the interface for appropriate actions, undoing erroneous actions and user actions that do not result in a system response); and

- **Perceived Usability** - Subjective measures from a user satisfaction questionnaire.

The MUSiC methodology has had a reasonable level of success in terms of commercial uptake (particularly through its harmonisation with ISO/TC 159/SC4/WG5 and its contribution to ISO JTC1/SC7 (ESPRIT 1993)) and has made a significant contribution to HCI practice through the improved facilities for video analysis and eliciting feedback. The method requires expertise in human factors and is typically applied by human factors experts. The results are fed back into systems development via a usability report.

Whilst MUSiC can be beneficial in terms of improving human-computer dialogue, the laboratory based approach limits the extent to which the analysis can ensure the system is suitable for working practices in a specific organizational context. Whilst it is reasonable to claim that an unusable system will adversely affect task performance, the implicit assumption that user acceptance of a software product improves task performance does not hold. Thus, whilst MUSiC metrics may indicate efficiency of operation and user feedback under laboratory conditions is positive, such a system is not necessarily well-matched with working practices where users must interact, perform other tasks and handle interruptions.

2.4.4 Summary

Having reviewed two specific examples of interaction between HCI and software engineering and considered a selection of usability metrics, some general points can be made regarding HCI's contribution to improving information systems and software development. Firstly, as illustrated by usability engineering, HCI has implicitly assumed a distinction between the functionality of a system and its interface. This assumption is oversimplified, but cannot be overcome without a single theoretical basis for integrating HCI and software engineering analyses. Similarly, cognitive psychology (e.g. GOMS), work psychology (e.g. task analysis) and social (e.g. scenario analysis) analyses do not integrate to provide a clear picture of how design can proceed. Recommendations may be made regarding interfaces for the efficient performance of tasks by individuals, but these cannot be integrated to provide a rigorous analysis of how presentation affects the interpretation of information and how human-computer interaction affects the social context in which information is understood. Secondly, as with software engineering, there is a need for better recognition of the social dynamics of software development activity and how this is affected by the involvement of actors, such as human factors specialists, who have domain specific knowledge.

2.5 Information Systems

The Information Systems community has engaged in relatively little consideration of the information systems and software development lifecycles *per se*, addressing lifecycle issues in the context of information systems development methodologies. Many structured methods for systems development, such as SSADM, assume a conventional waterfall model or some variant. Such methodologies are not considered here because they have similar strengths and weaknesses to software engineering approaches based on the same model. The focus is upon those methodologies that have aimed to facilitate a user-centred approach to development and whether their implicit models of the IS development lifecycle are substantially different to those already discussed. Having considered the IS lifecycle, approaches to IS investment evaluation are then reviewed. As with IS development, the range of methods is considerable and only a representative selection is considered.

The field of Information Systems emerged from concerns that organizational and human factors need to be addressed when designing technical systems. Thus, the history of the discipline itself represents a continued development in treatments of user participation. One of the first information systems texts, developed on behalf of the National Computing Centre, was on systems analysis (Daniels and Yeates 1971). The approach to analysis assumed a waterfall type model consisting of analysis, design, implementation and testing. Owing to the severe constraints of ICTs at that time, the analysis process addressed many technical design considerations. Analysis was effectively regarded as making trade-offs between user needs and the need for efficient data structures and processing algorithms. In terms of user involvement, the approach is consistent with the waterfall model of software engineering, with the analyst using interviews, questionnaires and observations as the basis for systems design.

Over the 1970s and 1980s, information and communication technologies have spread throughout the organization, increasingly being used to *support* users' tasks, rather than their use *being* the main tasks of specific groups of workers. Consequently, this twenty year period saw an increasing impact of ICTs on job design and work organization. In response to these changes, the IS community began to consider technological developments in terms of their organizational change implications (see, for example, Damodaran et al. 1980, Eason 1988 and Olle et al. 1988). User participation in the development process was recognised as important to user acceptance of software systems and to the effectiveness of computer-supported work activities.

2.5.1 Effective Technical and Human Implementation of Computer Systems (ETHICS)

One of the earliest methodologies for addressing organizational needs when developing computer systems is Effective Technical and Human Implementation of Computer Systems (ETHICS) (Mumford and Weir 1979, Mumford 1995). The primary motivation of ETHICS was to discourage the trend for job design to be regarded as an afterthought that had to fit around a technical implementation of an already fixed design. The first step in the development of ETHICS was the realization that the design stage of systems development must consider both technical *and* work design. In considering whether ETHICS can, in principle, achieve its aim of developing satisfying jobs around the use of computer systems, the following key premise must be considered:

> "At each stage [of design and implementation], technical *and* human needs are taken into account, so that the system is designed specifically to meet both technical and human objectives at one and the same time." Mumford and Weir (1979, p.3)

The basic structure of ETHICS is simple, consisting of four stages: diagnosis, design, monitoring and evaluation. The activities performed at each stage are balanced: (a) to provide equal weighting to human and technical needs; and (b) to result in improved job satisfaction.

The diagnosis stage consists of three activities. First, user satisfaction questionnaires are used to collect data on the pre-change system. Statistical analysis is performed upon the resulting data to identify any source of dissatisfaction. Where any are found, observations attempt to establish the cause of dissatisfaction. Finally, employee groups discuss the identified problems in order to establish *why* the present situation is unsatisfactory. At the design stage, the results of the diagnosis are used by the design team (which includes some employees) to formulate human objectives. Technical objectives are then established and social and technical solutions are identified. Having identified alternatives, an attempt is made to establish feasible systems designs by combining social and technical solutions. These solutions are ranked in terms of preference and the best alternative selected for implementation.

ETHICS does not address development and implementation in detail, but describes a monitoring procedure whereby deviations from the human objectives can be identified. Although no explicit support is given for doing so, the project team are expected to intervene if such deviations are identified. The final stage of the ETHICS method involves evaluating the change process using the tools from the diagnosis stage. This evaluation enables the team to learn from this experience as developers of socio-technical systems.

As noted above, the main premise of ETHICS is that the joint consideration of human and technical objectives will lead to systems that satisfy users and provide a satisfactory technical solution. Thus, the parallel structure during the early stages of diagnosis and objective setting is significant because it indicates that human and technical issues are treated separately until well into the design stage. This early division encourages a conceptual framework for analysing and designing systems that, whilst a significant improvement over the 'technical imperative,' is, itself, limiting in terms of user-centred design.

Consider the nature of, and distinction between, human and technical objectives. Objectives are formulated by stakeholders in the system. Thus, one must ask which stakeholders contribute to formulating which objectives. Whilst it is clear that workers who will be end-users of the computer system are the primary specifiers of human objectives, the origins of technical objectives are less obvious and not discussed in detail by the authors. Mumford and Weir (1979) assume that the technical objectives are set by technical experts within the organization. Whilst this may, on the whole, reflect systems development practice (particularly in the 1970s), it ignores the broader objectives of the managers who authorised the technical change. If management objectives replace technical objectives in ETHICS, however, then diagnosis becomes a process of reconciling conflicting sets of human objectives; a process for which ETHICS provides no explicit support. If one continues to assume that certain management objectives are isolated and are faithfully reflected in the objectives of the technical experts, then one limits the use of ETHICS to technical systems that provide only efficiency gains.

The divide between the social and technical systems is effectively an assumption that management follows scientific management principles when developing its technical objectives and these objectives are essentially at odds with workers' interests. Whilst this may have been a plausible assumption regarding computer systems in the 1960s and 1970s, the assumption is now less valid because ICTs have more significant impacts upon the nature of work than achieving productivity gains (Vincent 1990). Increasingly, ICTs are used within organizations to interconnect individuals and, thereby, accompany organizational restructuring, rather than being confined to job design for primary users. Where workers interact *through* ICTs rather than *with* them, the initial division made by ETHICS is constraining. A broader analysis of the management philosophy and nature of working practices is required.

2.5.2 Soft Systems Methodology (SSM)

Whereas ETHICS is specifically targetted at problems of computerisation, Soft Systems Methodology (SSM) is a general approach to analysing systems and guiding intervention. The method's breadth in scope makes it applicable to a wide range of problem situations, but less specific in its guidance for designing and developing systems and/or their technical components. Although centred around the same seven stages, explications of SSM have developed considerably since the early 1980s. Initially providing techniques for analysis and modelling (Checkland 1981), criteria for evaluating the quality of SSM models and the quality of transformations in the activity system have been introduced (Checkland and Scholes 1990). Furthermore, increased attention is now paid to cultural analysis, with techniques for modelling the social system (in terms of roles, norms and values) and the political system being provided (Checkland and Scholes 1990, Checkland and Holwell 1998). SSM has had a significant influence upon IS research. Perhaps owing to its lack of prescription on the development of technological artefacts (see, for example, Savage and Mingers 1996) and the cross-cultural difficulties experienced in its use (Gregory and Lau 1999), SSM has not, however, had a comparable impact on IS practice. To explore its potential contribution in this respect, the lifecycle model and methods for user participation that SSM can support are now considered.

SSM regards organizational forms as systems, which are described as processes made up of a number of transformations. This processual view, as noted by Checkland and Scholes (1990, p.312), is more compatible with a prototyping approach to systems development than a project-oriented model such as the waterfall lifecycle. In addition to being processual in nature, prototyping is specifically aimed at providing a learning mechanism for the development of technological artefacts. Checkland and Scholes (1990) claim that SSM can be used to provide a mechanism for modelling information flows based upon a widely agreed model of the organization's activities. By continually reflecting upon the model, the organization is supposed to be able to review its information strategy. Whilst this may be the case, it must be noted that the effectiveness of such an approach to information strategy is dependent upon who participates in the development of the model. In the case of strategy development, it is likely that senior managers will dominate the process of analysis and this will have a substantial effect upon the nature of the model of the human activity system. No explicit mechanism for managing user participation in such an application of SSM is supported.

In terms of developing the ICT systems embedded in human activity systems, rather than managing the use of ICT resources, the contribution of SSM has not been made explicit. This is primarily because, despite the more normative extensions for evaluating methodology use, its contribution is solely at the analysis stage. This limitation in scope is a significant weakness for Soft Systems Methodology as an approach to systems development. Although the analysis may lead to an indication of information needs for the various working practices, there is no direct means of translating the analysis into a specification for the engineering of the ICT components. As argued above, the use of evolutionary prototyping, which SSM claims to be compatible with, only ensures that the system will satisfy certain aspects of the users' requirements. Other aspects of the system, such as detailed processing, still need to be analyzed and defined with sufficient clarity for them to be engineered. Such a process is critical to the accuracy and reliability of information systems, and reasonable satisfaction of both is necessary for user acceptance. Thus, whilst progress has been made in terms of using SSM to support information systems analysis, a much more comprehensive methodology is required for improving IS development practice.

2.5.3 Multiview

Of the various methodologies that address human and organizational issues, Multiview (Avison and Wood-Harper 1990) is one of the more comprehensive, consisting of the following five stages:

1. Analysis of human activity;

2. Analysis of information;

3. Analysis and design of socio-technical aspects;

4. Design of the human-computer interface; and

5. Design of technical aspects.

For this reason, the methodology is considered here in some detail.

Multiview attempts to combine elements of conventional software development and the soft systems and socio-technical approaches already discussed. It does so by proposing a contingent approach to systems development, whereby the application of Multiview techniques depends upon the characteristics of the problem situation.

Multiview does not follow a conventional lifecycle, but models systems development as an iterative complex of logical and social activities. The methodology does not provide an integrated perspective to development but, as its name suggests, draws upon several perspectives: Checkland's soft systems approach; Mumford's socio-technical approach; and conventional systems development. Consequently, in considering Multiview's success in addressing user and organizational issues in developing effective systems, the main focus of attention is whether the diverse analyses lead to a coherent and satisfactory solution. For this reason, the interconnections between Multiview's development stages are of particular interest.

In applying Multiview one begins with an analysis of the human activity system using SSM's rich picture technique. The purpose of this stage is to gain agreement between the analyst, client and problem owner regarding the nature of the present (unsatisfactory) situation. The main stakeholders, activities, issues, concerns and potential conflicts are emphasised. Having established a high-level description of the problem situation, the boundaries of the system are then defined in terms of a *root definition*. This definition explicitly states the world view, or *weltanschauung*, from which the analysis is being conducted. By doing so, the analysis stage can stimulate the consideration of different perspectives on the problem situation so that the most appropriate model for describing the problem is chosen. At this stage, general agreement will have been reached about the *nature* of the problem. No detailed and measurable objectives are defined, however, and no specific process for doing so is provided by Multiview. This is a weakness of the method (and the methods it draws upon) because success and failure are not clearly defined for the software system beyond conformance with the documentation. In this respect, Multiview is document driven, which is convenient for dealing with contractual issues, but not always satisfactory in terms of ensuring user acceptance.

The next stage of analysis is to develop an abstract conceptual model of the activity system within the bounds of the root definition. This model is constructed by the analyst, customer and problem owners in a normative manner, stating how the system *ought* to work. This model is compared with the current performance of the system to formulate a plan for improvement. As noted by the authors of Multiview, this initial stage of analysis is more accessible to users than the equivalent stages of conventional systems analysis.

Having modelled the overall system in general terms, attention is focussed upon the information flows that support and interconnect its component activities. This stage uses conventional modelling techniques to perform a functional decomposition of the problem leading to the documentation of information requirements using entity-relationship and data flow diagrams. Given the significant amount of effort put into the modelling of the human activity system, some benefit in terms of the efficiency of information analysis or the quality of its outcomes would be expected at this stage. The most likely source of benefit in this respect is the choice of an effective viewpoint for modelling the problem situation. Particularly if advised by an IS professional, the model of the activity system ought to be a convenient perspective from which to describe information flows. A further advantage is that the functions to be used in the information model are abstracted from the conceptual model, thereby making the information modelling process more user-centred than the conventional approach to systems analysis. It should be noted, however, that the end-result is still in the format of a conventional information model. Consequently, the mechanisms for ensuring its use remains user-centred in the subsequent design stages must be considered.

The first stage of design is the use of Mumford's socio-technical approach for, amongst other things, partitioning tasks between the user and the computer system. Consequently, Multiview is exposed at the design stage to the same limitations as ETHICS. A further complicating factor for Multiview is ensuring the compatibility of the socio-technical analysis with the already agreed model of the human activity system developed using SSM techniques. Such reconciliation could be problematic because the socio-technical approach is based upon a particular view of industrial democracy, which may be at odds with assumptions implicit in the *weltanschauung* of the rich picture and the conceptual model.

Having established the technical system's boundaries and requirements, existing approaches to interface and software design and development are applied. Such an approach is desirable on the grounds of the availability of technical skills. Perhaps the most significant deficiency of Multiview, however, is that it provides no explicit framework for evaluating the likely social and user impacts before or during development. General objectives for the changes to the human activity system are formulated during analysis, but these need to be translated into technical objectives for the software in order to guide the development process. Without such a mechanism, there is a risk that the technical implementation will become distanced from the systems development process in a manner similar to that illustrated by Clegg's organizational and psychological tools (see section 2.4.1).

The recent development of Multiview 2 (Avison et al. 1998) aims to address some of the limitations identified above by paying more attention to software development and implementation. Multiview 2 replaces the original five stages of Multiview 1 with an 'interpretive scheme,' which includes organizational analysis, information systems modelling, socio-technical analysis and software development. Clearly, this revision does not address the problems stemming from the socio-technical approach or the possible conflicts between the socio-technical and SSM perspectives. It does, however, reduce the conceptual distance between social objectives and technical design by placing more emphasis on the implementation process. Furthermore, the contingent nature of the methodology is much more intuitive than that presented in Multiview 1. The emphasis upon each part of the interpretive scheme is different for the use of a software package, for example, in comparison with the development of bespoke software. Until a more detailed explication of Multiview 2 is available, however, a more detailed analysis of its success in addressing the weaknesses of the original methodology cannot be performed.

2.5.4 Information Systems Investment Evaluation

The analysis of Multivew illustrates the importance of epistemological issues, particularly different 'world views,' on the conceptualization and resolution of information systems problems. Particularly when considering choice between alternative systems developments and other evaluation issues, it has usually been assumed that increasing profits is the aim of any investment. Whilst this might be the officially stated aim and, perhaps, the aim of shareholders and (indirectly) senior executives, the world views and goals of stakeholder groups are usually more diverse (Moorhead and Griffin 1992). Consequently, one must begin studying information systems investment evaluation from the premise that, although stakeholders want the benefits of change to exceed the costs, there is no single perception of what costs and benefits are. To gain an understanding of the likely impacts and perceived impacts of IS-related change, one must, therefore, consider both the causal relations between changed system components and the causal attributions made by stakeholders between perceived costs and perceived benefits.

The delineation of causal relationships between recorded costs and recorded benefits, however, presents numerous practical difficulties, the most significant being that: (a) some factors are difficult to quantify (Couger 1987, Parker et al. 1988, Strassmann 1990, Farbey, Land and Targett 1993); (b) accounting for risk and uncertainty leads to a range of possible evaluation outcomes (Powell 1992, Dos Santos 1994); and (c) the audience and problem characteristics must be accommodated by the evaluation process (Hirschheim and Smithson 1987, Symons 1991, Avgerou 1995). The most common approach to IS evaluation is to apply standard financial appraisal techniques. These methods have been criticised, however, for failing to address any of the three practical difficulties adequately (Strassmann 1990). Consequently, several IS specific quantitative methodologies have been proffered (Parker et al. 1988, Strassmann 1990, Kaplan and Norton 1992) but their take-up has been low (Willcocks and Lester 1994). More recently, there has been a significant shift away from quantitative approaches to IS evaluation (Symons 1990, Symons 1991, Symons 1994, Serafeimidis and Smithson 1994, Avgerou 1995, Serafeimidis and Smithson 1996).

The foremost criticism of capital investment appraisal techniques is that strict adherence to accounting principles precludes the recording of a significant proportion of an investment's costs and benefits: any factor that cannot be defined, observed and measured (i.e. intangibles) should be excluded from the analysis (American Accounting Association 1966, Ernst and Young 1992). To acknowledge these criticisms, however, is not to declare such methods unsuitable for IS evaluation. It must simply be recognised that, due to the restrictions required to ensure a valid accounting analysis, the results provide only a subset of the information required to make a balanced decision concerning any investment. It is often the case, however, that the satisfaction of all accounting criteria is not possible and internal consistency must be sacrificed in favour of external relevance or *vice versa*. In other words, information quality involves trade-offs in the light of the purpose(s) for which and the user(s) for whom the analysis is intended:

> "For internal purposes, subjective estimates, especially forecasts, may be used because of their high degree of relevance even though they possess a very low degree of verifiability." American Accounting Association (1966)

Despite, or perhaps owing to, this element of subjectivity, accounting guidelines serve as a useful protocol for the conveyance of information to support an IS evaluation and any subsequent decisions. Provided that the decision maker is aware of the principles of accounting (and any necessary deviations), conflicts with the decision framework and subsequent limitations can be recognised. Nevertheless, observations that such mechanisms are consistently poor in an area dominated by hard-to-quantify benefits (Strassmann 1990) suggests either that such an awareness does not exist or that the use, in combination, of accounting and non-accounting information for decision-making is unsuccessful. This poor performance has been seen by some, however, as a failure of capital investment appraisal techniques to find the 'true' value of information technology (Parker et al. 1988). Consequently, attempts have been made to develop techniques that combine accounting and non-accounting information to provide a more comprehensive analysis, to accommodate the views of multiple stakeholders and, in some cases, to explicitly quantify the risks and uncertainties associated with the various evaluation criteria. A representative selection of quantitative techniques is discussed below to demonstrate the fundamental problems of attempting to find the 'true' value of an investment in this manner.

2.5.5 Cost-Benefit Analysis

Cost-benefit analysis aims to extend capital investment appraisal to address two of the three deficiencies identified earlier (Farbey et al. 1993):

1. providing a mechanism for the investor/decision maker to consider the costs incurred and benefits realised by other stakeholders; and
2. enabling considerations of factors that have no immediate monetary value.

It is evident, however, from a comparison of cost-benefit analysis and the accounting techniques from which it was derived, that cost-benefit analysis is not so much an extension but a relaxation of the rules that maintain the integrity of the accounting analysis. Furthermore, because the modification of these rules is not accompanied by any change in the appraisal process, neither deficiency is addressed satisfactorily.

Considering point 1 above, a capital investment appraisal is produced for a specific audience; this is made quite explicit in the accounting literature (American Accounting

Association 1966, Ernst and Young 1992). Thus, if the finances of several groups will be affected by an investment in different ways, a comparison of accounting analyses produced for each of them would reflect this. Capital investment appraisal rightly avoids any attempt to reconcile divergent stakeholder views because this kind of political activity is beyond the scope of accounting methodology. Any overall view of the investment should, therefore, be the product of negotiations between the various interested parties (Self 1970).

The second point concerns the acceptability of notional values as part of an evaluation (Self 1970). In the case of cost-benefit analysis, the evaluator must derive monetary value for all criteria to ensure that the analysis is complete. Where the criterion cannot be subject to exchange under market conditions, monetary values are meaningless (Self 1970) and the numbers necessarily arbitrary. Even where quantities exist from which a monetary value can be approximately derived, the estimation will be a matter of judgement, the validity of which will always be open to question. It is also necessary that the criteria included in the cost-benefit analysis are independent; a demand which is impossible to satisfy where the definition of the criteria are themselves matters for debate. It is unreasonable to expect an evaluator to place a monetary value on customer satisfaction, for example, that is independent of a resulting increase in sales, yet this is a necessary condition for the factors to be included in a rigorous cost-benefit analysis. Furthermore, trying to provide an analysis that suits all stakeholders, rather than addressing an identifiable group, makes the identification of an acceptable interpretation of what the criteria mean to the audience an intractable problem.

The combination of these changes to standard accounting techniques, rather than producing a clearer view of the investment's overall value, are likely to lead to disputes over the precise amounts attributed to particular criteria, thereby detracting from the more fundamental issue of what the criteria represent. A more sensible approach would be to expose the areas of conflict via an open debate to arrive at an understanding of why the various groups differ. This demands that only readily quantifiable factors are included in capital invesetment appraisals that are produced with specific audiences in mind.

2.5.6 Return on Management

Whereas cost-benefit analysis and Information Economics (see Parker et al. (1988) for a detailed explication and Hemingway (1997a) for an discussion of the method's key deficiencies) are concerned with extending investment appraisal to make it more complete, Strassmann (1990) criticises the basic premises of accounting methodology, arguing that the focus upon investment in capital, at least with regard to information technology, is fundamentally misplaced. To support his claim, Strassmann conducts an extensive analysis of the relationships between various measures of IT performance and returns to shareholders. The investigation reveals that no direct relationship between these measures and business performance, so defined, can be demonstrated. Strassmann subsequently contends that management, not capital, is responsible for business performance and develops the Return on Management calculation on this basis. Intuitively, the relationship between management and business value is appealing and it is not entirely discounted here. The Return on Management calculation (Figure 2.7), however, is of questionable validity.

$$\textbf{Return on Management (ROM)} = \frac{\textbf{Management Value Added}}{\textbf{Management Costs}}$$

Figure 2.7: The Return on Management Calculation (From Strassmann 1990)

Based on the assertion that management information systems affect only management performance, the ROM calculation is purported to isolate the contribution of management to the organization and the costs that management incur, indirectly demonstrating the effects of MIS. In order to claim that MIS causes increases in management value, however, it is also necessary to assert either (a) that ROM is affected only by MIS investment or (b) that when all other factors affecting ROM are eliminated there is a correlation between MIS investment and the remaining components of ROM. Although the calculation of management value eliminates all costs, it does not exclude all benefits that are derived from non-management expenditure, thereby failing to satisfy either assertion (a) or (b) above. The effects of investment in management, in the form of IT or otherwise, are not, therefore, distinguished from the effects of other types of organizational change. Furthermore, management value added (and implicitly management) is defined as the value remaining when all costs have been accounted for, implying that all benefits in an organization are due to management, thereby suggesting that **all** discretionary activity within an organization is regarded as management for the purposes of ROM analysis. An examination of Management Value Added and Management Costs reveals that Return on Management is essentially a profitability calculation with non-management costs excluded from the denominator and is, therefore, only meaningful for business unit and higher level analyses, if at all. Thus, it might be suggested that ROM could be used by the highest levels of management to draw some form of comparison, for example, between different lines of business and how they utilise their resources.

As already noted, because the benefits resulting from non-management expenditure are not excluded from the analysis, Return on Management is sensitive to changes in all categories of cost. Strassmann justifies this by arguing that good managers will be able to achieve, for example, higher returns on investment and negotiate better contracts with suppliers (Strassmann 1990, p.88). Although this may be true, the performance of management in no way accounts for all changes in the prices of inputs; there are, unsurprisingly, influences in the environment that affect the component factors of the ROM calculation. Furthermore, whereas profitability calculations can only be improved by increasing net profits or reducing costs, Return on Management can be improved simply by shifting costs from management to an operational cost category and is, therefore, much easier to manipulate. The questionable nature of some definitions, of which 'management' is an example, further exposes the method to numbers games. Thus, although Strassmann offers many useful insights regarding investments in information technology, the actual method conveys little useful information to the evaluator.

2.5.7 Decision Analysis Techniques

In contrast with accounting techniques, methods for decision analysis do not rely upon the use of monetary values. Moreover, decision theorists have developed rigorous methods for eliciting preferences and subjective probability estimates (Alt 1971, Wright 1984, French 1986). A number of different approaches to decision analysis have emerged and distinctions have been made between decision types that require different treatments. In terms of information systems evaluation, two types of decision are of interest: one-off decisions made by individuals; and decisions made by groups. Whilst specific applications of some of these techniques to information systems have been reported (Parker et al. 1988), the focus here is upon the relevance of decision theory in general.

When an individual encounters repeated instances of similar problems, he or she develops standard responses that are known to be successful. For such classes of problems, the problem situation is perceived holistically through its association with long-term memories of similar situations (Larichev 1984). In contrast, decisions such as IS investments are sufficiently different on each occasion to be regarded as unique. Such problems must be explicitly analysed by the decision maker in order to identify a course of action. Decision theory aims to support explicit analysis of this kind by providing mechanisms for eliciting values for decision criteria and decision rules for determining appropriate choices.

In considering the contribution of decision analysis to IS evaluation, it should be observed that IS investments are characterised by very high levels of uncertainty. The understanding of the problem situation is sometimes so poor that requirements cannot be stated, even in quite broad terms. Consequently, a successful decision analysis must begin by supporting the decision maker in describing the problem situation in terms of its key attributes. It is this stage, however, that is most poorly addressed by the decision analysis literature. Indeed, as noted by Harrison (1995, p.9-10), it is typically assumed that users will be able to identify the key attributes without difficulty and the main problems are in formulating subjective valuations of each attribute in a manner consistent with the mathematical requirements of normative decision theory.

A further limitation of decision analysis is that the representation of the problem is often not consistent with the decision maker's intuitive perceptions of the problem (Hogarth 1986). Consequently, although the analysis may be informative, it may reduce the confidence of the decision maker that the decision he or she takes is correct. In the case of IS investments, this is critical because the championing of a chosen course of action has been shown to be critical to the success of IS development (Palvia and Chervany 1995). Given these basic limitations, decision analysis techniques for individuals are not considered further, although the empirical findings of decision analysts are considered in the review of psychological issues presented in chapter three.

Considering group decision analysis, a key factor is the negotiation of an agreed set of criteria for decision making and a mechanism for assigning values. This type of process more closely reflects *ex ante* IS investment evaluation, where different stakeholder groups weight criteria differently and have a vested interest in ensuring their weightings are reflected in the group's decision analysis. The social dynamics of the group clearly impact upon the decision outcome. Thus, a challenging question for decision theory is to determine the status of the group's analysis and what role it has in determining the overall decision outcome. Research into this aspect of IS evaluation has increased since the early 1990s, with processual analyses leading to significant insights (see, for example, Symons 1990, Symons 1994, Bryson and Currie 1995). Models and theories of social interaction suitable for analysing the evaluation process are considered in detail in chapter four.

2.5.8 Summary

The above analysis of information systems development and evaluation indicates several basic limitations of current approaches. It is clear from the analysis of information systems methodologies that a key requirement for effective systems development is the successful involvement of customers and users. Information systems has made a key contribution in this respect, using tools such as rich pictures, which are more inclusive than conventional systems analysis techniques. As illustrated by SSM, however, user participation in the analysis is ineffective if the analysis cannot be used to improve the quality of technical design. Multiview has made some progress in this respect by attempting to account for user-centred analyses at the design stage. The current limitations of Multiview and other methodologies, however, indicate that there are a number of conceptual difficulties to be faced in integrating multiple analyses of the problem situation.

Information systems evaluation during the early stages of the lifecycle is extremely limited in what it can achieve. Furthermore, the difficulties faced in improving on current methods are considerable. Standard investment appraisals continue to dominate practice, despite having severe drawbacks for evaluating ICT investments. This is, at least in part, a result of their familiarity and acceptance for use in other types of investment decisions. In attempting to extend accounting-based methods, arbitrary values have proven to be potentially misleading, particularly when they are expressed in terms of familiar financial units. The analyses conducted by Strassmann (1990) demonstrate that effective evaluation requires an identification of the causal factors that are changed by technologies. The limitations of his Return on Management calculation, however, illustrate how difficult it is to establish causal relations when analysing working practices. The decision theory literature provides a number of insights into the problems of IS evaluation. The findings are somewhat frustrating because the literature has documented many obstacles to effective evaluation but relatively little progress has been made in overcoming them.

2.6 Key Issues

The ultimate aim of any development involving information and communication technologies is to produce a high-quality technical system that is perceived to be beneficial to at least some of the actors in the social system in which it is embedded. By reconsidering several current perspectives on ICTs, it has been possible to identify several factors that are central to achieving this aim.

A central theme in the above analysis was the importance of considering the physical situation in which ICTs are developed and used. In terms of the wider social system, this suggests that the goal of information systems development should be to integrate ICT artefacts into an effective and satisfying activity system. In the context of firms, which develop most large information systems, this equates to the integration of ICTs into effective working practices.

As illustrated by information systems methodologies, the analysis of working practices is primarily concerned with the social dynamics. The above review has shown, however, that such analysis does not readily contribute to technical design. Furthermore, current methods for modelling social activities are not sufficiently consistent with the psychological models of HCI to provide a coherent design rationale that accounts for individual and social factors. A significant step forward in terms of improving the development of ICT-based information systems would be the development of an integrated basis for considering the psychological and social aspects of systems design.

The recognition that social and psychological analyses cannot be regarded as distinct is particularly evident in the consideration of information content versus information presentation. Current HCI techniques focus primarily upon general presentation characteristics, but are poorly informed in terms of how the style of presentation affects information use in a social context. Conversely, social analyses pay considerable attention to communication and meaning but do not relate these issues to the design of information artefacts.

Considering the design of ICT artefacts, similar conceptual divisions between physical and logical design and between interface and functionality have been shown to be problematic. The recent emphasis upon logical design considerations has led to the design of complex ICT systems as wholes. This approach results in a highly complex systems development process, which has typically had quite poor results in terms of systems quality. It was suggested above that closer attention to the physical characteristics of ICTs at the design stage may help to improve the quality of systems components and, thereby, raise the quality of ICT systems. The field of information systems has proposed a number of development methodologies that are more radical than software engineering approaches. These methodologies have made significant steps forward in addressing some user and organizational issues. In making these issues more visible, however, information systems researchers have revealed some difficult philosophical and conceptual issues that demand further attention. The epistemological issues relating to the notions of 'organizational' and user needs and objectives have been shown to be of particular significance when addressing information systems problems, including those of investment evaluation. For this reason, the issues relating to individuals are now considered in chapter 3 and the concepts of social organization and organizational form are examined in chapter 4.

Chapter 3

The Individual

3.1 Introduction

Information systems are constituted by human activity. Consequently, understanding IS phenomena requires a detailed explanation of human needs, characteristics, behaviours and actions. The scope of explanation is extremely wide, ranging from the need to exploit sensory physiology to design effective messages through to the impacts on social organization of using ICTs to mediate human experience. As illustrated in chapter 2, theory can best support value judgements related to IS design if it is integrated. Developing an integrated perspective is difficult, however, for several reasons. The degree of understanding at each level of physiological and psychological analysis varies considerably. Furthermore, some levels focus on physical characteristics, whilst others concentrate on mental functions, largely ignoring their implications for human behaviour or their plausibility with respect to the brain's physiology. The most conceptually challenging aspect of integrated theory development, however, is the need to reconcile objectively describable processes with the subjective experience implied by cognitive concepts from perception to reflexive action.

Given the potential value of an integrated theory and the evident limitations of basing systems design around various incompatible perspectives, this chapter selectively reviews the cognitive science and social cognition literatures to indicate the present state of understanding and considers some of the basic issues that must be addressed by an integrated theory of the individual. Particular attention is paid to identifying those aspects of the person that have received little attention in the broad 'architectural' models of cognition, which are among the more complete models of mental activity.

3.2 Sensory Stimulation

A key element of any organism's ability to adapt its behaviour in response to its environment is some means of becoming informed about the environment's present state. The diversity of sensory apparatus across species illustrates how evolutionary adaptation has resulted in sensory physiologies that are stimulated by information from the environment that is most relevant to the organism's survival. In the case of human sensory physiology, the range of information that can be gathered via the different sensory modalities is enormous. The three stages of objective sensory physiology prior to perception demonstrate, however, that the vast quantities of sensory stimuli are extensively processed and integrated before reaching the brain (Dudel 1986):

1. phenomena in the environment map onto sensory stimuli;
2. sensory stimuli excite the sensory nerves; and
3. the propagation of excitation from sensory nerves to the central nervous system and resulting integration of stimuli.

Early visual processing, for example, facilitates discrimination between four environmental factors that combine to determine the stimulation of visual receptors: (a) geometry; (b) surface reflectance; (c) scene illumination; and (d) viewpoint (Marr 1982, p.41). The integration of stimuli by the sensory system results in a range of functional stimuli, such as intensity changes in the visual field, that serve as inputs to the subjective sensory physiology that facilitates perception (Dudel 1986).

3.3 Perception

As mentioned earlier, perception is the transition from sensory processes that are objectively describable to those that imply subjective experience. For this reason, Dudel (1986) identifies consciousness as a condition for the occurrence of sense impressions and perception, making them challenging to explain. The natural scientific studies of perceptual processes in psychophysics has, however, provided some valuable insights into the early stages of cognition.

Although it is evident that not all experiences are discursive (Anderson 1990), the primary research approach in psychophysics involves presenting specific stimuli to subjects and recording their statements of what they experience (Dudel 1986). Sometimes, reliable correlations can be established between objective measurements of nerve activations and subjective reports. This has not led to the development of an integrated perspective, however, for primarily methodological reasons:

> "As far as the methodologic approach to subjective sensory physiology is concerned, we can in principle ignore all that we know about sense organs, receptors and centers in the brain - our interest is directed entirely toward the response of the human subject to a sensory stimulus." Dudel (1986, p.15)

The subjective nature of sense impressions is reflected in the measurement systems used by psychophysics. Although measured stimuli, such as particular wavelengths and intensities of light, are objectively measured, the objective measurement of intensity in terms of action potentials in the nerves, for example, is replaced by one of the following methods:

1. subjects estimate how much stronger a sensation is compared with a standard sensation unit, providing a ratio scale;

2. the comparison of stimuli intensities across modalities (e.g. the more intense a light, the harder the subject presses against a force meter); or

3. the measurement of intensity of sensation in steps of just noticeable difference.

Subjective scales are similarly constructed for the spatial, temporal and affective aspects of sensation. The ability to establish reliable correlations between objective sensory physiology and aspects of subjective experience, such as perceived intensity, provides a critical link that allows physiological properties to be exploited for information and ICT artefact design, as illustrated by the following examples:

- Knowledge of visual acuity, for example, enables estimates to be made of optimal test size on different types of displays (Sutcliffe 1995, p.17);

- Knowledge of how elements in an image are grouped together can be used to construct graphs to make information extraction easier (Kosslyn 1994); and

- Knowledge of the auditory system enables the generation of sounds (for example, ambulance sirens) so that their source is more easily located (Catchpole, McKeown and Withington 1998).

3.4 Memory

Both observations of adaptive behaviour and the developmental nature of conscious activity (Mandler 1992) indicate that memory of experiences is an important brain function. Unlike sensory processes, however, the physiology of human memory is not well understood. Consequently, models of memory have been developed from a functional perspective, often relying on assumptions drawn from computer science and artificial intelligence. The symbolic architecture proposed by Newell (Laird, Newell and Rosenbloom

1987, Newell, Rosenbloom and Laird 1989, Newell 1990, Newell 1992), for example, was originally proposed as an artificial intelligence technique before being proposed as a serious model of human memory and cognitive processing. More recent models have drawn upon physiological evidence concerning human memory to develop connectionist networks that are 'neurally inspired.' The symbolic and connectionist architectures are contrasted below and appraised in the light of available empirical evidence to illustrate the range of viewpoints on human memory.

3.4.1 Symbolic Architectures

A defining characteristic of cognitive science is its reliance upon computational modelling to provide predictions that can be compared with psychological and behavioural evidence. It is of little surprise that computers have been used to implement many of these models and to provide simulation data for comparison with observations of human performance. Somewhat less obvious is the justification for using the computer as a metaphor for the human mind. Nevertheless, the influence of the digital computer on the symbolic architectures of human cognition (Laird et al 1987, Anderson 1990) is apparent in their three basic premises:

- intelligent systems are physical symbol systems;
- cognition can be described in information processing terms; and
- cognitive processes are rational insofar as they use available knowledge to achieve goals.

The first premise was explicitly formulated by Newell and Simon (1976) as the physical symbol system hypothesis, which states that the ability to manipulate symbols and symbol structures is a necessary condition for intelligence. A symbol is defined as a pattern that denotes an external (perceived) entity or an internal symbol structure. The denotation of internal symbol structures is significant because it enables all knowledge within a system to be applied to solve a problem, despite being distributed throughout the brain (Newell et al. 1989).

The functioning of a symbolic architecture is illustrated here using Adaptive Cognitive Theory (ACT) as an example. This theory is one of the most complete symbolic architectures developed and has a good record of matching its models of psychological phenomena with relevant empirical evidence (Anderson 1990, Anderson 1993, Anderson, Matessa and Lebiere 1997). The more recent versions of the theory - ACT* (Anderson 1990) and ACT-R (Anderson 1993) - propose two distinct types of memory: declarative and procedural. Declarative memory stores discursive knowledge such as that derived from the perception of an object. Procedural knowledge represents tacit skills developed by practice. Although the two memories have some commonalities, they are distinguished by the following factors (Anderson 1993):

1. **Reportability** - Declarative knowledge is at least potentially discursive, whereas procedural knowledge is not;
2. **Associative Priming** - The recall of an element of declarative knowledge increases the likelihood (and speed) of recall of associated elements (for example, 'beach' may prime 'sand');
3. **Retrieval Asymmetry** - Procedural memory is stored as condition-action pairs and can only be activated by the condition and not the action;
4. **Acquisition** - Declarative knowledge is acquired from the environment, procedural knowledge is acquired by practice;
5. **Retention** - The retention of decalarative and procedural knowledge, even in the same domain, seems to be largely independent given current empirical evidence; and
6. **Dissociation** - Evidence suggests that the capabilities for acquiring declarative and procedural knowledge are largely independent.

The basic components of declarative memory are chunks, which are sets of around three or four elements. Elements may be patterns, such as alphabetic characters or other chunks; the latter permitting hierarchical organization. The elements in a chunk are ordered in some manner, such as a linear sequence or spatial arrangement. What a chunk encodes determines what information can be held by its elements (see figure 3.1). All facts derived from the environment are encoded using chunks. The encoding mechanism is fast, but memory organization is relatively inefficient.

The procedural memory of ACT provides a mechanism for representing skills that are: (a) gradually acquired; (b) relatively permanent; and (c) performed efficiently. The memory consists of rules, which are condition-action pairs, similar to the if...then statements of computer programming languages (see figure 3.1). Productions are fine-grained, so a number of them must be applied in sequence to achieve a goal. Rules are made generally applicable by the use of variables in their conditions. This generality increases the likelihood that several rules may match the stimuli in working memory at any time. For this reason, goals are encoded as part of the condition so that productions only fire when they are relevant to the goal in working memory.

The main concerns about symbolic architectures stem from the lack of any clear relations between the components of the architectures and what is now known about the physiology of the human brain. Consider the physical symbol system hypothesis. Newell (1992, p.427) explains its purpose as follows:

> "Symbols provide distal access to knowledge-bearing structures that are located physically elsewhere within the system. The requirement for distal access is a constraint on computing systems that arises from action always being physically local, coupled with only a finite amount of knowledge being encodable within a finite volume of space, coupled with the human mind's containing vast amounts of knowledge. Hence encoded knowledge must be spread out in space, whence it must be continually transported from where it is stored to

```
event1
  isa jumping
  agent fox1
  object dog1

fox1
  isa fox
  speed quick
  color brown

dog1
  isa dog
  character lazy
  number plural
```

Chunks of Declarative Knowledge

```
NEXT-COLUMN
IF      the goal is to solve an addition problem
        and C1 is the rightmost column without an answer digit
THEN    set a subgoal to write out an answer in C1
```

An Informally Specified Production Rule

```
NEXT-COLUMN
 =goal>
    isa addition-problem
    object =array
 =array>
    isa numberarray
    columns ($ =column2 $)
 =column2>
    isa column
 - bottomrow +
    answerrow blank
 - (=array>
    columns ($ = column2 = column1 $)
 = column1>
    isa column
    answerrow blank)
==>
 - subgoal>
    isa write-answer
    column = column2
 !push! =subgoal
```

A Formally Specified Production Rule

Figure 3.1: Examples of Declarative Chunks and Production Rules in ACT-R (From Anderson 1993)

> where processing requires it (distribution does not gainsay this constraint). Symbols are the means that accomplish this required distal access."

As the above quote illustrates, symbolic architectures function very much like digital computers, with symbols being passed from one part of the brain to another. Upon seeing a dog, for example, a symbol denoting 'dog' would be activated in long-term memory. If the person wanted to describe what he or she saw, then the symbol would be transmitted from long-term memory to the area of the brain concerned with speech production. This perspective seems to be at odds with psychological evidence in several ways. First, this approach regards intelligent processing as syntactic. Consequently, "symbol systems have enormous difficulty with context, especially in determining which areas of memory are relevant." (Chown and Kaplan, In Newell 1992, p.443). In contrast, models of memory with organizations that reflect the structure of the environment are context sensitive, as illustrated by connectionist architectures. The SOAR architecture proposed by Newell does provide some context sensitivity through the potential to have many rules that make use of the same symbol, with the appropriate rule being selected on the basis of the presence of other symbols (serving as contextual cues). A second criticism is that a neural basis for the symbol processing model has not been provided. Whether the symbol-level description can be convincingly explained at the physical level remains to be seen, although some critics claim that this is implausible (Hinton and Anderson 1989). A third point, relating to the information processing approach, is that symbolic architectures are unlikely to perform operations of human complexity in a sufficient time scale, given current knowledge of the speed of neuron activations (Gregory 1987). This claim must be qualified, however, by the observation that symbolic architectures could be implemented at the physical level in many different ways and the implementation chosen will have a substantial impact on performance. The final premise of symbolic architectures - that human cognition is rational insofar as its capacity allows - seems to be a reasonable approximation of human performance (Simon 1951, Simon 1982, Gigerenzer and Goldstein 1996). Although symbolic architectures at present are more regular in their performance than humans, this may be because the models are underspecified. It is possible, however, that this hypothesis may eventually have to be revised.

The symbolic approach was the first to be used for developing general cognitive architectures. As such, it has provided insight into some mental processes, articulating them with sufficient precision to facilitate comparison with actual human performance. They are abstract representations, however, that have only limited correspondence with psychological evidence. Their lack of attention to neurophysiology, as indicated above, also raises concerns about their validity. It is in recognition of these concerns that connectionist architectures have been developed to describe human cognition in 'neurally inspired' terms (Rumelhart 1989). Particular attention is paid by connectionist models to factors such as the speed of neuron activation and the possible implementation of cooperative, parallel processing (Anderson and Hinton 1989). For this reason, it has been claimed that such models are more accurate descriptions of human cognitive phenomena than symbolic architectures.

3.4.2 Connectionist Architectures

Connectionist models of memory can be positioned on a continuum, depending upon their representation of memory as specific or distributed (Hinton and Anderson 1989). Neuron specific representations are similar to symbolic architectures, consisting of networks of neurons that correspond with specific symbols. For this reason, they are not considered here. Distributed representations, in contrast, store knowledge implicitly in the connections between neurons, rather than explicitly at the site of a neuron. In this respect, distributed representation provides an implementation level model that is not obviously consistent with the physical symbol system hypothesis.

The basis of distributed associative models is the correspondence of a physical stimulus with a pattern of neuron activations. Activation patterns are paired in a cause-effect manner, such that specific environmental stimulation activates one pattern in memory that, in turn, activates another (see, for example, Feldman 1989). Learning takes place by adjusting the weights associating the elements of the two patterns. Various techniques for calculating weights have been proposed, including relaxation (Rumelhart 1989) and spreading activation (McClelland and Rumelhart 1981). Connectionist models of this sort capture several key perceptual and cognitive traits. The pattern-based approach naturally deals with similarity, for example, thereby providing a basis for: (a) the recognition of partial stimuli; (b) context sensitivity; (c) analogical reasoning; and (d) generalization/induction. Significantly, such behaviours are achieved without the need for explicit rules or - in contrast with symbolic architectures - an executive to apply rules to memories. Thus, cognition is not information processing in the conventional sense.

3.4.3 Summary

As already noted, the evaluation of cognitive models is made difficult by the lack of concrete metrics for testing their behaviour. The comparison of radically different architectures is further complicated by their description of cognitive processing at different levels of analysis. In terms of the three levels of cognitive modelling (Pylyshyn 1989), symbolic architectures are functional descriptions, whereas connectionist architectures are implementation level descriptions. As suggested by the outlines of the two approaches and their successes in accounting for some empirical evidence, both models provide some valuable insights and are not necessarily contradictory. Given the current state of knowledge in this area, it is, perhaps, sensible when selecting a memory model to consider which level of analysis is most appropriate and which aspects of memory behaviour are most relevant.

3.5 Working Memories

The benefits of long-term memory to the adaptive capabilities of an organism are clear. Knowledge of what has happened in the past can guide future actions so that needs are more likely to be satisfied and goals attained. Long-term memories are only useful, however, if they can be retrieved and utilized in conjunction with perceptions of the current environment. Working memory functions refer to much of this capability. The distinction between short-term and long-term memory was made in the 1950s on the basis of differences in the causes of forgetting for very recent and less recent memories. In light of these empirical findings, Broadbent (1958) proposed a short-term memory that was subject to trace decay over a period of seconds and a long-term memory for which the main cause of forgetting was interference between similar memories. Subsequent research challenged this view by showing that trace decay could account for long-term memory phenomena and interference could account for some short-term forgetting (Baddeley 1986). By the late-1960s and early-1970s, however, most memory models consisted of the following:

- sensory memories for each modality;
- a short-term memory store; and
- long-term memory.

Significantly, it was generally held that memories necessarily were transferred from sensory memory to short-term memory and then to long-term memory (Baddeley 1986). Research during the late-1960s and early-1970s began to bring this model into question. Craik and Watkins (1973), for example, demonstrated that the time an item remained in short-term memory did not affect long-term learning.

In response to the difficulties encountered in explaining short-term memory phenomena, an alternative explanation of learning was proposed based around the concept of levels of processing (Craik and Lockhart 1972). This framework suggests that the greater amount of processing to which a stimulus is subjected, the more likely it was to be remembered. Recognising a word, for example, is less likely to lead to its retention than comparison of the phonological similarity of the word with a reference word (for example, determining that dog sounds like log) (Baddeley 1986, p.23). A related finding is that maintenance rehearsal (or rote learning) is ineffective for retaining familiar words, but quite effective for unfamiliar words or sounds (Horowitz and Prytulak 1969). Thus, it seems that the amount of learning effort is a key determinant of the probability of retention.

Baddeley (1986) suggests that the levels of processing framework has considerable heuristic value in the area of verbal learning, although some exceptions have been identified empirically and some deficiencies in the framework have been identified (for example, it is a linear model, whereas much of the processing involved is now known to be parallel with feedback between stages). A particularly valuable qualification is the point made by Morris and Liebert (1977) that the retrieval situation affects how memories are best encoded. When one needs to recall memories on the basis of phonological cues, for example, then phonological encoding is most effective. Recall based upon semantic cues, however, will be more effective when semantic encoding is used. How knowledge of retrieval tasks affects encoding is unclear, however, although it has been found that visual processing is generally less effective than phonological which is generally less effective than semantic processing (Baddeley 1986). Indeed, extensive research into the role of working memory in learning and comprehension led Baddeley and Hitch (1974) to propose a model of working memory that consists of three components: an articulatory loop, a visuo-spatial scratch-pad and a central executive. The concept of sensory working memories has potentially significant implications for information systems design given that human communication relies heavily on language (i.e. the articulatory loop), whilst human-computer interaction relies more extensively upon visuo-spatial processing. For this reason, these aspects of working memory are now considered in detail, followed by a brief review of the concept of an executive.

3.5.1 Auditory Working Memory

The role of phonology in short-term memory was clearly illustrated by Conrad and Hull (1964) who found that subjects made more errors when remembering sequences of letters if the letters sounded similar. Subsequent research showed that the effect was caused by articulation, rather than hearing. Furthermore, a study by Ellis and Hennelly (1980) demonstrated that articulation time affected the number of items that could be held in short-term memory. Such evidence strongly indicates the existence of a phonological store of between 1 and 2 seconds of speech, the contents of which can be maintained by sub-vocal rehearsal (Baddeley 1986). The phonological store can be 'accessed' by both the articulatory processes and the auditory system, although only input from the articulatory processes guarantee rehearsal (Baddeley 1986). As demonstrated by Salamé and Baddeley (1982), however, noise (i.e. input from the auditory system) has a significant effect on the retention of visually presented words because it interferes with the use of the phonological loop for remembering words. Whilst sounds, particularly speech, interfere with phonological tasks, use of the phonological loop does not prevent visually based phonological tasks from being performed (Besner and Davelaar 1982). In other words, reading comprehension does not demand use of the auditory working memory, although it does often seem to be used, especially for tasks such as the disambiguation of homophones.

3.5.2 Visuo-Spatial Working Memory

Although less decisive than explanations of phonological working memory, an increasing body of evidence suggests the existence of a distinct visuo-spatial component to working memory (see, for example, Baddeley 1986). It remains a subject of contention, however, as to whether this memory function is visual, spatial or visuo-spatial in character (Kosslyn 1980). Nevertheless, several interesting findings are of interest in terms of information and ICT artefact design.

First, the interaction between visual and auditory memory functions are again corroborated by the ability to use auditory information to construct spatial memory (Baddeley 1986). Second, studies showing that spatial tasks can both enhance and interfere with visual memory tasks indicate that there is at least some spatial component to the memory function (Wright, Holloway and Aldrich 1974). Third, subjects are able to rotate representations of objects (Baddeley 1986), although errors suggest that this ability relies upon either non-visual representation or imagery of limited accuracy (Nickerson and Adams 1979). Fourth, visual information, particularly letters and words, are encoded into the phonological loop by adults, although this encoding is not immediate and, therefore, relies upon a visuo-spatial working memory (Kroll 1975).

Whether imagery is a component of memory or is purely phenomenological remains a matter for debate (although Baddeley (1986) regards the issue as a pseudo-question of limited value). It is clear from empirical studies that relationships between information and ICT artefacts and working memory are complex, affected by both development and experience. Hitch and Halliday (1983), for example, suggest that conversion from visuo-spatial to phonological encoding emerges at around the age of ten years. Such developments in the use of working memory must be accommodated if information and ICT artefacts are to be designed for different age groups (as is the case with learning difficulties). Studies of expert abacus users (Hatano and Osawa 1983), for example, also indicate that the overlearning of skills for ICT artefact usage have an impact on the individual's use of working memory. Thus, not only must the general properties of working memory be considered to improve the effectiveness of ICT artefacts, the impact of different ICT artefact designs on learning and problem solving abilities should be considered.

3.5.3 Executive Control in Working Memory

The central executive function refers to the limited capacity of the individual to attend to functional stimuli and/or active memories (Baddeley 1999). As such, the notion of an executive raises questions about consciousness (Mandler (1992) provides an outline of the different viewpoints). The only substantive finding relating to working memory is that subjects are only able to cope with a limited number of sensory inputs and with stimuli of only limited complexity (Oakhill, Yuill and Parkin 1988). This finding, combined with evidence of interference effects (see section 3.6.2), suggests that working memories are interrelated in some way and that this interrelation is one factor determining what is conscious to an individual at any given time. Beyond this very limited understanding, however, very little is known about the executive function in human cognition (although a recent workshop focused upon the issue and published its proceedings in late July 1999 - see Miyake and Shah 1999).

3.6 Use of Long-Term Memory

The preceding sections have considered how cognitive processing makes use of immediate sensory inputs by relating them to past knowledge, stored in long-term memory. This section expands upon this discussion by briefly reviewing the tasks in which long-term memory is involved. The utilization of long-term memory involves three interrelated tasks: encoding, storage and retrieval. Most of the studies of long-term memory task performance have focused on particular task domains. Given the aims of this thesis, the general discussion presented here is substantiated using examples relevant to social interaction (e.g. face recognition) and information artefacts (e.g. the interpretation of written information).

Perception and cognition are active processes in which stimuli are related to stored representations in memory. Although this process must clearly have a beginning, the earliest stages of human development can be ignored for the purposes of studying information systems development. Memory encoding will typically form a part of an activity, such as looking for a friend in a crowded room. One will determine that many of the persons in the room are unfamiliar precisely because there is no exact match between their appearances and knowledge of persons in long-term memory. At the same time as one fails to match stimuli with existing knowledge, some of the stimuli may be stored in memory. Furthermore, although one might fail to recognize a single face in a room, many stimuli will be recognized to be of familiar types, for example an unfamiliar person is still recognized as a person and features such as eyes and arms can be distinguished.

Early research into recognition suggested that the encoding process consisted of several stages. In reading text, for example, a person first detects letter features, then letters, perhaps syllables, and then recognizes entire words. More recent research indicates that although there are levels of this sort, they do not operate serially, with feedforward and feedback relations evident (McClelland and Rumelhart 1981). It should be noted, however, that some reservations remain regarding models of this sort (see, for example, Massaro 1988).

The complexity of encoding is further increased by the interrelations between the sensory modalities. When reading text, for example, there appears to be both direct links between visual stimuli and the relevant semantic knowledge, and a process mediated by phonological representation (Baddeley 1990). Smyth et al. (1994) illustrate phonic mediation with the simple example of the ease with which we read ‘wunss uppon a tyme’ as ‘once upon a time.’ This is possible because the phonological encoding of the first phrase approximates that of the second. In purely visual terms, however, the two phrases have quite different appearances. Conversely, human ability to deal with homophones (such as piece and peace) indicates that phonic mediation is insufficient to explain reading, although phonic mediation is better substantiated than the direct semantic process (vanOrden, Pennington and Stone 1990).

3.6.1 Encoding

In light of the above discussion, it is clearly desirable that, out of the many possible ways of encoding information, information should be encoded in a way that facilitates effective matching or retrieval. This is impossible, however, because a person will not know what the future uses of a new item of information will be at the time of encoding. Thus, a person must process information according to current task demands (Morris and Liebert 1977). Demanding tasks require more processing effort, and this increased processing results in more complex associations between memories, resulting in increased accessibility in the future (Johnson-Laird, Gibbs and deMowbray 1978). The task contingency of memory encoding is illustrated by the findings of Nickerson and Adams (1979) that the details of familiar objects, such as coins, are rarely remembered because their functional properties are sufficiently indicated by shape, size, colour and inscribed numbers. Details, such as the direction in which the head faces are unlikely to be decision relevant, except for domain experts, in this case coin collectors.

An important implication of memory encoding involving relations to existing memories is that recall will be improved by encoding as many attributes, relations and other links as possible (Baddeley 1990). As the number of links representing a particular memory increases, it becomes more likely that the memory will be distinctive in some respect. In contrast, experiences of many very similar experiences are likely to lead to difficulties in recalling specific instances. A related implication of the task contingent nature of memory is that learning can be improved by setting information acquisition in the context of performing a meaningful task (Morris 1979). Specifically, evidence suggests that memory retention is a function of the number of positive decisions made with respect to a memory item (Hanley and Morris 1987). It is important in setting tasks that they are relevant to the user and that the presentation of the information is meaningful to the learner, as illustrated by Morris, Tweedy and Gruneberg (1985), who found that one factor explaining the difference between expert and novice retention was that terminology made some information meaningless and, therefore, extremely difficult for novices to encode in a manner that facilitates effective retrieval.

3.6.2 Storage

Given the task contingent nature of long-term memory, the primary storage considerations are: (a) whether stored memories may be lost; and (b) whether adding to memory affects what is already stored. Forgetting was first investigated by Ebbinghaus (1885), who demonstrated that memories exhibit exponential decay. This finding has largely stood the test of time, although few if any psychologists have suggested that memories are forgotten. Indeed, Bahrick (1984) suggests that memory decays exponentially for some time before decay eventually stabilizes, from which point on performance for those memories does not deteriorate. There are two exceptions to the normal pattern of memory decay. First, the decay of memories that are recalled frequently is counterbalanced by reactivation (Rundus 1973) and the addition of further associations between the memory and memories relating to the tasks in which the memories are recalled. Second, memories associated with strong affective responses are resistant to forgetting, a phenomenon known as flash-bulb memory (Smyth et al. 1994, p.288-291).

Considering the effects of storing new memories on the ability to retrieve older memories, retroactive interference and proactive inhibition are the main empirical findings of interest. Retroactive interference occurs when recall of items memorized at T_1 is impaired by interference from items memorized at a later time T_2 (Baddeley 1990). The studies by Loftus (Loftus and Palmer 1974, Loftus 1979) of the effects of questioning on the memories of witnesses illustrate the significance of retroactive interference. Loftus (1979), for example, found that the use of different words affected estimates of the speed of a car in a film of a car crash. The word 'smashed' led to an average estimate of 40.8mph, 'collided' 39.3mph, 'bumped' 38.1mph, 'hit' 34.0mph, and 'contacted' 31.8mph. Proactive inhibition refers to the interference of earlier learning on later learning (Baddeley 1999). Comparatively little research has been conducted in this area, however, with mainly anecdotal evidence being provided (see, for example, Baddeley 1999). Although interference effects may be of significance in terms of understanding errors in repeated task performance where, for example, parameters must be held in the user's memory, it is important to note that the effects have only been studied for the memorization of simple lists. It has yet to be demonstrated whether similar effects will occur during more complex memory-based tasks (Smyth, Collins, Morris and Levy 1994).

3.6.3 Retrieval

Memory retrieval is a continuous and automatic process. At any given time, the activation of memories by sensory stimuli can have one of three possible results:

- the stimuli do not match with any memory trace;
- one memory trace is strongly activated; or
- several memory traces are activated to a similar degree.

Only the second of these possibilities guarantees memory retrieval. A failure to activate any memory trace indicates that the stimuli are novel, whereas multiple activations suggest that the stimuli are similar to several previous experiences and, consequently, further information is required to discriminate between them. The importance of single activation versus multiple activation is illustrated by the context sensitivity of memory recall. The context dependent nature of retrieval from memory was demonstrated by Godden and Baddeley (1975), who showed that considerable variations in environment affected the ability to recall items from a list. Less extreme environmental changes have been shown to have effects (Smith 1988). The findings generally indicate that the closer environmental conditions are at retrieval to those in which encoding took place, the better memory retrieval will be (see Smith 1994). This seems likely to be because precise cues will lead to a stronger activation of memories of single instances than of the categories to which they belong.

Although memory retrieval is effective with partial stimuli (for example, we can often recognize partial images of objects), the recall of related groups of instances can be impaired. Brown (1968), for example, found that subjects provided with a partial list of US states found it more difficult to recall the names of other states than those provided with no initial information. A likely explanation for this is that the presentation of some US states increases their activation levels in memory. Consequently, when attempting to retrieve US states from memory, those already presented are more likely to be recalled, effectively suppressing the required information. This type of effect has been suggested to be a likely explanation of 'tip-of-the-tongue' experiences (Smyth et al. 1994). Phenomena of this type have implications for information systems analysis and design practices, such as deciding when it is appropriate to provide information as memory aids to users. Providing instances of typical transactions to an end-user and asking him or her to recall exceptions, for example, could increase the likelihood that the analysis will be incomplete.

A final point to note with respect to memory recall is that its primary purpose is not to retrieve knowledge verbatim and be able to verbalize it. Memory is utilized to guide decisions and actions and, consequently, is reconstructive. This has several implications, ranging from the reliability of eye witness evidence (see section 3.6.2) to the improvements in recall achievable through the structuring of presented information (Bower et al. 1969). An interesting finding, given the increasing use of visualization and graphics in information systems, is the tendency for reconstructing imagery to be inaccurate, yet significantly increase the subject's confidence that the recalled information is accurate. Moreover, it has been found that confidence in inaccurate imagery is often higher than confidence in correct non-imagery that is recalled from memory (Morris 1992).

3.6.4 Improving Memory Performance

Research into the improvement of memory encoding and recall is of potential interest to information systems both in terms of systems analysis and the design of information systems and information artefacts. The cognitive interview technique (Geiselman 1988), for example, was developed to improve the amount and accuracy of information recalled from witnesses. The process could be drawn upon to improve systems analysis techniques and, in its current form, be used to reconstruct the events leading up to major events, such as systems failures:

1. reinstate the encoding context;
2. encourage free recall
3. encourage subject to recall the event from the perpective of other persons involved in the event;
4. encourage the subject to recall the event sequence in different orders, for example, in reverse chronological order.

In terms of designing information artefacts, the studies of Herrmann and Petro (1990) and Intons-Peterson and Newsome (1992) provide some indication of how the cognitive modelling of human memory can contribute more towards the improved design of information and ICT artefacts. Herrmann and Petro (1990) identify the following classes of commercial memory aids:

- **Memory prosthetic** - a device that facilitates memory performance, such as a wristwatch with an alarm;
- **Memory corrector** - a device that can assist in correcting memory errors; and
- **Memory robot** - a device that performs memory tasks for its owner, such as a thermostatic control.

Memory prosthetics were found to be by far the most common type of memory aid (61 out of 71 memory aid types). Observing that "almost all that is known about the utility of commercial memory aids comes from anecdotal reports made by consumers," Herrmann and Petro (1990, p.447) suggest that the utility of memory aids is affected by a number of factors including: (a) the importance of the memory problem to the user; (b) the user's ability to deal with the memory problem without memory aids; and (c) the effectiveness and convenience of memory aid use. They also suggest that active reminders, typically containing audible alarms, tend to be more effective than passive aids, although they have a tendency to become distracting and, consequently, can fall into disuse. Furthermore, it has been noted that memory aids might reduce the memory performance of the user (Estes 1980).

3.7 Learning

The use of knowledge of past experiences to guide future actions requires an ability to identify similar types of phenomena and make generalizations about them. Cognitive science models of these abilities are based upon mechanisms for categorizing the knowledge stored in long-term memory. Inevitably, the choice of memory architecture essentially determines the way in which categories are formed. Given that aspects of categorization can be empirically studied much more readily than precise memory function, the choice of memory architecture should, given present knowledge, be constrained by empirical evidence of categorization activities. Thus, in evaluating alternative models of memory and categorization in human memory, it is useful to consider the following aspects of human cognition, which have been subject to substantial empirical study:

1. the categorization of and generalisation about nominal and natural kinds;
2. the perception and categorization of phenomena based on partial stimuli;
3. the encoding of personal and functional attributes;

4. generalization does not considerably inhibit the ability to recall specific experiences; and

5. communication about specific phenomena and general categories.

One way of categorizing experiences is to establish definitions that include the singly necessary and jointly sufficient conditions for category membership (Smith 1989). The necessity condition stipulates that every instance belonging to a category must possess that attribute. The sufficiency condition states that any instance that possesses all defining attributes must be a member of that category. Empirical evidence suggests that this approach to categorization applies well to nominal kinds, but does not provide a satisfactory explanation of categorizing natural kinds, which do not have (or have usually imperceptible) necessary and sufficient conditions. To illustrate this limitation, consider the nominal kind 'grandparent' in contrast with natural kinds, such as animal species. A grandparent has the necessary and sufficient condition of being 'a parent of a parent' (Smith 1989). In contrast, 'lion' does not have necessary and sufficient conditions that can be readily ascertained (even if this can be established genetically, this information is not accessible by direct sensory stimulation).

Considering the second issue above, the absence of information about any necessary condition will lead to a failure to classify or to misclassification. Moreover, this type of categorization will be sensitive to small information deficiencies, rather than exhibiting the 'graceful degradation' characteristic of human performance (Rumelhart 1989). The third requirement - including personal and functional attributes - implies the encoding of non-necessary attributes into a category definition. The attribute 'retired,' for example, might be applied to 'grandparent,' although this is only a stereotypical trait. Such non-necessary attributes cannot be included within a category specification because they contradict the necessity requirement. The inadequacies of using necessary and sufficient conditions have led to their rejection as a viable explanation of human categorization. Investigations of the problems with this approach have, however, made the significant contribution of demonstrating the importance of similarity and typicality relations in deriving categories from experiences.

A more realistic approach to dealing with natural kinds is to define categories in terms of the typical attributes of instances in that category. Such an approach inevitably produces fuzzy categories, which often overlap. Consequently, some ambiguity and occasional inconsistencies in classification result, although these behaviours seem to reflect human performance reasonably well. By far the most popular techniques for categorizing on the basis of typicality are prototype-based (Smyth et al. 1994). Prototypes are typically represented by schemata, which identify the attributes relevant to a category, the relative importance of the attributes in determining category membership, the relationships between attributes, and default attribute values. Prototype categorization typically relies upon the frequency of perceptible attributes of phenomena (Smith 1989) and is, thus, particularly useful in dealing with natural kinds.

Instances of phenomena are matched with prototypes on the basis of similarity by a process of 'best fit.' Consequently, incomplete descriptions of phenomena can still be dealt with, albeit with reduced accuracy. Again, this is an advance because it is similar to human performance in terms of the second requirement identified above. The third requirement - allowing the encoding of personal and functional attributes - is also satisfied by prototypes because the approach does not require category descriptions to consist of only necessary attributes. It is quite possible, for example, for a tomato to match the prototype for a vegetable owing to its functional attributes, even though it only satisfies the necessary conditions for being a fruit. The use of schema as the basis for determining category membership has a further advantage that composite categories can readily be dealt with. The prototype for 'vehicle,' for example, can easily be augmented to provide the prototypes for 'car' and 'motorcycle,' even if this requires attributes of the supertype 'vehicle' to be removed. The more formal category specifications preclude such complex category relationships, which seem to be characteristic of human categorization (Smyth et al. 1994).

Considering the final requirement - the ability to support communication - symbolic prototype-based categories appear to offer satisfactory support for communication about similar phenomena, making specific reference to perceptible, personal and functional attributes. This appears to contrast with the exemplar approach, which claims that only instances are stored, with categories being inferred. The ability to communicate about general concepts and specific instances, the final requirement identified above, is easily implemented with the symbolic architecture. Symbols that denote concepts are simply associated with other symbols. It is useful at this point to bear in mind the assertion made by Marr (1982) that the implementation can only be considered valid if it is meaningful at the semantic/knowledge level. The association of concept symbols with other symbols has limited support in this respect (see section 3.4.1).

As noted above, connectionist models based upon distributed representation deal effectively with similarity. As illustrated by Anderson and Mozer (1989), such memory models, if they incorporate feedback, will effectively categorize experienced phenomena, grouping together similar stimuli. Models of this sort behave in a manner very similar to prototype models. Salient features that are common to many stimuli will lead to large adjustments in the strengths of associations between patterns of neurons. Less salient features will lead to associations being developed more slowly and uncommon features will serve to individuate instances that have unique characteristics (Anderson and Mozer 1989). The resultant behaviour of the system is promising in terms of the requirements set out above.

Natural kinds are also dealt with by connectionist models in an effective manner. Initially, all instances with gross similarities will be grouped together. Over time, the most salient phenomena are well-established and their associations become saturated so that they no longer strengthen, or strengthen very slowly. Less salient attributes, however, continue to develop associations leading to the division of the initial categories to reflect finer distinctions between instances. In the example developed by Anderson and Mozer (1989), for example, the letters C, K, N, W and X are all placed in the same category initially. With repeated exposure to instances of the letters, however, separate categories are formed for C, N and W, and K and X. With further experience, distinctions are made between instances of all letters. One limitation of this approach is that it does not explain nominal categories, which may be learned by a special case of this process or in a different way.

As already noted, distributed representation is reconstructive, with partial stimuli resulting in the activation of a complete pattern in memory (see section 3.6.3). Of course, a partial stimulus may not provide sufficient information to determine a specific category. In the example above, a partial stimulus may result in both 'K' and 'X' being activated because the less salient features that differentiate between the two categories are not contained in the stimulus. This type of behaviour exhibits a graceful degradation similar to that of human recognition and categorization.

The fourth criterion set out above - enabling specific instances to be recalled - also appears to be satisfied in a way that bears similarity to human performance: exactly similar instances will activate specific memory patterns, whilst unfamiliar ones will activate more general patterns relating to categories. Furthermore, recognizing instances still makes category information available, although some form of external cue will be required to encourage the subject to 'switch' from using instance to category knowledge. The final requirement - supporting communication about specific instances and general categories - can also, in principle, be satisfied because associations can be developed between any patterns, including patterns specific to instances and more general patterns that approximate to category prototypes.

3.8 Representation

Having recognized elements of the environment and activated/recalled relevant knowledge, the ability to engage in complex interactions with the environment requires that the elements are put together to provide an overall impression of the environment. Explanations of how persons comprehend their environments are central to the study of higher cognition. Early accounts of reasoning suggested that formal logical rules were used to represent and reason about the environment. Two observations bring into question the plausibility of this claim. First, human performance in applying logical rules is generally poor and can only be improved by extensive training in logic (Smyth et al. 1994). Second, and most significantly, the logics proposed are insensitive to the content of the sentences used to represent the problem whereas human reasoning is highly content sensitive. Consider, for example, the *modus ponens* rule:

a $\rightarrow$ b

a

therefore b

This works well in some real world cases, particularly those in which the referents are linguistic categories:

x is a dog $\rightarrow$ x is a mammal

x is a dog

therefore x is a mammal

Modus ponens fails, however, for much real world reasoning where the referents are concrete particulars:

x has a headache $\rightarrow$ x should take paracetamol tablets

x has a headache

therefore x should take paracetamol tablets

Clearly, this is not generally true because x may, for example, be allergic to paracetamol. Whilst it is possible to construct more complex rules, observational evidence suggests that a person does not recall all complexities at once. Furthermore, which complex rule was appropriate could only be determined by inspection of the referents being used in each case. Examples from spatial reasoning also demonstrate the importance of representations in enabling comprehension and reasoning, For example, the propositions 'the square is to the left of the circle,' 'the triangle is to the left of the square' and 'the triangle is to the left of the circle' create difficulties for purely propositional reasoning. If, however, one has a model of the shapes drawn on the surface of a cylinder, then there is no contradiction in shapes being both to the right and left of one another. The representation is necessary for the meaningful comprehension of the three sentences. Such examples do not, of course, demonstrate that propositional representations are never used. They do, however, corroborate the studies of Mani and Johnson-Laird (1982) that propositions may be used when representations are indeterminate and alternatives, such as mental models, used when determinate representations can be constructed.

A more plausible model of reasoning can be constructed by representing semantic content in memory so that the sentences used in reasoning have referents that facilitate comprehension (de Vega et al. 1996). A person will use initial evidence to construct a representation and proceed to search both memory and the environment for counterexamples and additional factors to augment the representation. Seeing bottles of medication by a patient's bed, for example, may trigger the inclusion of combined drug effects into the representation, which then leads to a memory search for possible instances including paracetamol. One approach to achieving this type of representation is to use mental models.

Unlike propositions, which represent everything using predicates and arguments, mental models use tokens to represent instances, the properties of tokens to represent the properties of instances, and the relations between tokens to represent relations between instances. Thus, models can be regarded as isomorphic to the aspects of the world that they represent (Johnson-Laird 1983). The theory of mental models makes three key predictions:

- the more models that are required to reason about a problem, the longer problem solving will take and the more error prone it will be;
- erroneous conclusions should be consistent with the premises rather than inconsistent with them; and
- general knowledge can influence the mental process of deduction.

Given the problems with propositional representation outlined above, the last of these predictions suggests that mental models are potential improvements as descriptive accounts of human reasoning.

The use of analogues in visuo-spatial reasoning has been given some empirical support by experiments that indicated subjects were performing the equivalent of mental rotation (Shepard and Metzler 1971). These claims that a continuous representation was being used were further supported by Kosslyn who subsequently proposed an analogue theory of representation (Kosslyn 1980). The evidence is not decisive, however, and much of it could readily be explained using propositional representations (Pylyshyn 1981), which are discrete and, consequently, imply that subjective experience of imagery is an epiphenomenon (Pylyshyn 1981). Evidence that subjects make errors when reasoning about physical objects (Hinton and Anderson 1989) has also been used to argue against analogue representations, although such evidence only discounts claims that analogue representations are always accurate, rather than discounting analogue models altogether.

3.9 Social Cognition

Although the above discussion of cognition has paid some attention to the context sensitive nature of human action, little has been said about the role of cognition in social interaction. To address this issue and to provide a point of contact between the cognitive issues addressed in this chapter and the social issues addressed in chapter 4, the social cognition literature is now explored. Although something of an oversimplification, the social cognition literature can be seen as augmenting the 'what' and 'how' of human cognition with the 'why' of human motivation. In so doing, social cognition theories provide a basis for explaining the relationships between personal values and the *raisons d'être* of social institutions.

3.9.1 Causal Attribution

Causal analysis is a central component of an individual's attempts to explain events. As with other mental acts, the imputation of cause is made with varying degrees of intentionality, typically receiving most attention when events are unexpected or undermine one's beliefs (Hastie 1984). Everyday causal attribution is explained by attribution theories. Consistent with cognitive accounts of rationality, these theories assume that causal analysis is valuable to the individual because it assists in predicting and, to some extent, controlling the future (Fiske and Taylor 1991).

Recognizing that causal inference must be based upon different kinds of evidence, Kelley proposed two models of causal attribution. The covariation model (Kelley 1967) derives attributions using evidence from many similar events, whereas the use of causal schemata (Kelley 1972) enables attributions to be made with respect to a single experience. Both models enable attributions to be made to other persons, environmental factors and to oneself. The motivational basis for the models is uncertainty (and the resulting loss of self-confidence) arising from, for example, a lack of support for one's beliefs, the inability to solve a problem, or a high level of ambiguity in available information (Fiske and Taylor 1991, p.33).

The basic principle of covariance analysis is to attribute cause where one type of event is always followed by events of another type. Where one type of event is sometimes followed by another and sometimes not, covariation is lower and a causal relationship is less likely. Kelley (1967) suggests that a person will consider three factors when making attributions based upon covariance. Distinctiveness refers to the extent to which covariance is common across many 'effects', given the same 'cause' (in the case of attribution to oneself, for example, the subject considers 'does this always happen to me?'). Consensus refers to the extent to which covariance is common across many 'causes' (the subject might ask him- or herself 'does this happen to others or just to me?'). Consistency over time/modality refers to the 'strength' of the covariance (does it happen every time or just some of the time? does it happen in all types of situation or just for specific types?). High levels of distinctiveness, consensus and consistency are likely to lead a person to make a causal attribution to an observed actor (i.e. attribution of a personality trait). In contrast, low distinctiveness, high consensus and low consistency are likely to lead to a causal attribution being made to the entity being acted upon by observed actors (i.e. an entity has a property that leads all persons to act upon it in a particular way) (McArthur 1972). Clearly, the type of attribution made will have significant implications for the likely future behaviour of the individual making the attribution. The covariation model proposed by Kelley is similar in many respects to the statistical procedure, analysis of variance (ANOVA). Several empirical studies have brought into question whether variance analysis is a valid approximation of the process by which causal attributions are made (Hilton and Slugoski 1986, Hilton and Knibbs 1988).

Descriptive accounts of causal attribution, such as White (1988), identify a range of influencing subjective attributions of causality: (a) cause precedes effect; (b) temporal contiguity; (c) spatial contiguity; (d) perceptual salience; and (e) covariation. The first four of these seem to be learned by children and continue to play a significant role in causal attribution throughout adulthood, particularly in unfamiliar situations. In domains of expertise, adults often acquire more complex attribution rules, including distal causality, attributions to multiple causal factors and the use of the representativeness heuristic (Fiske and Taylor 1991). An important observation relating to information search and acquisition is that people tend to distinguish between events (including emotional responses) and intentional actions (Zuckerman and Feldman 1984). When attempting to establish the causes (i.e. reasons) of general types of action, subjects usually look for similarities in the responses of numerous other actors. When trying to establish why a particular individual acted in a particular way, however, subjects will primarily seek multiple examples of that person's behaviour.

One difficulty raised by the possibility of multiple and distal causes is explaining how they are apprehended by an observer. Whilst spatio-temporally contiguous causal relations might be learned on the basis of sensory stimuli and without extensive cognitive processing (as demonstrated by operant conditioning), more complex attributions suggest that some form of abstract mental representation is being used. Hilton and Slugoski (1986) propose that this is achieved by comparing events with one's world knowledge, which takes the form of 'scripts.' A script describes the typical ordering of events for a class of causal relations. Events are compared with appropriate scripts to determine causes. Where an unexpected outcome occurs, this will be attributed to any abnormal condition that distinguishes the particular event from others in the same class.

Although helping to explain some empirical evidence, the abnormal conditions model raises the question of how scripts containing normal causal relations are formulated. Consequently, the theory by itself is incomplete. It does, however, emphasize the importance of taking into account the role of knowledge of past experiences in making causal attributions. A more complete view of causal attribution might combine some of the simple attribution rules identified above with some form of covariance analysis and, using empirical evidence, model the relationship between general rules and domain knowledge.

Whilst normative models of causal attribution have proven useful starting points for developing an appreciation of human performance, various systematic deviations from these models have been observed. It is worth paying careful attention to these findings because they provide potential insights into some dysfunctional aspects of systems development and have influenced therapeutic techniques, such as reattribution training (Forsterling 1985), that might contribute to improving systems development practice.

The most well known deviation from covariation models of attribution has been termed the fundamental attribution error (Fiske and Taylor 1991). This 'error' refers to the tendency for subjects to attribute the behaviours of persons whom they observe to their personality/disposition without accounting for situational characteristics, such as the person fulfilling a role or acting according to appropriate norms and conventions. Although not yet fully explained, it is generally agreed that a major reason why situational characteristics are downplayed is that they are less salient to the observer. Factors present in the locale of action are likely to be more static than the actor and, therefore, will not be well attended by the observer . Roles and norms, however, only become apparent through their influence on the person being observed and, consequently, are more likely to be regarded (with some justification) as part of the actor's disposition. It is interesting to note that this attribution is learned, as indicated by the observation of young children making causal attributions to concrete situational factors (White 1988). Moreover, the tendency is more pronounced in Western cultures (Fletcher and Ward 1988). The plausibility of the salience account of the fundamental attribution error is reinforced by evidence that the reverse is true of self-attribution. When evaluating one's own actions, one will tend to be aware of the many environmental and other situational effects that influence one's actions (Taylor and Fiske 1978). One is also aware of one's own feelings and intentions, making it more plausible that the action is a result of one's current mood, rather than stable personality traits.

The false consensus bias (Marks and Miller 1987) refers to the tendency to assume that one's behaviour is more typical than it actually is. An equivalent bias has also been found with respect to the evaluation of groups to which an individual belongs (Hewstone 1989). Whilst it is unreasonable to expect that a person who does not conduct systematic samples will make unbiased estimates, some possible reasons for systematic bias are of interest. Marks and Miller (1987) suggest that the bias may result from the tendency to associate with similar others, thereby leading one to experience a biased sample of behaviours. They also suggest that the need for self-confidence and the avoidance of anxiety may lead subjects to regard their own beliefs and values as appropriate, thereby biasing their view of what is typical and atypical. A final point to note is that personal experience is the only detailed source of evidence upon which a person can draw when reasoning about others' beliefs and values and, consequently, some bias in reasoning is inevitable.

3.9.2 Social Categories and Schema

Categorization has already been considered as a general cognitive function. It is briefly reviewed here in terms of its implications for social interaction, along with the closely related concept of a social schema. Schemata are cognitive structures that represent generalized knowledge about concepts or types of stimuli (Fiske and Taylor 1991). Schema include the various attributes of concepts/types and the relations between them, forming an holistic impression that facilitates top-down, theory driven mental processes. In contrast, the use of sensory information available at any given time provides the basis for bottom-up cognitive processes. Whether processing is predominantly top-down or bottom-up depends on the nature of the information available and the motivation of the individual.

An early application of categories and schema to social cognition was the study of person perception. Asch (1946) suggested that overall impressions are created by relating various attributes of an individual to form an integrated whole. One consequence of this model is that attributes take on different meanings according to the overall context of the situation. Intelligence, for example, is usually regarded differently when it is an attribute of an honest or dishonest person. An alternative to the holistic model of person perception is the elemental approach, whereby attributes are treated in isolation and summary judgements are made about individuals on the basis of average/total attribute scores. Fiske and Neuberg (1990) suggest that both models explain different aspects of person perception behaviour, again depending upon information availability and motivational issues.

As illustrated in section 3.7, the formation of categories plays a fundamental role in cognition. It is, therefore, not surprising to discover that many aspects of an individual's social behaviour can be explained in terms of the development and use of categories and the related concept of social schema. The roles of categories and schemata in social interaction are threefold:

- the individual generalizes across experiences of past social events in order to anticipate the outcomes of present and future social events;

- the individual is able to develop plans for his or her own actions based upon their anticipated consequences; and
- perceived probabilities of success affect future actions, which on a social level leads to routinization of certain kinds of activities and the elimination of others.

The third point - that social activities become regularized - has been given the most attention in the literature on social organization and is, consequently, discussed in chapter 4. The remaining two points are considered here.

As noted in the discussion of categorization (section 3.7), categories are not developed in terms of necessary and sufficient conditions. Research in social cognition has raised alternative explanations for how social categories are formed. First, it has been suggested that instances are assigned to categories depending upon how well they match a prototype or ideal/extreme type that serves as the 'centre' of the category (Hayes-Roth and Hayes-Roth 1977, Barsalou 1985). This type of generalization over persons and social events suggests a top-down, schema-driven, approach to social action that is strongly influenced by prior experience. In contrast, exemplar models (Smith 1984) and associative models (Wyer and Carlston 1994) suggest that person perception and other aspects of social cognition are bottom-up (i.e. data-driven). According to these views, categories are stored as collections of instances, rather than being defined in abstract terms. As pointed out by Fiske and Taylor (1991), both views have their limitations and it seems that a combination of the qualities of top-down and bottom-up models will provide a more satisfactory account. For this reason, prototype, associative and exemplar models are briefly reviewed in turn and summarized to indicate how an integrated model of the role of categorization in social cognition might arise.

3.9.3 Prototype Models

Observations that subjects rate, for example, dogs as more typical pets than spiders led social psychologists to describe categorization as the comparison of experienced instances with prototypical instances. By classifying objects and events in this way, a person would be able to use general schemata for acting in classes of situations. Consequently, it would be possible for an actor to draw upon relevant past knowledge to guide actions in similar situations that arise in the future. A person could, for example, have a general schema for eating at a restaurant, which would consist of being seated at a table, ordering from a menu, and so on (Fiske and Taylor 1991). More specific schemata might be developed for eating at a Chinese or fast food restaurant. Recognition of objects (such as chopsticks or waiters) would prompt specific schemata, whereas the failure to activate a schema for a specific type would lead to the use of the more general schema as a guide to action. This model has featured significantly in social cognition because it appears consistent with observations and intuitive understanding of much social interaction.

A prototype, according to this view of social categorization, consists of the central features of a category. Consequently, it is possible that a prototype instance is never actually experienced by an individual, although most of the features of many category members will usually be similar to those of the prototypical case. Furthermore, some features will be more diagnostic than others. Apples and oranges are typical fruits, for example, whereas a tomato is not. A key reason for this is that tomatoes are often used as vegetables rather than as fruits, despite having many of the typical physical properties of fruits. Thus, functional and personal properties can be regarded as important for social categorization, consistent with the general assertion made in section 3.7.

The value of the prototype model for explaining social perception arises from the combination of the hierarchical organization of categories and the fuzziness of category boundaries. Cantor and Mischel (1977), for example, showed that subjects provided with information about the particular features of a person tend to 'fill in' other attributes using the typical values of the nearest prototype. A subject will make very different personality attributions, for example, if told that someone is a builder or a lawyer. Furthermore, considering the relative diagnosticity of traits, Skowronski and Carlston (1987) found that more weight is usually given to negative or extreme attributes. Thus, personality attributions will be considerably different for male and female builders and for white middle-class and immigrant American lawyers.

3.9.4 Exemplar Models

Although prototype models have formed a major component of the social cognition discipline, there are a number of weaknesses with this account of evidence concerning social perception. One of the more serious of these deficiencies is the difficulty in accounting for subjects' abilities to recognize specific instances. If an abstract prototype is stored in memory for categorization purposes, then memories of specific instances must be stored separately. This not only seems inefficient (Nosofsky 1988), it contradicts the view that prototypes are recognized even though they have never actually been experienced (Neumann 1977). Further limitations include the failure to account for alternatives to prototypes, such as defining some categories in terms of ideal or extreme types (Barsalou 1985), as may be the case with categories of scholars, priests and criminals, for example. Categories also seem to be dependent upon goals and knowledge, indicating that they may be more contingent than the hierarchical model suggests (Fiske and Taylor 1991). This is reinforced by the difficulty in defining categories (especially social categories) so that even a significant number of their defining features are necessary. Exemplar models of social categorization attempt to overcome some of these limitations.

Exemplar approaches begin with the view that all instances are encoded in memory and that category assignment is achieved by assessing the similarity of an instance with each instance of all categories. The new instance is assigned to the category to which, relatively speaking, it is most similar (Nosofsky 1988). This approach overcomes some of the above difficulties in accounting for both knowledge of instances and knowledge of categories.

Furthermore, as shown by Nosofsky (1991), exemplar models accurately predict some categorization behaviours. The exemplar approach is, however, very similar to the prototype model in that it regards categorization as based around central tendencies emerging from experienced instances. One potential criticism of this model is its inefficiency. Whilst Nosofsky (1988) criticizes the prototype approach for being storage inefficient, the exemplar appears as if it may be inefficient in terms of its processing demands because it implies extensive pairwise comparison. This cannot be determined, however, unless specific commitments are made regarding the physical instantiation of the exemplar process. It would, no doubt, be infeasible for a von Neumann computer, although not necessarily for the highly parallel human brain.

3.9.5 Associative Models

Reviewing the cognitive architectures outlined in section 3.4, it is evident that connectionist architectures effectively function in terms of central tendencies when used for recognition, because strengths of association are determined by the frequencies of exposure to different combinations of stimuli. Thus, the typical features of category instances will have the highest strengths, whilst unusual features will be highly predictive. All of these behaviours corroborate the social perception studies referred to above.

Connectionist models of categorization are similar to exemplar models in that they group instances together if they share many common attributes. The connectionist models differ, however, in that they adjust the weights for category membership if a categorization error is made, whereas exemplar models operate on the basis of correct categorization only. A further difference between connectionist and exemplar models, which proves to be advantageous, is that the networks in connectionist models operate on combinations of features as well as operating on the individual features, enabling the subject to identify which of several cues is more predictive (Shanks 1991).

3.9.6 Self Perception

Although the preceding sections have distinguished cognition from personal values, little has yet been said about the personal value system. Central to any understanding of such a concept is self perception - the way in which a person conceives of his or her personality and establishes a sense of identity. Early analysis of the self led to the emergence of Freudian and, more recently, ego psychologies, which both rely upon the concepts of id, ego and superego (Gregory 1987). Ego psychology serves as the basis for the theory of structuration (Giddens 1984) and some of its implications for information systems are appraised in chapter 4.

The Freudian division of the personality is essentially arbitrary (Gregory 1987, p.271) and provides only a limited basis for explaining and planning human activity. The social cognition literature offers an alternative framework of theory and empirical research, which is considered here. As will become evident, although many of the findings of social cognition researchers are of potential value for understanding information systems related changes, the ontological and epistemological bases of the various theoretical claims are not entirely convincing.

Self perception research in social cognition has focused primarily upon the individual's representation of his or her own personality attributes, social roles, experiences and goals (Fiske and Taylor 1991). This is desirable from an information systems perspective because it emphasizes the role of self perception in decision making and action. The literature reflects a concern with three basic issues regarding the representation of self: the need for accurate diagnostic information concerning one's self; the general goal of maintaining a consistent representation of self across the range of situations and possible actions; and the tendency to create a positive view of oneself (Fiske and Taylor 1991). As illustrated by references to maintaining consistency, it is generally assumed that an individual has different self-concepts for different kinds of events. These do not reflect different personalities, but different portrayals of personality that the individual attempts to manage.

For this reason, Fiske and Taylor (1991, p.182) distinguish between chronic and working self-concepts, as follows:

> "Which aspect of the self influences ongoing thought and behaviour depends in large part on which aspects of the self have been accessed. The aspect of the chronic self concept that is accessed for a particular situation is termed the *working self concept*. . . the working self concept is important because it guides social behaviour and in turn is modified by feedback from the situation."

The chronic self concept consists of numerous schemata, each of which represents personal experiences in a given domain. The schemata are established in long-term memory and it is well established that they are resistant to change (Backman 1988). Whilst this can be limiting, the stability is critical for the maintenance of a person's identity. Less visual information is stored about the self than about others and there is some indication that more reliance is placed upon verbal encoding (Lord 1987). It is acknowledged that the schemata are affect laden (Fiske 1982), although little indication of the role of proprioceptive stimuli is given in the social cognition literature.

As noted above, the working self concept consists of several schemata active at a particular time. The activation of schemata is automatic and unintentional, responding to environmental circumstances (Bargh and Tota 1988). Once active, schemata act like perceptual filters, with important dimensions of the self-concept having well-developed schemata and, consequently, resulting in related perceptions being processed more rapidly and extensively than those relating to unimportant aspects of the self (Fiske and Taylor 1991). There is also some evidence that schema relevant attributes are perceived more readily than aschematic attributes, whilst the availability of a relevant schema will affect how associations are made between observations. Thus, the schemata active as the working self concept can be likened with interpretive frameworks. As schemata reflect goals and priorities, they also serve as guides to future action. In the absence of external constraints, a person will act in a manner consistent with his or her self concept, thereby reinforcing identity, avoiding anxiety and other negative emotions, and increasing the likelihood that personal goals will be realized. Like expert knowledge, highly developed schemata will yield quick responses in particular domains, whereas novel but related situations will often result in lengthy delays because the subject must process a large amount of relevant knowledge (Fiske, Lau and Smith 1990). Underdeveloped schemata, in contrast, will provide rapid, but often inaccurate, responses for all situations in a given domain.

It has been suggested that personal goals and values are sometimes reflected in the construction of possible selves (Markus, Cross and Wurf 1990). This model proposes that a person is capable of constructing a (partial) self concept that reflects what he or she would like to be. This process involves an extensive amount of information acquisition. At some point, often determined by external factors, the focus will shift from information acquisition to learning by direct experience and, in doing so, eventually realizing some or all aspects of the constructed 'ideal self.' If, for example, a person wishes to appear more competent at using a computer system, he or she might construct an image of him- or herself touch typing, using different types of software, helping colleagues with their computing problems and so on. The person may then acquire instruction in learning to type and so on. Eventually, as expertise increases, the reliance on, for example, teachers and reference books will decline.

A related theory is the self-discrepancy theory proposed by Higgins (Higgins 1987, Higgins 1989). This theory also suggests that a person makes comparison between his or her current and ideal selves. Higgins goes on to suggest that discrepancies between the two self concepts can lead to agitation-type emotions, such as fear, restlessness and anxiety. Moreover, similar affective responses may result from a discrepancy between a person's current self concept and the ideal of that person conveyed by a family member, friend or colleague. Higgins also suggests that confusion and/or indecisiveness may result from contradictions between self concepts. This might occur if, for example, a person constructs a self concept idealized with respect to certain moral values and another that emphasizes traits relating to popularity with peers. Self-discrepancy theory also suggests that having the ability to achieve goals but failing to do so can lead to fatigue or listlessness (i.e. motivational depression), whilst being unable to achieve goals can lead to a sense of hopelessness. In this regard, an interesting observation is made by Linville (1987) who suggests that individuals with complex self concepts (usually arising from a varied lifestyle) tend to have less extreme emotional reactions than those with more simple self concepts. The explanation offered for this is that failure for a person with a complex self concept only relates to a small part of the self concept, whilst other parts of it may be perceived to be successful. In contrast, a person whose life is centred upon very few types of activity will perceive failure as much more significant because it impacts upon a large proportion of what represents personal achievement.

Clearly, self perceptions have very significant implications for how individuals react to changes in their social environments. Thus, an appreciation of stakeholders' self concepts may contribute to the more effective planning and implementation of IS-related organizational change. One obstacle to conducting such an analysis is that, whilst the chronic self concept may be very stable, working self concepts and self descriptions vary with context (McGuire, McGuire and Cheever 1986). Self descriptions often emphasize attributes that make the individual distinctive in a given context, although prolonged social interaction usually results in self descriptions (and the descriptive language itself) being aligned with the self descriptions of others (Wurf and Markus 1990).

As changes inevitably destabilize social interaction, they may lead many individuals to perceive a loss of control. This often results in increased stress because the individual feels that he or she is less able to work towards an ideal self and to achieve personal goals. Several mechanisms for helping an individual to regain a sense of control (not necessarily to regain control) have been identified by Thompson (1981).

- behaviour control
- cognitive control
- information control

- decision control
- retrospective control
- secondary control

It is interesting to compare the findings of social cognition researchers in this area with those of the socio-technical approach to organzational change. Deci and Ryan (1987) and Taylor and Brown (1988), for example, suggest that control over and flexibility of one's behaviour leads to greater motivation and interest, better conceptual learning, higher self-esteem, higher levels of trust and greater persistence at behaviour change. This is consistent with the claims of Mumford (1995) and others regarding job enrichment. Increased control has also been observed to have negative effects, however, where the associated increase in accountability makes subjects feel a greater need to perform.

3.10 Decision Making and Social Inference

From the perspectives of both developing and using information systems, the process of collecting and combining information for reasoning and decision making is of critical importance. A considerable amount of work has been conducted in comparing human decision making with normative decision rules under the general heading of behavioural decision theory. This work and some broader findings relating to information search strategies are reviewed here.

A prerequisite to any reasoning and decision making is the acquisition of relevant information. Although it is usually possible for a person to recall some relevant information from memory, existing knowledge is often only sufficient to guide the acquisition of information in the current situation. It is possible that subjects will sometimes draw upon specific past experiences, but it is usually the case that generalized knowledge, in the form of a model or theory about a problem domain, will be utilized (see sections 3.7 and 3.9). The use of theories to guide information search suffers from two main weaknesses: (a) the theory may be wrong and misguide the search; and (b) the theory may be used without consideration of available contextual information, leading to reduced search performance. It should be noted, however, that a person will be more or less confident in his or her theories and the level of confidence will influence the extent to which it is relied upon to guide information acquisition.

Having collected relevant information from the environment, it must be analyzed and integrated with existing knowledge. Several problems can arise at this stage. First, direct experience is often limited to a few instances of a given type of situation or problem. Despite sample size being low, a person will often generalize these experiences in order to guide his or her action. Several potential problems have been identified in this regard. First, it seems that decision makers are relatively insensitive to sample size, failing to recognize the possible biases present in a small sample, such as the undue influence of a few extreme instances (Tversky and Kahneman 1982a). Second, where information is collected by consulting others, the biases arising from social organization are often overlooked. Consulting friends, for example, is unlikely to provide representative responses because social groups are typically composed of individuals with similar interests and attributes (Fiske and Taylor 1991, p.351). Third, regression toward the mean is generally overlooked when predicting performance (Kahneman and Tversky 1982a). On the basis of exam results, for example, it is usually overlooked that very good and very poor performances are unlikely to be repeated. This judgemental 'error' has been suggested as a possible explanation for the common belief that punishments are more effective than rewards (Kahneman and Tversky 1982a).

A further observation, the so-called dilution effect, suggests that a prediction or decision made using only relevant information will be moderated in response to the provision of information that is irrelevant to the decision. A possible explanation for this effect, at least with respect to social judgements, is that the decision maker utilizes all information to construct a model of the person or situation in question, drawing upon prototypes to develop a more complete model (see section 3.7). This model, perhaps supplemented by memories activated by the relevant and irrelevant information, is then used as the basis for decision making. Tetlock and Boettger (1989) found that the dilution effect increases with increased accountability for one's decisions. This is possibly explained by the increase in accountability motivating the use of a wider range of information sources so that decisions appear to be based on more thorough evidence and, consequently, better justified.

In order to discover how well a person is capable of integrating information regarding many similar instances, numerous studies into the use of base rate information have been conducted. It has typically been observed that too much attention is paid to specific experiences rather than to statistical information. It should be noted, however, that few of these studies have explored the use of base rates by those individuals who have experienced a large sample of similar events (i.e. domain experts) (Borgida and DeBono 1989). Of those that have, it has typically been found that experts are considerably less biased than novices, suggesting that, whilst the use of presented statistical summaries is often inadequate, the 'statistical' properties of experiencing large samples are reasonably sound. Furthermore, Ajzen (1977) has noted that base rate information seems to be utilized to a greater extent where its causal relevance is perceived, suggesting that some base rate errors are exaggerated by experimental design. Bar-Hillel (1980) expands upon this observation by suggesting that relevance is the primary concern, being defined by factors including causal relevance and representativeness. A further observation, made by Hinsz et al. (1988), is that the credibility of information sources seems to play a determining role in the use of information.

Difficulties observed with information integration have been found in terms of probabilistic reasoning. The conjunction error (Tversky and Kahneman 1983), for example, describes the tendency for decision makers to believe that multiple pieces of evidence suggest that an outcome is more likely. Although this sounds intuitively correct, decision theorists argue that it is inconsistent with statistical reasoning because the conjunction of two events is necessarily less likely than either event (Tversky and Kahneman 1983). Whilst the statistical case holds for uncorrelated events, it is reasonable to expect that most information integration activity is performed in order both to guide and predict future outcomes and, thus, is likely to focus on associating evidence to identify causal chains. Thus, although it is statistically less likely that a manager will be both charismatic and intelligent than charismatic or intelligent, causal reasoning suggests a preference for the presence of both characteristics and would, therefore, guide the appointment of a manager. When a decision maker is asked whether it is more likely that a manager will be charismatic and intelligent or just intelligent, he or she may well interpret the problem in terms of selection problems and, thereby, respond 'incorrectly.' A related point is made by Fiske and Taylor (1991) who suggest that decision makers construct 'explanatory frames' for events, which are deemed more plausible as they become more detailed. Furthermore, it is possible that prototypes or schemata are used in decision making, with conjoined characteristics having better fits with prototypes than single attributes (see section 3.7). Similarly, when analyzing events, Locksley and Stangor (1984) found that multiple causes are preferred to explain rare events and single causes for common events. Given the complex nature of causal reasoning in a dynamic environment, such a heuristic is likely to be functional on the majority of occasions. Thus, whether 'conjunction errors' demonstrate flawed reasoning or merely show that normally adaptive behaviours do not work in some experimental situations is a matter for further investigation.

Given the reliance upon causal reasoning to explain some properties of human decision making, it is instructive to reconsider how covariation information is processed by decision makers (see section 3.9.1). Nisbett and Ross (1980) suggest that, in comparison with normative models, human decision makers do not perform especially well, primarily as a result of their search strategies. A fairly robust finding is the tendency to search for confirmatory evidence, but pay scant attention to disconfirming evidence (Arkes and Harkness 1980), despite both being equally important for identifying causal relations. It seems likely that this finding represents a general focus of attention on action, with less interest being expressed in what conditions were absent when something happened or what conditions were present when nothing of interest happened; a point supported by the findings relating to scenario generation (Kahneman and Miller 1986). Nevertheless, several possible ways of improving performance on covariation tasks have been identified, including (Fiske and Taylor 1991, p.371):

- assessments of covariation are more effective for small samples;
- assessments of covariation are more effective when information is presented simultaneously rather than serially;
- assessments of covariation are more effective when information is presented in summary form;
- instructions regarding the assessment of covariation must be clearly communicated;
- repeated exposures to information tend to lead to improved accuracy of assessment;
- the presence of competing information sources usually leads to reduced performance in covariation tasks;
- covariation assessments are made more effectively by decision makers with a sound appreciation of non-covariation; and
- the problem domain is familiar and the information presented in a readily codable form.

So far, the focus has been upon information acquisition and organization. Various other 'errors and biases' have been found to arise as a result of the way in which the task or decision is presented to the decision maker (Bar-Hillel 1982, Kahneman and Tversky 1982b, Kahneman and Tversky 1982c, Tversky and Kahneman 1982a, Tversky and Kahneman 1983, Gigerenzer 1994). The most well-known framing effect relates to the assessment of risks. Tversky and Kahneman (1981) found that problems phrased in terms of gains led to risk avoidance, whereas problems phrased in terms of losses seem to encourage risk seeking behaviour. The empirical study used to demonstrate the effects of framing referred to loss of life and it has been argued that this may, in itself, lead to unusual modes of decision making because the stakes involved make such decisions particularly difficult. An alternative explanation, however, is that the different descriptions of the problem lead to different models of the problem situation being constructed. Problems described in positive and negative terms, for example, can convey different impressions regarding the population involved. The choice between saving 0 lives and 600 lives and between losing 600 lives and losing 0 lives is only guaranteed to be equivalent for a population of 600. Reviewing the experiment by Tversky and Kahneman (1981), it seems that there is scope for misinterpretation in this respect. Thus, framing effects of this sort are open to question. Nevertheless, sensitivity to problem descriptions must clearly be considered as a potential factor influencing decision making.

The final stage in the decision making process is the application of a rule of choice. Several judgemental heuristics have been identified that seem to be used as convenient decision rules for a wide range of decision types. The representativeness heuristic (Bar-Hillel 1982, Kahneman and Tversky 1982c, Tversky and Kahneman 1982b), for example, appears to be used quite effectively where precise, detailed information is unavailable, as in the case of personality judgements. In such cases, classification tends to proceed on the basis of best fit, with little attention paid to the frequencies of occurrence of each class in the environment (i.e. base rates). As already noted, all available information tends to be used for classification/model construction, regardless of its relevance to the problem domain in question. Although a frequently effective rule, the representativeness heuristic is sometimes applied in inappropriate problem domains. Gamblers often assume, for example, that some random sequences are more typically random than others (Kahneman and Tversky 1982a) (it is interesting to note, however, that Fong, Krantz and Nisbett (1986) found that randomness is more likely to be appreciated with respect to machines, but not with respect to a person's actions).

The availability heuristic suggests that decision makers base some of their frequency judgements on the ease with which they can recall instances from memory (Tversky and Kahneman 1973). In light of connectionist models of human memory, it seems plausible that frequency of experience will strengthen memories and, in turn, increase the likelihood of recall. Consequently, the availability heuristic is a plausible cognitive model and seems likely to be valuable in situations where precision is not required. There are potential limitations to the effectiveness of the heuristic, however, such as the influence of situational context. If asked how many people of one's own age have children, for example, a person at work may recall work colleagues, whilst a person at home may recall friends. Given that one is more likely to have met one's friends' children than the children of work colleagues, the different contexts may lead to different probability estimations. It has not been established, however, whether such context sensitivity can be particularly detrimental. Furthermore, it has been suggested that availability is more often used to identify an appropriate schema to use to guide decision making, rather than for making judgements.

Another heuristic that has received considerable attention is the use of anchoring and adjustment (Slovic and Lichtenstein 1971). This heuristic describes the tendency for decision makers to centre their estimations around presented information and is clearly important in terms of information systems. In general, it seems that this heuristic is useful for making predictions, although it is clearly dependent upon the accuracy of the information source providing the 'anchor' information. Little is known, however, about how well decision makers detect misinformation.

Looking at the human decision making process as a whole, it seems that, despite numerous possible sources of error and misjudgement, performance is adequate in most situations, being most problematic when attempting to utilize both presented information and knowledge derived from experience. This interpretation - in contrast with the claim that humans are poor decision makers because they do not conform to normative standards - suggests that a great deal more study needs to be conducted into information design and the development of information and communication technologies that support the entire decision making process. Although normative models are valuable when information is constructed with the intention of using normative decision rules, it is important to recognize that personal experiences lead to the development of knowledge that is best utilized via less formal judgemental heuristics. Furthermore, care must be taken not to confuse the satisfaction of a normative standard with the needs that motivate human decision making (Taylor and Brown 1988).

3.11 Summary

Information systems needs to understand the individual not just as a user but as a social actor who uses information systems and ICTs. This kind of understanding requires consideration of human needs and motivations, cognition, behaviours and actions, and how they are affected by social interaction. Psychological studies relevant to each of these issues has been reviewed in this chapter to show that, although the issues have individually been addressed, the 'conceptual distance' between the explanations is too great for such a diverse body of knowledge to provide a practicable basis for developing information systems theories. In terms of developing a more integrated theoretical basis for information systems, the discussion in this chapter has shown the following issues to be particularly relevant:

- Human sensory physiology is highly selective in terms of information acquisition and is, consequently, a major determining factor in how the environment is interpreted. This must be taken into account when presenting information (a similar point has been made by Kosslyn 1994).

- As shown in section 3.3, there is quite a marked divide between accounts of objective physiology and accounts referring to subjective experience. This divide results from ontological and methodological commitments made by different areas of cognitive science.

- The limited understanding of the brain's physiology has led to the formulation of a number of currently plausible accounts of memory organization and cognitive processing. The competing theories have some broad similarities and some of their main differences primarily result from differences in their chosen levels of analysis.

- There is a substantial body of evidence suggesting that the working memory function is composed of distinct short-term working memories for audition and vision, which are interrelated. The nature of the interrelationship is not entirely clear, but experiments have demonstrated several implications for information processing and reasoning tasks (Baddeley 1986, Baddeley 1990, Baddeley 1999).

- Long-term memory is reconstructive. Thus, perception, recognition and recall can essentially be regarded as different aspects of the same cognitive function. A specific stimulus pattern can be simultaneously encoded into long-term memory, activate similar patterns in memory and result in the construction of a mental representation that includes related memories.

- The impacts of memory aids and other ICT artefacts on memory performance was found to be poorly understood (section 3.6.4).

- Categorization and generalization play very important roles in structuring knowledge so that it serves as an effective guide to action. These processes bear some similarities to, but are not as precisely defined as, their formal logical equivalents. Instead, categorization seems to be based upon similarity, rather than the matching of specific attributes, and generalization seems to arise from the central tendencies arising from the associations between many similar instances.

- There exist two competing views regarding the form of mental representations, the analogue and propositional models. Both provide useful explanations of some evidence but are inconsistent with other research findings.

- The basic processes of categorization and generalization are extended in complex ways to accommodate knowledge of the social environment. The basic processes appear to be essentially the same, however, indicating that social judgements rely upon categories that are vaguely 'defined' by central tendencies and that reasoning about social action is highly sensitive to the content and context of particular problem situations. In general, the most specific knowledge to be recalled from memory will be utilized in judgement and decision making. This account is different from normative decision theories, which are based upon content-free and context insensitive decision rules.

- The central role of the senses in providing information to the associative memory results in social categories being formed around observable objects and events. Consequently, there is a strong tendency for observations to be attributed to actors, who are usually the most salient features in social situations. Consequently, there is a tendency to reason about social interactions in terms of the dispositional attributes of individuals, with less weighting being given to the context in which actions are observed. The opposite is true of an individual's self concept because environmental factors are particularly salient when reasoning about oneself and one's actions.

- Individuals construct numerous self-concepts for themselves, each relating to particular types of social situations. Motivations are related to world knowledge via the construction of an idealized self-concept that represents the self in a desirable future state that can be realized by acting in the current social situation. The various self-concepts will rarely be consistent and individuals often experience stress and anxiety because actions that are desirable with respect to one self-concept are undesirable with respect to another. Furthermore, environmental and social pressures may result in individuals acting in ways that are not consistent with any of their self-concepts.

Chapter 5 uses the above summary of psychological issues as the basis for developing a theory that aims to address some of the problems identified in chapters 1 and 2.

Chapter 4

Social Organization and Organizational Forms

4.1 Introduction

The previous chapter presented a detailed analysis of the individual, including an examination of the individual as a social actor. This chapter considers the actor in a social context and considers the manifestation and development of organizational forms. Given the concerns of this thesis with evaluation and investment decisions, particular attention is paid to the economic nature of the firm. An important part of this analysis is the comparison of the individual actor, as described in chapter 3, with actor characteristics as assumed by theories of social organization. Whilst considerable simplification is to be expected, any contradictions between psychological and social analyses will clearly be problematic for the IS community, which must balance user impacts against organizational costs and benefits in order to ensure the effective use of computer-based systems.

The following section begins the study of social organization with a brief review of different conceptions of social structure. Giddens's theory of structuration (Giddens 1984) is then examined in detail as one of the most significant contemporary social theories explicitly to address the ontological status of social structures. The implications of theories of social organization for the development of economic models of intra-firm and market activity are then considered, followed by a brief analysis of change management issues, which are key considerations for any development process.

4.2 The Nature of Social Organization

Epistemological and methodological issues have always played significant roles in the development of the social sciences. The standpoint adopted with respect to these issues has significant implications for the accounts of social organization presented by different theorists. From the late 19th century onwards, one of the most significant influences upon social theory was the positivist conception of the natural sciences. Some social theorists advocated the application of natural sciences methods - as characterised by logical positivism - to the social sciences, whilst anti-positivists held that this was impossible owing to the nature of the subject matter. The central concern is essentially the distinction between behaviour and action. Behaviours are observable and subject to 'objective' description. Like the weather, social behaviours are not readily predictable owing to their complexity, but there are both general tendencies and low-level regularities. Action, in contrast, implies intention. Thus, even if 'what' and 'how' are describable in objective terms, why particular actions were taken by an individual in a given situation cannot be explained without some recourse to subjective interpretation on the part of the observer. The following review of early social theories shows how the emphasis on behaviour or action translates into more or less deterministic social theory.

Weber (1949) addressed many of the issues raised by positivist and anti-positivist views in sociology. He asserted that human behaviour was no less predictable than some natural phenomena and sufficient to enable individual actors to predict others' responses sufficiently well to engage successfully in social interaction. Predictability did not, however, imply determinism or fatalism because he made a further distinction between two types of rationality. Purposeful rationality is choice based upon the comparison of different means and ends (much like evaluation as assumed by 'rational economic man'), whereas value rationality is the pursuit of an assumed preference (for example, profit maximization). The distinction played an important role in Weber's subsequent work, leading him to rule out the possibility that scientific/rational analysis could validate or reject value judgements. In considering political organization, Weber concluded that bureacracies maintain existing values but do not create or change them (Weber 1947). In government, therefore, politicians pronounced value judgements as policy goals, whereas bureaucrats worked towards political goals by making rational, rule-based choices.

The form of bureaucracy described by Weber has had a major influence on theories of organization design. Moorhead and Griffin (1992) identify the following key elements of Weber's bureaucracy in relation to organizations, rather than just the bureaucracy of government:

- rules and procedures;
- division of labour;
- hierarchy of authority;
- technical competence of workers;
- separation of ownership;
- rights and property associated with jobs not individuals; and
- documentation.

Although, as Moorhead and Griffin (1992, p.588) suggest, Weber may have "conceived of a bureaucracy as a powerful and efficient form of administration," he was clearly also aware that the benefits of 'the organization' were achieved at the expense of workers, as illustrated by his discussion of the division of labour:

> "The Puritan wanted to work in a calling; we are forced to do so. ... This order [the modern economic order] is now bound to the technical and economic conditions of machine production which today determine the lives of all the individuals who are born into this mechanism, not only those directly concerned with economic acquisition, with irresistble force." Weber (1930)

The frequent use of the machine analogy and the mention of 'irresistible forces' is suggestive of social structures that exist as constraints on individual actors. Weber does not make an ontological commitment to social structures, however, but advocates a strong methodological individualism (Giddens 1995). The determinism apparent (not actual) in the quote above can be explained on two levels. First, owing to Weber's interest in politics and organization at the societal level, he makes a sharp distinction between those who make value judgements (i.e. politicians) and those who operate according to purposeful rationality within the boundaries set by politicians (i.e. bureaucrats). This distinction is inevitably blurred when applying the concept of bureaucracy to the firm because (assuming the complete divorce of ownership and control) even the broadest of value judgements (such as the organization's mission) are made by appointees who also have responsibility for policy implementation. Second, at the societal level of analysis, social organization can appear to the individual actor as irresistible forces such as the economic and educational systems, whereas this is less true of individual perceptions of organizational forms. Clearly, one must be careful not to take Weber's mention of 'social forces' too literally, as indicated by his advocation of methodological individualism. The constitution and roles of organizational forms and other structural properties of social organization must, however, be explained.

To Durkheim (1933), writing at around the same period, the emergence of internally differentiated societies was a consequence of their increasing complexity. As a society undertakes an increasing number of activities, specialization becomes essential. Although Durkheim does not provide a detailed validation of this claim (Giddens 1995), focusing more upon relations between society and individuals than upon the nature of differentiated social organization within society, he makes some interesting points with respect to the emergence of 'occupational associations.' A claim of significance to this thesis is that the formation of firms and other institutions is an essential component of the division of labour. In terms of political institutions, Durkheim's explanation of why this is the case often refers to moral issues. With respect to capitalist markets, however, he adopts the premise that firms emerge from the class divide between those who have property and resources and those who provide their labour. Consequently, he describes the firm as characterized by a relationship of exploitation, similar to that described by the principal-agent problem in economic theory. It is for this reason that Durkheim advocates the regulation of markets by government. The nature of the relationship also suggests a motivation for the development of other institutions, such as trade unions. Durkheim recognizes, however, that rules are not sufficient. A well-informed labour market is also essential if organizational forms are to contribute to an effective labour market because nothing should prevent individual workers "from occupying the place in the social framework which is compatible with their faculties" (Durkheim 1933).

The discussion so far has pointed to the emergence of organizational forms, including firms, but provided little detail about the processes of development involved. It has also indicated the importance of both clear statements of ontological commitment to social phenomena and the need to understand economic phenomena that involve individuals working in differentiated organizations within a society. These issues have continued to represent major themes in sociology and economics, and are now considered in turn. The ontological issues are discussed in terms of a critique of Giddens's theory of structuration, which is the most significant recent social theory to be based upon an explicit treatment of the ontological status of different aspects of social organization. Economic considerations are then addressed in an analysis of the roles of technology and information within firms. This provides a sharp focus on the issues of most concern to the IS community, whilst addressing the general socio-economic concerns identified above.

4.3 Theory of Structuration

One of the most significant recent developments in social theory is the theory of structuration (Giddens 1984). This theory directly addresses the division between naturalistic social theories, which regard social structures as objective phenomena serving to constrain individuals' behaviours, and hermeneutic social theories, which focus upon individual action with little interest in identifying social entities. The theory of structuration effectively dissolves some of the conceptual differences between the two perspectives by redefining the ontological status of a number of key sociological constructs. In doing so, Giddens aims to overcome the conceptual difficulties inherent in each perspective.

Whilst not advocating methodological individualism, as such, structuration theory places a strong emphasis upon the reflexive character of human conduct - an idea imported from hermeneutic social theory - and the explanation of social phenomena in cognitive, especially linguistic, terms. The basic ontological shift made by the theory of structuration is the substitution of the dualism of the agent as social subject and society as social object with the duality of structure, whereby actions are based upon and reproduce social relations extended over space and time. Thus, social structure does not exist independently, but is present in the memories that guide individuals' actions. The structural properties of social organization are reproduced intersubjectively through their physical manifestation during human activity. This and other concepts of structuration theory have significant implications in terms of the models of the individual and of social action that they entail. Thus, for the purposes of this research, it is useful to develop an analysis of structuration theory in terms of its implications for: (a) human cognition; (b) the human body and physical presence in the environment; and (c) social organization *per se*.

4.3.1 Cognition in Structuration Theory

As already noted, structuration theory places significant emphasis upon the psychology of the individual when explaining the reproduction of social phenomena. Thus, an important first stage in reconciling theories of human cognition and human action is to compare and contrast the psychological assumptions underlying structuration theory with the cognitive science models presented in chapter 3. Given the criticism that cognitive science typically use models of disembodied cognition (Anderson et al. 1997), particular attention is paid to the emphasis of structuration theory upon the role of the body in human performance. The aim of this analysis is to bring out the similarities and differences between the two perspectives and to suggest how the cognitive level, essential for effective ICT design, can be redeveloped using ideas from structuration theory and to identify the key issues involved in synthesising the two perspectives.

Just as perception and, therefore, cognition are regarded in cognitive science as subject to a continuous stream of environmental stimuli, the theory of structuration speaks of experience as a durée. Owing to its high level of abstraction, however, structuration theory offers no account of the workings by which experience affects the individual's memory in such a way as to facilitate future action. Nevertheless, several psychological characteristics are assumed that can be directly contrasted with those of cognitive science. Influenced by ego psychology's concepts of id, ego and super-ego (Gregory 1987), Giddens bases structuration theory upon three levels of consciousness: the basic security system; practical consciousness; and discursive consciousness. The basis in ego psychology implies a number of deviations from explanations phrased in terms of perception, cognition and action as proposed by cognitive science models. The most basic difference, related to the disembodied nature of cognitive science models, is the claim by structuration theory that perception is an active process. Whereas cognitive models tend to consider perception as passive, presenting an 'objective image' to the cognitive processes that make sense of it (see section 3.4), structuration theory regards perception as bound-up with action. Specific aspects of the environment are focussed upon by the individual and, consequently, other aspects of the environment are barely recognized, if perceived at all. Furthermore, the integration of the senses is emphasised in contrast to the reliance of cognitive science research upon accounts of the senses as operating with considerable independence (see sections 3.2 and 3.3).

Considering higher level cognition, it is useful to examine the notions of practical and discursive consciousness in terms of the empirical claim of cognitive science that increased exposure to situations that the individual assigns to the same category leads to improved performance in terms of recognition and subsequent action. With sufficient experience, cognitive science claims, such actions are performed with very little explicit cognitive attention. Relatedly, structuration theory considers many routine skills to be so familiar to the individual that they are performed in an almost automatic way, with little, if any, explicit attention. Furthermore, structuration theory considers the division between practical and discursive consciousness to be unstable, changing over time. This is consistent with the cognitive science models, which suggest that cognitive demand changes over time. An account of which actions can or cannot be explained by the actor, however, is provided by neither theoretical account.

The concept of practical consciousness is basic to the central claim of structuration theory that social activity relies heavily upon routinization. For any particular social encounter, the individual is acting in a social context. The actions are constituted by the individual's knowledgeability of social situations, and the experience feeds back into the individual's memory. Thus, agents' knowledgeability is constitutive of social encounters and recursively implicated in social activity. Such a perspective is not inconsistent with the accounts of memory and learning already presented in sections 3.4 and 3.7, although such models assume discrete episodes rather than durées. This inconsistency is present, however, within the cognitive science paradigm, with perception considered in terms of a stream of stimuli, yet categorization being based around discrete entities and events. Structuration theory does provide some insight into how this inconsistency can be addressed: social engagement, because it is constituted in time-space, occurs in an episodic manner, thereby suggesting that 'events' can be isolated by high-level cognition as the basis for categorization. This explanation relates to the structurational account of motivation and value.

Day-to-day, routine activity should, according to structuration theory, be considered as a process of continuous action, not as a set of events. This action stems from the acquisition of knowledgeability about 'how to go on' in routine social contexts. This knowledge of how to act in highly familiar social contexts is part of the practical consciousness. Giddens claims that the use of routinized knowledge, which is constitutive of social activity, suggests "...a general motivational commitment to the integration of habitual practice across space and time." (Giddens 1984, p.64). In other words, routinized behaviour not only reduces cognitive demands by regularising basic social activity so that it requires little conscious cognitive effort (e.g. tacit knowledge of turn-taking in conversations), it is also claimed to avoid anxiety by satisfying the individual's need for 'ontological security.' Such motivational factors are beyond the scope of cognitive science models and comparisons cannot be made (although the discussion of self-perception in section 3.9.6 makes reference to anxiety). Evidence of adverse effects of eliminating such routinisation does, however, suggest that the concept of ontological security and its roots in routinisation, which can be considered in cognitive science terms, does have some validity. The claim that routinization is constitutive of social activity also has implications for considering the importance of environmental context in cognitive performance.

Action, according to structuration theory, requires the individual to have the power to intervene in the environment. As Giddens notes, "this presumes that to be an agent is to be able to deploy (chronically, in the flow of daily life) a range of causal powers, including that of influencing those deployed by others." (Giddens 1984, p.14). Thus, the agent's perceptions of power can be directly compared with the process of categorisation by which individuals attribute causality to their own actions and to other events in the environment. Given a complex, dynamic environment, which social activity surely entails, the difficulty of perceiving and delineating all relevant factors in the environment suggests that the actions of the individual, whose perceptual and cognitive abilities are selective and/or limited, will be of limited accuracy. This uncertainty and inaccuracy accounts, in cognitive terms, for some of the unintended consequences that, according to structuration theory, arise from the individual's actions. The other significant cause of uncertainty is the existence of other sources of intervention and change in the environment, most notably other actors.

Whilst the above discussion suggests a number of areas of overlap, which provide scope for reconciling accounts of human cognition and the structurational theory of human agency, there are a number of psychological traits assumed by structuration theory that fall outside of cognitive science. Perhaps the most significant of these, particularly in terms of this research, is the consideration of motivation. The suggestion that practical consciousness implies that many actions are not directly motivated has significant implications for the subjective evaluation of technological artefacts. If individuals cannot discuss their motivations, either directly or indirectly, except where routine breaks down and the individual is forced to exert cognitive effort in identifying a course of action, it seems unlikely that such individuals will be able to provide reliable subjective evaluations, particularly where the range of events to be valued is wide. Any detailed consideration of value concepts is difficult, however, given the limited attention paid to them by both structuration theory and cognitive science. For this reason, value concepts are discussed in socio-economic terms in section 4.5 and in psychological terms in chapter 5.

4.3.2 The Body in Structuration Theory

Models of higher level cognition have an almost exclusive focus upon mental processes. What is actually being 'processed,' however, is generally taken as given, with perception effectively regarded as a passive process of information acquisition. Structuration theory, in contrast, places a significant emphasis upon the body situated in an environment. Thus, the individual is regarded as an "...integrated system of perceptual and communicative modalities." (Giddens 1984, p.67), which is active in its perception of the environment. This advanced conception of human agency provides several insights into the effects of ICTs on social activity.

To understand the role of the 'integrated body' in social activity, it is necessary to consider the individual situated in his or her physical environment. Where the activity involves social interaction, the physical settings are described as *locales*. A locale is a region in space with definite physical boundaries that serves as a focal point for certain social activities. Locales are typically internally regionalised, providing differentiation of context for activities within the same locale. A workplace, for example, is a locale physically bounded by the construction of the building. Within the workplace, specific rooms or areas within rooms are used for different purposes. At a more day-to-day level, several individuals may form a bounded region simply by the positioning of their bodies during a conversation.

Giddens (1984) provides a classification of different modes of regionalization based upon four attributes: form, duration, span and character. The form of a region refers to its defining boundaries, which may be physical or symbolic. In episodes of co-presence, the region is usually demarcated using bodily posture, tone of voice, and other aspects of positioning. Such episodes are typically very brief in duration. Where types of interaction become institutionalized, however, more permanent forms of regionalization, such as the use of a specific room at regular times, may be established. Duration refers to the temporal regularities of social activity. Specifically, social interaction is episodic in form, with each encounter occurring within a defined time period. Spatial regions are also used differently at different times. There are, for example, typical working hours during which business districts are most active and typical times of the day when rooms of a house will be used. Span refers to the spatial extent of a region. A house, for example, is precisely defined by its physical construction and legal records of land boundaries. A city, however, has a much less well-defined boundary, with the distinction between urban, suburban and rural subject to gradual, but continual change. Character refers to the functions for which regions are used. In some societies, for example, the home serves as the centre for both family relationships and working activity, whereas modern capitalist societies tend to have distinct regions in which working and family lives take place.

In terms of information systems and ICT artefact design, the psychological and social impacts of regionalization are potentially significant, particularly where ICTs are used to mediate social interaction, rather than merely support it. Whether intentional or not, ICT-based information systems will lead to the formation of regions, which can be supportive or restrictive, depending upon their appropriateness for the social activity in question. How such boundaries are perceived is central to understanding the impacts of ICTs on social interaction and are, therefore, discussed in detail in chapter 5.

4.4 The Organizational Form

The previous sections of this chapter have considered social organization in general. One of the main themes of this analysis was that social structures do not exist as such, but can be discerned as regularities in the interactions of individuals over time. Some identifiable sets of social regularities are of particular significance in terms of information systems development because they are recognised within societies as specific kinds of interaction and are accorded legal status. These social phenomena are referred to here as organizational forms, specific forms of which include public sector bureaux and firms. This section first reviews a number of research studies investigating the nature of organizational forms generally, before considering the distinguishing characteristics of each of these types of organizational form in terms of their development and usage of information systems.

4.4.1 Structural Aspects of Organizational Forms

A major premise of organizational behaviour research is that the structure of an organizational form is a determinant of how successfully its goals are achieved (see, for example, Moorhead and Griffin 1992). To be able to appreciate the value of this premise, a clear understanding of the nature of organizational forms and their attributes and relations is required. This section reviews a selection of organizational theories to indicate how accounts of organizations and their characteristics have developed.

As indicated in the discussion of structuration theory, organizational forms cannot be regarded as concrete, objectively describable entities in the same sense as a physical particular, such as a building. It is not surprising, therefore, to discover many different ways in which the structural aspects of organizational forms have been interpreted and described. One of the earliest approaches to describing organizations was to consider the division of labour and the mechanisms used to supervise and coordinate the activity of the workforce. Beginning with the work of Smith (1776), this perspective came to be exemplified by the principles of scientific management and Fordism.

The division of labour was most rigorously applied within the manufacturing industries, where it resulted in the efficient use of labour resources by creating jobs that were very narrow and well-defined. Workers quickly became skilful at performing their very limited range of tasks and, because they had minimal discretion, made a standard contribution to a large process (such as a production line), often at a pace dictated by managers coordinating the process or its sub-processes. Even within manufacturing, an overemphasis on division of labour has been acknowledged as quite severely limited (Shaw 1981). The mechanistic nature of the resulting jobs led to very low commitment and job satisfaction, which could limit efficiency as well as the quality of process output. The use of tighter supervision cannot entirely overcome these difficulties and clearly offsets a proportion of the cost savings.

Considering how the level of supervision can impact upon the ways in which workers perform their tasks, span of control is a major factor in organization design. If the span of control is inadequate, workers may not perform sufficiently well and coordination problems could arise. Too much control, however, may alienate workers and have other harmful effects, as well as unduly increasing costs. Mintzberg (1979) suggests that the appropriate span of control is contingent upon a number of factors: (a) the degree of job specialisation; (b) task similarity within a given unit; (c) the type of information available to and needed by unit members; (d) members' needs for autonomy; and (e) workers' needs for access to their supervisors. Clearly, some of these factors reflect the interaction between span of control and other structural aspects of an organization, such as power and authority relations. Furthermore, none of these 'dimensions' is readily quantifiable and any decisions regarding changes to an existing span of control rely on the judgement of managers and their experience within the organization.

The degree of centralization within an organization refers to the extent to which decision making authority is retained by managers at the top of a hierarchy or delegated to lower level employees. As indicated above, centralization is closely related to span of control. A supervisor with many subordinates, for example, can only manage them effectively if each of them can work with considerable autonomy. Such autonomy is achieved either by making tasks highly programmatic or by permitting subordinates to make some decisions by themselves. Recognising this, Simon (1960) made a distinction between programmed and non-programmed decisions. Programmed decisions provide some degree of decentralization for certain types of well-defined activities (i.e. the programmatic tasks of workers have some conditional elements in them), whilst retaining a high degree of central coordination because only managers may alter the decision making rules. Non-programmed decisions require individuals to act within their authority using their own judgement. As discussed later in this section, some organization theorists have argued in favour of decentralization because of its benefits in terms of job enrichment and because participative management is considered to be more ethical than highly mechanistic forms of social organization.

The distinction between programmed and unprogrammed decisions has been expanded by some researchers, who have described organizations in terms of their levels of formalization (for example, Pearce and David 1983). A highly formalized organization is likely to rely more upon programmed decisions, particularly at lower levels in the organizational hierarchy. Furthermore, unprogrammed decisions are likely to be well-documented and communicated via formal communication channels. In contrast, informal organizations rely less upon official reporting structures and unity of command, exhibiting a greater dependency upon organizational culture to ensure that workers make decisions that conform with the basic rationale of the organization.

In some respects, the formalization of an organization indicates the general nature of its coordination mechanisms. Analyzing coordination in more detail, reporting and authority arrangements usually reflect the relative importance of one or two of the following:

- business functions (e.g. finance and marketing)
- products or services
- customers
- geographical regions

The choice of organizing arrangements has a signficant impact upon the type of communication channels and authority relations required for the organization to function effectively. Evidence also suggests that such arrangements affect the development of expertise and the degree of innovation within the organization (Moorhead and Griffin 1992). The use of business functions, for example, is usually effective for developing expertise in the functional areas; reinforced by the institutionalized professions within developed countries (Mintzberg 1991). The effectiveness of functional organization for exploiting expertise is, however, somewhat limited because there exists no automatic flow *through* the organization to ensure that skills are used to create value for customers. In contrast, organization by product or service provides a clear process by which single products or groups of products/services can be provided. This type of specialization can inhibit innovation, however, because opportunities for transferring best practice from one product/service to another are more likely to be overlooked. Furthermore, organization by product can expose major customers to unnecessary complications and inconsistencies because they have multiple points of contact with the organization. Similarly, geographical organization can inhibit the spread of best practice and is only likely to benefit customers that operate in specific regions, possibly creating problems for geographically dispersed firms. Thus, whilst some form of organization is essential to retain control over an organization, the choice of structure must be made with careful attention paid to the organization's customer base and product/service characteristics. Furthermore, information systems are clearly required to provide communication channels that reinforce the strengths of the adopted structure and addresses its limitations.

In addition to the findings about particular types of structural arrangement, a conclusion of many researchers is that organization structure needs to balance authority and responsibility (Miner 1982). Authority is the power to act within an organization that is legitimized by its attachment to an organizational role. Responsibility refers to the association of rewards and sanctions with the outcomes of particular kinds of action. Ensuring that authority and responsibility are well aligned is important in several respects. Both motivational and ethical considerations suggest that praise and blame are best given to the person(s) who carried out a course of action. The ability to assert effective control over the organization also requires that the two are aligned. Alignment can be difficult, however, where efficient production is best achieved using group organizations, for example, but the society in which the organization exists tends to assign responsibility to individuals.

Whilst the above comments suggest that authority arises in a top-down manner within an organization, its existence in a wider social context, combined with the ethical and motivational issues, suggest that this is not entirely the case. Barnard (1938) was one of the first organizational theorists to propose the individual worker as a *source* of authority within the organization. Barnard's acceptance theory argues that the manager's authority depends upon the subordinate's acceptance of the manager's right to issue directives. Whilst this theory has some limitations, it rightly points out that the labour contract does not ensure absolute compliance, but only ensures that the worker will act in accordance with those directives that seem reasonable with respect to his or her terms of employment. As noted earlier, the employment relationship is somewhat exploitative. Regulation and the power of institutions advocating workers, however, provide the worker with some protection.

4.4.2 The Organizational Environment

The structural aspects of an organizational form are a product of both internal and external changes (however the organization's boundary is defined). Thus, in addition to considering the social activity that directly constitutes any observed social 'structure', it is essential to gain an appreciation of the external pressures and opportunities for action that motivate at least some internal activity. Traditionally, such issues have been described in terms of the organization's environment.

Given that any social structures are interpretations of the regularized activities of many individuals, it is not surprising to find that social institutions overlap because a single actor will participate in a number of social institutions and 'belong' to several organizational forms. Thus, as already noted, an individual worker is simultaneously a participant in a wider society that has implications in terms of allocating authority and responsibility to individuals and groups. The organization theory literature has paid particular attention to the following social factors and groupings:

- human resources - trade unions, labour markets, employment agencies, education systems
- competitors and suppliers
- customers
- shareholders (who may be internal or external to the organization)
- physical resources
- government - laws, taxation
- financial institutions - banks, venture capital

Given that the environment in its broadest sense refers to all activity that falls outside of the organizational form's boundary, it is both highly complex and subject to continuous change. This has two main implications: (a) there is a need for intelligence gathering so that the actors in the organization are aware of how the environment is changing; and (b) on some occasions, actions of organization members will need to change in response to changes in the environment.

Structural contingency models of organizations (the first of which was Lawrence and Lorsch 1969) suggests that an organization should attempt to match its structure with its environment. Given the definition of organization adopted here and the potential breadth of the environment, the plausibility of achieving such a match is open to question. Firstly, achieving complete information about a large and dynamic environment is extremely costly, if possible at all. Secondly, the amount of information processing effort required to determine the best course of action would be extremely high (Simon 1982). Thirdly, it is questionable whether a completely adaptive organization could be regarded as an organizational form because the highly routinized practices that constitute an organizational form imply some form of continuity of activity and, therefore, some (not necessarily intentional) resistance to change.

A second type of model of the relationship between organizational form and environment emphasizes the dependency of the organizational form upon the environment as a source of labour, raw materials and other resources. These models avoid the third weakness of the structural contingency models by describing two types of resource exchanges: interorganizational exchanges and making changes to the environment (Daft 1995). Attention to interorganizational linkages clearly emphasizes the identity of organizational form as an identifiable structure, whilst also acknowledging that its very existence is dependent upon a wider social organization. Recognition that organizational forms serve as parts of the changing environments of other organizational forms is also a strength of the model, making social organization appear less fatalistic than the structural contingency model suggests. Furthermore, the suggestion that it is sometimes 'easier' or 'preferable' for an organizational form to exert pressure for change rather than changing itself is suggestive of the resistance to change implied by routinized working practices and established social norms and also provides a means for actions where no amount of internal change could counteract adverse environmental change (for example, where legislation threatens the commercial viability of an enterprise).

4.5 Economic Models of Intra-Firm Activity

Within the firm, information systems have two significant economic impacts. First, information systems can facilitate effective resource allocation and usage as part of the firm's ongoing business. Second, information systems consume a firm's resources and must, therefore, be subject to evaluation. The main obstacle to effective evaluation is that, as indicated by these points, many of the costs of information systems are direct, whereas most of the benefits and a proportion of the costs arise from changes in the way other business activities are performed. Given this scenario, a sound and detailed economic model of the firm is required to contribute to the analysis of the impacts of IS changes and as the basis for evaluation methodology. Furthermore, if such analysis is to be consistent with analyses of user and non-economic considerations, the basic assumptions made by economic theories must be closely examined.

The models of the firm used in most, if not all, IS evaluation theories are limited in significant respects by their conformance with the assumptions of the conventional economic theory of the firm. One exception to this is Return on Management (Strassmann 1990), the limitations of which are described by Hemingway (1997a). The conventional model of the firm is really a set of simplistic assumptions about firm behaviour that provide sufficient detail to model some aspects of entire markets (Cyert and March 1992). Its use as the basis for evaluation and investment decisions, consequently, tends to be somewhat fatalistic, favouring labour-saving technologies as a means of profit maximization. More significantly, from an IS perspective, the implicit assumption that the firm operates using complete information about resource allocation and the means of production precludes the analysis of information systems changes. As a result, many investment appraisal techniques only function effectively when evaluating investments that provide primarily efficiency gains, particularly those attributable to automation (Farbey, Land and Targett 1994). Nevertheless, these simple quantitative techniques have remained more popular with IS practitioners than alternatives that take into account social and political factors by way of qualitative analysis (Ballantine, Galliers and Stray 1994).

4.5.1 Development of Economic Theories of the Firm

A milestone in the understanding of the economic functions of firms is *An Inquiry into the Nature and Causes of the Wealth of Nations* by Smith (1776), where the division of labour and other key aspects of work organization were explained for the first time. The focus upon capital as the primary source of power (Marx 1867) and the importance of the simplification and automation of tasks to provide efficiency gains have had a lasting impact upon much financial and management practice, particularly in Western economies. Consistent with the notion of social structures as self-reproducing, these impacts have tended to reinforce Smith's conception of the division of labour in four ways:

- an increased specialisation of tasks leads to a reduction in the individual worker's skills and a corresponding increase in manual dexterity;
- the reduction in the number of skills an individual requires leads to a reduction in the amount of time spent switching between tasks;
- increased specialisation and division leads to the precise specification of many tasks and, consequently, it becomes possible to develop machinery to perform some of these tasks more efficiently; and
- individual workers are no longer aware of the overall process of transformation that leads to economic value, leading to the need for specialist coordinators and a reliance upon the unconscious social reproduction of institutionalised activities in order to maintain the firm as a viable means of production.

Furthermore, the assumption of complete information, the existence of sets of unambiguous preferences for decision making, and full knowledge of all available means still have significant influence over management practice, despite being repeatedly questioned by economic theorists (Marx 1867, Coase 1937, Hayek 1945, Simon 1951, Alchian and Demsetz 1972, Cyert and March 1992). Despite the obvious problems these assumptions create for the study of information systems, the IS community has not yet succeeded in formulating a more realistic basis for IS evaluation. These issues are addressed here through the analysis of three factors:

- the nature of evaluative information;
- the nature of economic institutions; and
- the role of technology in organizational forms.

4.5.2 The Nature of Evaluative Information

The main purpose of evaluative information is to facilitate choice between costly prospective actions. As indicated above, the evaluation of information systems is unique because the provision of evaluative information is one of the key benefits of the system itself. Until the second half of the twentieth century, however, relatively little attention was paid by economists to information, presumably because the assumption that all relevant information is available implies that information acquisition and management are of little interest. Empirical evidence demonstrates, however, that the information available for decision making within an organization is widely dispersed and the collation of information pertinent to any given decision is severely limited (Travica 1998). Consequently, the modelling and analysis of a firm's information system must be regarded as critical to evaluating and improving its performance.

Evaluative information can be regarded from an economic perspective as describing opportunities for future resource usage. Considering the firm as a whole, such information can be partitioned according to the three classes of resources: (a) money; (b) labour; and (c) equipment. As already noted, acquiring evaluative information is a costly process and the amount of information acquired and used will, consequently, be limited by the expected benefits the information will yield in terms of improved resource management. The search mechanism used to identify useful information is also a key factor in determining what information is acquired and, consequently, how decisions are made (Simon 1967). The search will typically be based upon the knowledge of individuals occupying decision making roles, with their effectiveness being determined by a number of factors, including access to information and familiarity with, and availability of, information sources. Other determinants of the search and acquisition processes are the information users' interests, motivations and responsibilities, which were assumed by conventional economic theories to equate with profit maximisation for the firm. This assumption has been demonstrated to be inadequate, failing to reflect the impacts of personal motivations and values on the performance of workers (Cyert and March 1992).

It is evident that, due to the many different sources of uncertainty for actors, evaluation is always subject to severe limitations. It is also quite clear that the way in which information is collected, presented and used has a considerable bearing upon the evaluation outcome and any subsequent actions (Tufte 1983, Tufte 1990, Tufte 1997). In terms of information systems evaluation, the sheer diversity of functions that ICTs can perform makes any anticipation of the effects of IS changes even more uncertain. All of these factors are closely related to the nature of the planned action, the type of technology being considered and a number of other factors. Farbey et al. (1995) propose a benefits evaluation ladder as a mechanism for classifying project types in such a way as to draw out some of these issues. They go on to suggest that the evaluation of the different types of systems, because of the different sources of uncertainty that they face, should be conducted in different ways.

As a final point, it is interesting to note that uncertainty regarding economic performance increases as greater flexibility is achieved through the use of advanced technologies and decentralised working arrangements. Decentralised decision making, for example, reduces the level of control that any individual has. Clearly, such flexibility is desirable to some degree. A more effective approach to evaluation would be one that considers the interrelationships between the technological artefacts and persons using information to improve resource allocation within the firm. Such an approach requires an understanding of the second of the factors identified earlier: the nature of economic institutions.

4.5.3 The Nature of Economic Institutions

Having noted the serious deficiencies in the conventional economic model of the firm, Cyert and March (1992) propose a more detailed explanation of the processes of resource allocation, based upon the assumption that firms are coalitions. Although the model is pertinent to exclusively economic issues, its detachment from the political and, most significantly, the institutional aspects of organizations is limiting when considering the wider impacts of change. Analysis of the after-effects of a change and the development of new procedures and norms is, therefore, not well supported. This is a serious deficiency from the point of view of IS evaluation methodology, as evidenced by the high rates of post-implementation failure, where information systems have to be withdrawn because part or all of the organization is unable to settle into new, stable forms of work organization (Zuboff 1988). For this reason, both negotiation-based views of change and social theoretic models of institutionalisation are considered here.

4.5.3.1 Negotiation in Organizations

Actual behaviour in the day-to-day running of a firm is complex and varied, with most agreements, whether reached by consensus or compromise, being achieved on an informal basis. The nature of the behaviour leading to agreement may vary from hostile and antagonistic to co-operative. The institutionalization of working practices and social roles, however, provides an overwhelming precedent and directed constraint upon the vast majority of negotiation and decision making.

Negotiations take place at a variety of levels within an organization. Persons within a work group, for example, negotiate minor changes in their working norms; work groups within a department adjust their boundaries of responsibility; and departments negotiate arrangements for exceptions to their accepted policies, protocols and procedures. Many of these agreements tend to be co-operative because of their recurring nature and because agreement must be achieved within previously negotiated and more far-reaching commitments (Raiffa 1982), such as budgets and authority relations. Such negotiations, because they are between persons who work together on a frequent basis and who cannot trivially cease to interact, are highly interrelated and depend upon a measure of goodwill. Where agreements cannot be reached, there usually exists a 'mediator with clout' (Raiffa 1982), in the form of a more senior manager, who can quickly achieve a settlement even if it is, to a degree, imposed. It is into such a social context that computer-based information systems are introduced and the effects upon established relations can be dramatic. Thus, the economic issues of resource allocation, the political issues of negotiating agreements and the structural issues of conformance with established procedures and precedent within the organization must all be considered as part of an evaluation of the wider impacts of an information system.

With more significant, long-term negotiations, such as strategy formulation, agreeing annual budgets and evaluating large capital investments, the nature of negotiations is somewhat different. Firstly, because there is usually a large number of stakeholders involved, the negotiation process is more formal and often centres around a 'negotiation text,' such as a strategy document or list of evaluation criteria. As shown in section 3.10, the use of explicit representations significantly alters the nature of human problem solving, both for individuals and for groups. Because the agreement has significant long-term consequences and will enable or constrain many future activities, the negotations will be much more competitive, with interest groups seeking to maximise their future flexibility, autonomy and discretion. Most importantly, as Cyert and March (1992) note, the members will have different views of what the key objectives are and how they should be attained. Such differences arise from the different responsibilities and interests that the members have and are, therefore, unlikely to be reconciled as often attempted with dialectic approaches to evaluation (Guba and Lincoln 1989, Avgerou 1995). These differences in viewpoint are largely a result of institutionalized working practices, which arise from a combination of the stability of the organizational form and the considerable uniformity of trades and professions in Western economies.

The breadth in scope of the negotiation issues, combined with the abstraction required to establish common ground between parties with different viewpoints, results in high-level negotiations, the results of which are inevitably ambiguous. Although ambiguity can, itself, become an obstacle to reaching agreement, it also gives considerable flexibility in coming to some form of agreement. Indeed, it is often sufficient to agree upon broad principles within which all parties can function with a relatively low degree of conflict. Further complexities may arise, however, if representatives of parties to the negotiation need to ratify agreements with the stakeholder groups that they represent (the involvement of trade unions is an example of the need for ratification in a firm). In such cases, a 'network' of negotiations takes place, with representatives making provisional agreements, negotiating changes to these with their constituents and then renegotiating the original agreement. A technique frequently employed to ease such procedures is 'creative obfuscation' (Raiffa 1982), where agreements are phrased in deliberately ambiguous language so as to allow representatives to present different interpretations of the agreement to their stakeholder groups to ease the ratification process. Such activity has the significant implication that group representatives are not acting solely in the interests of their groups. This makes the assumption that all members of the entire organization are acting in unison appear extremely oversimplified. Such practices also leave the possibility for numerous conflicts in day-to-day operations and the scope for confusion over responsibility and authority, which can lead to political tensions within the organization.

A final negotiation based issue, of considerable relevance to IS evaluation, is the negotiation that takes place between the organization and providers of hardware, software and expertise. Raiffa (1982) observes that the balance of power in agreements involving sunk costs changes as contractual negotiations progress over time. Thus, for example, a firm evaluating different network architectures is initially in a strong bargaining position because it can readily take its business elsewhere. Many vendors are, therefore, keen to offer good deals in order to attract initial sales. Once a purchase has been made, however, the purchasing firm has sunk costs and will face increasing switching costs as the facilities are used and expertise in their usage is acquired in-house (Lacity and Hirschheim 1994). The vendor is now in a stronger negotiating position for future negotiations relating to services, technical support and updated hardware and software. Reports in the IS literature suggest that early outsourcing agreements were prone to the ruthless exploitation of such changes in the balance of power (Fitzgerald and Willcocks 1994).

Summarily, it can be seen that, whilst evaluation may be an ongoing and, sometimes, formalised negotiation process, it often yields vague outcomes, which are subject to a variety of interpretations. In many cases, such ambiguity, contrary to discussion in the IS literature, is largely functional, reducing the likelihood that negotiations will break down or interfere with the day-to-day operations upon which the survival of the organization depends. Thus, it appears that evaluation methodologies will benefit from recognition of the various levels of negotiation and the variety of consensus and compromise that is required to change an organization whilst maintaining the continuity of the basic operations that serve as its *raison d'être*.

4.5.3.2 Regularized Behaviour in Firms

Although negotiation explains much activity within organizations, a great deal of co-operative, and often unquestioned, activity maintains them. Some unquestioned activities result from the implementation of rules and procedures, whilst others arise from actions based upon precedent and past agreements, which are only clarified in novel situations, rather than subject to complete re-negotiation (Raiffa 1982). Thus, a description of the firm is grossly incomplete without some explication of co-operative behaviour and the institutionalisation of social processes.

In the theory of structuration (see section 4.3), social structures, rather than being regarded as objective entities, refer to the rules and resources that lead individuals to reproduce social behaviours. As such, regularization is basic to social functioning and, where it is stabilised across space and time, is constitutive of institutions. Regularization is, thus, central to the continued functioning of the firm. As already discussed, repetitive negotiations are typically quite co-operative, with levels of co-operation being very high where routine actions are concerned. Over time, many such actions are accepted as conventions within the organization, being taken as givens. Furthermore, where such actions require direction of resources, authority may be formalised and, consequently, the persons to whom it is allocated adopt roles within the firm. Some activities are also accepted, in a similar way, to become formalised procedures. Acceptance of these regularized practices is extremely important to organizations as it gives them a permanency beyond any individual's actions. Obviously, whilst crucial, institutionalization, by its very nature, inhibits radical change (Weick 1985). It is important, therefore, that the implication of all forms of negotiated change, from routine bargaining to strategic decisions, are seen in this light. Large scale changes, such as IS developments, will usually involve challenging at least some established practices and routines. A continual process of negotiation, in the form of change management, is essential at this stage.

4.5.3.3 Summary

The conception of the firm presented above briefly describes its activities as directed via a mixture of negotiated compromise and the unquestioned execution of previously established practices. Such a view is considerably different to the conventional economic theory of the firm and addresses some omissions of the behavioural theory developed by Cyert and March (1992). The above model also differs from the classical strategic view of organizations, because it rejects the idea that workers in the organization can and do perform their activities in accordance with explicit strategic plans. The proposed model also differs from consensus-based views of the firm in recognising that stakeholding is not necessarily synonymous with the holding of power. Thus, although it may be considered desirable to involve all stakeholders in the evaluation process, this may not be effective. The effectiveness of stakeholder involvement is also further complicated by the different levels of abstraction at which stakeholders (such as end-users, senior managers and shareholders) view the organization. Instead, this model recognises that any change in strategy, or change in operating procedures or technologies requires: (a) high-level negotiations between senior executives (and, perhaps, union or other representatives); (b) ratification of (or, at least no dissent from) the negotiated agreement; and (c) continued re-negotiation between those enacting the changes, eventually leading to a reformulation of the working norms. Thus, it is clear that the typical assumptions as to how the benefits of investment arise need to be expanded upon.

Whilst specialisation and automation may continue to be important aspects of effective work organization, they represent a narrow range of options for gaining benefits from investments. IS evaluation must, therefore, attempt to account for a more diverse array of approaches to organizational change and to manage the considerable uncertainty regarding the impacts of the IS on organizational practices. As a general purpose machine, the computer continues to provide greater potential for reorganizing work activity. The realization of this potential can only be fully exploited through the integrated design of technology and working practices in a manner that is sensitive to established organizational roles (Kelly 1982). This implies, however, considerable uncertainty regarding what changes will actually take place: the norms and regularized working practices that will become established following a large scale change, whilst subject to the influence of an 'organizational design,' cannot be predicted. In this regard, IS evaluation will always be limited.

4.5.4 The Role of Technology in Organizational Forms

Technology has often been regarded as a substitute for labour or a means for increasing labour efficiency. Even socio-technical theory (Mumford and Weir 1979, Mumford 1983, Mumford 1995), which aims to balance human and technical concerns within organizations, effectively assumes the same overall goal and, consequently, aims to moderate technological change in light of ethical and humanitarian goals. A brief consideration of why technological artefacts are used, however, suggests a radically different perspective ought to be adopted.

Technological artefacts, such as machines and furniture, are acquired by organizations for a number of reasons, including cost reduction, reduced headcount, increased productivity, retention of skilled workers and impressing customers. The key point to note is that technological artefacts are bought or developed for a purpose and the purpose is the primary determinant of their use. Purpose and intention are attributes of human action. Thus, considering a technological artefact as a mere substitute for labour is to obscure its purpose. Even where reduced headcount is the primary goal of the power group sanctioning the purchase of new machinery, for example, that group's intentions and goals result in the inanimate objects becoming part of an organized system. Technology, then, must be regarded as the purposeful exploitation of artefacts. Without a purpose to provide a context for use, technological artefacts cannot be regarded as technologies. To illustrate this point, one needs only to consider the many improvisations made on a day-to-day basis, such as using a screwdriver as a lever or a chair in place of a stepladder. Chairs, out of context, are associated with sitting because they were constructed with this general purpose in mind. In the context of a person wanting to reach a high shelf, the technology is the exploitation of the seat of the chair as a stable, elevated platform. The mental creativity and physical skill of the user - not the artefact - is constitutive of the technology.

The main consequence of adopting this view of technology is that the preceding discussion of social regularities, rules, negotiation and other aspects of organizational forms *are* discussions about technology; the acquisition of technological artefacts being somewhat incidental. Thus, the primary economic issues relating to technology are as follows:

- **Information and Communication Skills** - How can individuals and groups best develop skills for translating, for example, their goals, political and environmental pressures into articulations of the organizational form's technological requirements?

- **Technological Skills** - How can individuals and groups best develop cognitive and practical skills for designing and developing networks of technological activities to satisfy their technological requirements?

- **Technological Artefact Evaluation** - How can individuals and groups best be guided in their selections of appropriate technological artefacts, given that they must often acquire such artefacts with very little knowledge of their future technological requirements?
- **Technology Skills** - How can individuals and groups develop skills for incorporating and adapting available technological artefacts effectively to realize their technologies?

Economics and information systems evaluation have typically focussed primarily upon the third of these issues. Although important, methods for selecting technological artefacts are bound to be limited if organization members are poor at identifying their own needs within the context of the requirements of the organization as a whole. Similarly, if needs cannot be translated into effective transformation processes to satisfy the demands placed upon the organization, then technological artefacts are unlikely to help. Indeed, evidence suggests that poor business processes are rarely improved, and sometimes worsened, by the introduction of new technologies (Davenport and Short 1990). Clearly, a theory of social organization must address all four factors and their interrelations if it is to serve as an adequate basis for developing methodologies for information systems and the development of information and communication technologies.

4.6 Organizational Change

The discussion of economic issues indicated the importance of both continuity and change, concluding with the suggestion that organizations rely upon negotiation, with formal and informal groups playing a major role in negotiating and realising change. Given the difficulties usually experienced in introducing technological changes (DeLisi 1990, Kirveennummi, Hirvo and Eriksson 1999), it seems particularly important to conclude this study of social organization with a consideration of the nature of social change within organizational forms and how such change can be managed. The discussion begins by identifying the implications of regarding an organization as a focused, but dynamic network of social interaction, before considering how significant, innovative organizational change can be achieved.

4.6.1 The Adaptive Organizational Form

Some theories of organization development and change management have effectively regarded organizations as discrete entities residing within an environment (see, for examples, the McKinsey 7-S Framework, Peters and Waterman 1982, and the cultural web, Johnson and Scholes 1993). In the 1980s, these organizations were recognised as 'having cultures.' Consequently, many theorists included a 'culture box' in their model of the organization. As argued above, the notion of the organization as a discrete entity is misleading. It is more appropriate to describe organizational forms as complex networks of social interaction. Social interactions are greatly influenced by precedent and routine, which enables observers to identify the intense and regularized social interactions as an organizational form. Such forms are not concrete entities, however, but continually evolving social networks; any reference to 'the organization' being primarily a linguistic convenience. Taking this view, as Bate (1994) and others have done, an organization *is* a culture, not an entity with cultural attributes. It is distinguished from the wider culture in which it is embedded by virtue of its organization. Consequently, the management of organizational changes, such as information systems developments, must be regarded as a form of cultural change. This, of course, raises the question of how to manage an organizational culture. Clearly, many current approaches are inadequate because they encourage the alignment of culture and, for example, strategy or technology. This is only meaningful if culture is regarded as a component of an organization and not synonymous with it. Alternative approaches to change management must, therefore, be explored. Bate (1994) provides an indication of how an alternative approach might be developed:

> "...culture exists not so much 'inside' or 'outside' people as 'between' people. This conception helps us to see a cultural change effort as a form of social

> intervention aimed at altering the quality of 'between-ness' in some way or other." Bate (1994, p.15)

Owing to the dynamic nature of social interaction, actors are always acting in novel situations. Consequently, the performance of regularized practices implies their continual evolution in response to the 'internal' and 'external' pressures. Thus, it is instructive when considering organizational change to assess the present state of the organization, how it came into being and how it is adapting to present circumstances. The development of such an analysis effectively provides a diagnosis prior to the planning of any 'treatment' in terms of information systems development or other organizational change.

As illustrated by the discussion of economic issues, organizational forms are far from being homogeneous. At any given time, various individuals and groups will be relying upon established norms, negotiating compromises, enforcing or being subject to enforced solutions, and so on. An appreciation of organizational culture, therefore, requires an understanding of the many interacting subcultures it contains. This need is reinforced by the observation that attempts to enforce a culture within an organizational form usually result in the development of strong counter-cultures (Gabriel 1991). It is for this reason that Bate (1994) and others recommend that cultural change proceeds on the basis of accepting and understanding pluralist social organization. In terms of developing information systems, the main implication of a pluralist outlook is the need to accommodate the different languages that each sub-culture uses. Attempts to develop monolithic information systems, which use a uniform user interface, are likely to result in the emergence of an increasing number of unsanctioned communication channels between individuals within workgroups, because such workgroups are, in effect, different language communities. The modelling of communication flows between individuals and workgroups is, therefore, a key component of effective information systems development.

Whereas strict uniformity is likely to be rebelled against, too much diversity will result in a lack of direction that can lead to the fragmentation and possible dissolution of the organizational form. Some degree of control needs to be asserted and power relations must, therefore, be included in the analysis underpinning IS and other organizational developments. This is particularly important in established organizations because the institutionalization of power bases will typically present considerable resistance to change (Markus 1983, Cavaye and Christiansen 1996). In terms of exerting control within an organization, Bate (1994) identifies the following options:

- manipulating the ideological premises of action
- implementing programmes and procedures
- issuing orders and ensuring compliance

Control exerted through ideology is potentially the most powerful form of culture 'management.' It involves the manipulation and propagation of symbols and stories within the organization as means of defining what is and is not acceptable to a particular power group. The use of programmes and procedures to exert control is the most widely discussed aspect of cultural/organizational management, referring to the explicit, often documented, aspects of working practices, exemplified by the machine bureaucracy (Weber 1947). Whilst such explicit rules form a significant component of an organizational form, they cannot cover all eventualities and their rigid imposition is likely to be counterproductive, frequently resulting in their circumvention (see, for example, Zuboff 1988). Such rules and procedures do, however, play an important legitimizing role within an organization, making clearer the distinction between the actual power of actors and their authority.

The final mechanism for control - the use of power to ensure compliance - is the most likely to have adverse consequences for the organization, usually only having its intended effects over a very short period. Whilst compliance may well be achievable within some organizational rules and procedures, it often goes beyond them, bringing into question the legitimacy of the power holder's demands of other actors. The acceptance theory of authority (Barnard 1938) highlights an important point in this respect: labour contracts are grossly incomplete and much of the activity within an organization is performed on the basis of workers' acceptance of organizational norms, precedent and their superiors' established authority. The excessive use of 'brute force' to ensure compliance is, consequently, a risky strategy for change because it increases the likelihood that the institutionalised practices that lend legitimacy to an actor's demands will be brought into question. This risks destabilizing the actor's sources of power, if not the entire organization.

4.6.2 Large-Scale Organizational Change

Innovation in an organization is inevitably risky, involving the replacement of functioning working practices with others that will initially be problematic and are not guaranteed to function over the longer term. This risk is balanced, however, by the likelihood that changes within or without the organization will eventually lead to existing practices becoming dysfunctional or less relevant to the individuals and social organizations placing demands upon and/or satisfying the needs of the organization. Consequently, change is unavoidable if an organization is to survive over the long-term. Whilst evolutionary change of the type discussed above may often be sufficient, it sometimes is ineffective because habituation at the individual level (leading to institutionalization at the organizational level) implies that core activities often go unquestioned (Bate 1994). If such activities lose their relevance and the organization's performance suffers, the psychological models presented in chapter 3 suggest that actors may attribute these problems to external factors or feel anxious about their working environment without knowing why. If explicit indicators, such as profitability, demonstrate that the organization is experiencing difficulties, it effectively enters a period of crisis. Crisis situations are often a prerequisite to the recognition of the needs for, and the feasibility of, major organizational change. The difficulty of successfully implementing large-scale change is indicated by one possible response to a crisis situation: the so-called 'escalation of commitment' (Brockner 1992).

Given that working within a culture is a key reason why actors fail to see the cause of their problems, an important factor in stimulating cultural change is making assumptions explicit and creating a feeling of dissatisfaction with the *status quo*. It was suggested above that gaining an understanding of the history of an organization and its current adaptation to internal and external pressures were important in understanding organizational change. A useful next stage in an analysis is to identify current sources of dissatisfaction and any actors or groups with aims that are not achievable given the current state of the organization. Such analysis must address both the actors' personal aims and objectives and their views of the organizational needs. The resulting analysis will provide an indication of the different visions of the actors who can ease or inhibit organizational change. Coupled with a model of the power relations, communication channels and internal and external factors contributing to the organization's development, sufficient information is available to anticipate the directions in which the organization is most likely to develop and how these outcomes can be facilitated or obstructed. Communicating this information (especially selectively) may increase the chances of organizational change by encouraging actors to reflect upon taken for granted actions and their outcomes. One advantage of the approach to change management outlined here is that it anticipates likely sources of political conflict and will provide a broad indication of the relative power of advocates of the *status quo* and advocates of change. It also raises a number of serious ethical concerns, which are discussed in chapter 6.

4.7 Summary

This chapter has considered the nature of social interaction from sociological and economic perspectives, comparing the assumptions made about individuals' actions with the general characteristics of individuals as conceived in psychology. The analysis has shown that there are a number of substantive theoretical issues that must be carefully addressed if a theory of social (including economic) organization is to provide an understanding of information systems. Indeed, as illustrated by the use of conventional economic theory as the basis for IS evlauation, basic assumptions can sometimes preclude the identification of those aspects of social activity most relevant to planning and undertaking change.

In summary, the development of an integrated theory of information systems must address a number of issues where the analysis of the individual and the analysis of social organization overlap (as identified in the first list below). Having addressed these integration issues, the broader issues of social organization and organizational change must be addressed (as identified in the second and third lists below).

- **An account of the relationship between the individual actor and the wider social organization must address :**
  - the distinction between behaviour and action;
  - experiential learning and its role in constituting social activity;
  - the extent to which the continuous experience of the individual is made discrete by cognitive factors and the episodic nature of social encounters;
  - the extent to which routinized activities are treated by the individual as episodes;
  - the role of causal attribution in problem solving and mental representation;
  - the nature of personal preferences;
  - the ways in which individuals search for and process information when performing activities assigned to them via an organizational role;
  - how an awareness of normally routinized/automatic behaviours and their implications can be stimulated;
  - anxiety and causal attribution.
- **An account of social organization must address:**
  - the roles of value judgements and rules in social organization;
  - the role of physical locales in reinforcing social institutions;
  - the development of skills within an organization, including, but not restricted to, an account of the division of labour;
  - the ontological status and general characteristics of social structures and other aspects of social organization;

- the effects of programmatic and discretionary autonomy, authority and responsibility distribution upon actor commitment and satisfaction, and the consequences for supervision and coordination;
- the different power relations and communication channels that can be developed within organizational forms, and their appropriateness for different types of organizational activity;
- the ways in which different intelligence gathering (including interorganizational linkages) and communication arrangements affects the abilities of actors to adapt their behaviours to accommodate changes in the environment;
- the relationship between the arrangement of communication channels and access to them and the efficiency of resource allocation and usage within an organization;
- the differentiation between environmental uncertainty, uncertainty arising from maladaptive organization, and uncertainty arising from the flexibility of the organization to adapt to its environment;
- the nature of formal and informal negotiations, the impacts of institutionalisation and precedent on negotiation, and the roles of power holders as 'mediators with clout';
- the use of negotiation texts as the basis for negotiation in organizations and the impacts of their substitution for ICT artefacts;
- the objectives of power groups and their representatives during negotiation activity;
- the changing balance of power during internal and external negotiations;
- technology as the purposeful exploitation of artefacts.

- **An account of organizational change must address:**
  - the institutionalisation of rules, procedures and roles;
  - the dynamics of subcultures within an organization;

– the significance of ideology and symbolism, procedures and rules, and power and authority in realising organizational change.

Chapter 5 uses the above summary of social and organizational issues and the summary from chapter 3 as the basis for developing a theory that aims to address some of the problems identified so far by this thesis.

Chapter 5

The Socio-Cognitive Theory of Information Systems

5.1 Introduction

Chapters 1 and 2 of this thesis identified a number of limitations of current information systems theory and practice. Chapters 3 and 4 have examined the literature of several reference disciplines to determine the nature and complexity of these issues, and to indicate ways in which they might be addressed. One of the points made in chapter 2 was that several levels of social analysis are required to understand any information system. Drawing piecemeal upon different reference disciplines, such as psychology, sociology and economics is an inadequate way of achieving this because inconsistencies between models from one discipline and the basic assumptions of another may arise. An integrated perspective is required to provide a comprehensive account of IS phenomena and to provide a rigorous basis for the development of practical methods and techniques. The purpose of this chapter is to present the socio-cognitive theory of information systems as a first step towards providing such an account.

In developing information systems theory, it is instructive to note the correspondence between the practical need for explanations that are "sufficient to serve as the basis for purposeful and effective intervention" (section 1.3) and the assertion in section 1.5 that the capacity to manipulate phenomena provides a substantial component of the justification for including them in a theory. Combining this observation with the other key points from the above discussion - theory is based upon the interpretation of evidence, is limited by the precision of categorization and does not have a truth value - strongly suggests that the development of a substantive theory of information systems will inevitably be a dialectical process involving conceptual development, empirical research and the application of the theory to address practical IS issues.

Given the various conceptual issues identified in the preceding chapters, conceptual development is chosen here as a starting point for developing a socio-cognitive theory of information systems. The following sections describe the phenomena of interest, making explicit the ontological commitments made to each type of entity and the causal relations defining each type of event. Given that reference must be made to the goals and intentions of actors, the epistemological issues relating to value judgements are also addressed. The practical implications of the theory for IS and ICT design are then considered in chapter 6.

5.2 Communication

Events in which communication takes place are, in part, constitutive of information systems. Communication is the encoding of a message by a sender, the transmission of the encoded message via a communication vehicle and the decoding of the message by a recipient. Encoding is the embodiment of a message into a series of signals that, as a whole, physically instantiate the message in a communication medium, resulting in a communication artefact. The communication medium and the communication vehicle may or may not be identical. Where medium and communication vehicle are physically distinct,

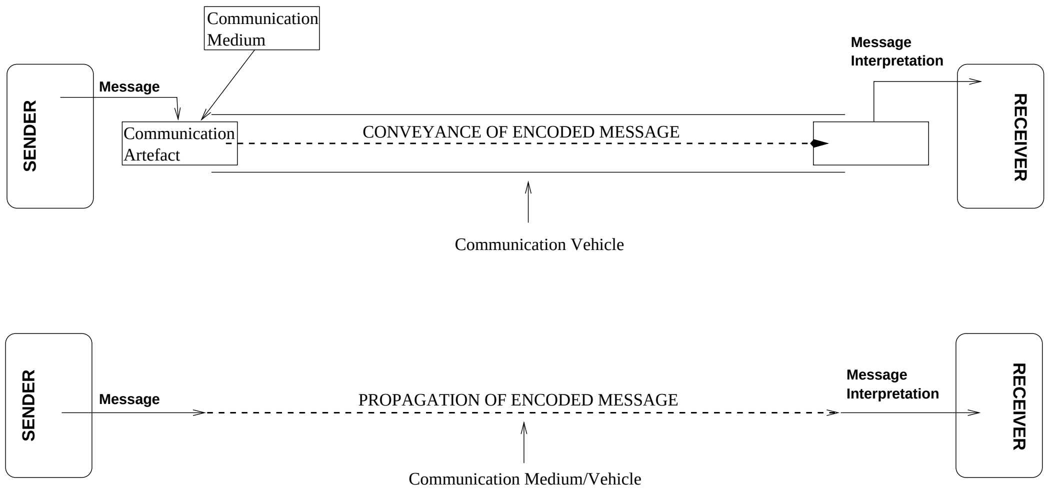


Figure 5.1: Two Types of Communication

transformation is termed conveyance. Where medium and vehicle are identical, transmission is termed propagation. Transmission is the movement of communication media from one physical location to another. Decoding is, ideally, the reverse of the encoding process. The various components of communication are shown in figure 5.1. As shown in the figure, encoding, decoding and transmission all involve technology artefacts. It should be noted that all entities referred to above are concrete particulars, with the exception of the message.

5.3 The Message

To give the concept of communication its full meaning, the concept of message must be defined. Unlike the other components of communication, however, the message is an abstract concept that is difficult to define in a precise and uncontentious manner. The realist ontology of this thesis suggests that an appropriate starting point for defining 'message' is the relationship between the message and the concrete elements of communication. Thus, an initial attempt might be *whatever is encoded and transmitted as a series of signals.* This definition has the advantage of referring only to concrete entities but it is essentially circular, shedding no light on the attributes of messages that make communication artefacts distinguishable from other concrete entities (for example, what features of a memorandum distinguish it from a piece of paper with meaningless marks on it).

More progress can be made in describing messages by referring to *some kind of organization*. A memorandum, for example, is distinguished from any other piece of paper not by the marks on the paper itself but by virtue of the arrangement of those marks. Organization is clearly an important attribute of a message, as illustrated by the fact that *equivalent* messages can be encoded into different media, resulting in different communication artefacts (for example, the memorandum could be transcribed onto audiotape). Specifically, a sender will arrange properties of a communication medium to signify something. The variety of encoding mechanisms raises a further problem, however: there is nothing in this model of communication to indicate that the receiver will decode the signals to yield the signification encoded by the sender. The organization that was encoded must have been derived from an organization within the sender but this only accounts for a sender being able to decode his or her own messages. When the message is transmitted from one subject to another, communication implies some form of *intersubjective organization*. Given that sender and receiver are distinct entities, there are two possible sources of intersubjective organization:

- the sender and receiver are somehow ‘constructed’ so that they contain sufficiently similar organizations to be able to encode and decode each other’s messages; and/or
- communication itself results in sender and receiver establishing sufficiently similar organizations to be able to encode and decode each other’s messages.

An important implication of regarding messages as being constructed upon the basis of intersubjective organization is that the calculations of information theory (see section 2.2.1) are limited definitions of efficiency because they only consider the efficiency of message transmission. One might, for example, use the calculation shown in figure 2.2 to derive the most efficient alphabet for transmitting a given class of messages. The alphabet is only the most efficient for communication, however, where encoding, decoding and message interpretation are perfect. Where, as in the case of much human communication, various types of redundancy reduce errors in encoding/decoding or interpretation, the most efficient communication mechanism may well not involve the most efficient transmission mechanism. An holistic approach must be taken to analysing and developing the communication system. Given that the sender and receiver in human communication are both individuals, intersubjective organization is considered from a psychological and then social perspective in the following sections.

5.4 The Individual

To understand human communication, it is necessary to explain the underlying intersubjective organization in terms of similarity of physical construction and similarities arising from interaction between individuals. Clearly, such an explanation will have both social and cognitive components and be firmly grounded in human physiology. As illustrated in chapter 2, distinctions between explanations of each component create difficulties for information systems design and evaluation. Consequently, an integrated explanation is provided in this section as the basis for an integrated theory of information systems.

5.4.1 Sensory Stimulation

It is evident from section 3.2 that the senses are highly selective in what they extract from the environment and that sensory stimuli are subject to considerable processing prior to perception. Evidence shows that the properties of the objective sensory physiology are usually very similar across individuals and the early stages of perception often correspond closely with the objective stimuli (Dudel 1986). Considerable variation does result, however, from disease, congenital and hereditary disorders and environmental factors (i.e. injury and adaptation). Such factors can have major impacts upon an individual's ability to use different types of communication and technology artefacts.

Attribute	Range	Design Implications
Frequency	Usually sensitive to 20Hz to 16kHz range; most sensitive to 2000-5000Hz range; distinguish changes of around 3Hz; different frequencies result in different patterns of neuron activation.	Use of auditory signals can relate urgency of signal with sensitivity to frequency.
Sound Pressure Level (SPL)	Sounds at different frequencies become audible at different threshold SPLs. Frequency and SPL in combination determine the subjective impression of loudness (e.g. 4000Hz at 70dB SPL normally perceived as loud as 1000Hz at 80dB SPL).	Design of auditory signals must account for fact that constant power level for a speaker may be perceived as too loud or quiet depending on signal frequency.
Auditory Orientation in Space	Only possible with binaural hearing; under optimal conditions, can locate sound direction within accuracy of around 3 deg; discrimination of sounds from front and behind depend on shape of outer ear	Record stereophonic sound using an accurate model of the head, with microphones positioned in positions of the eardrums.
Working memory	Contains equivalent of approx. 2 seconds of normal speech	Provision of auditory information that must be retained during task performance should fall within this limit.

Table 5.1: The Characteristics of Audition and their Design Implications

A full account of sensory physiology is not provided here because reports of detailed empirical studies are available elsewhere (Schmidt 1986) and the functioning of objective sensory physiology raises few theoretical issues of interest to the development of information systems theory. The typical characteristics of the sensory modalities and the variation in performance are of interest, however, in terms of their design implications. Tables 5.1 (derived from Klinke 1986) and 5.2 summarize selections of these characteristics for audition and vision - the modalities of most interest to ICT artefact design. Although they do not yet play a major role in ICT artefact design, mechanoreception and proprioception are summarized in table 5.3 to reflect the fundamental role of the body and the self-concept in human communication.

Attribute	**Range**	**Design Implications**
Depth	Cues from binocular fusion, size, overlap, haze effects and parallactic shifts during head movements	Design of 2D displays must account for 3D interpretation by human visual system.
Flicker Fusion	Above 22Hz, apparent and real motion indistinguishable when changes take less than 120ms	Causal attributions may not be inferred if animation does not operate within these limits.
Tracking	50% of visual field per second maximum, fixation point drifts over time, reflex saccade caused by intrusion into peripheral vision	Constrains design of representations for tasks requiring visual tracking.
Colour	Normal vision is trichromatic, although some individuals unable to discriminate certain colours	Risk of misinterpretation of display can be reduced by colour usage.
Brightness	Local adaptation of retina regions leads to afterimages, light adaptation takes time	Very large variations in brightness on display or between display and environment may reduce task performance.

Table 5.2: The Characteristics of Vision and their Design Implications

5.4.2 Perception, Recognition and Recall

The individual is subject to a continuous stream of sensory stimuli. At higher levels of cognition, this continuity is reflected in the durée of subjective experience (see section 4.3.1). The discrete events normally considered by decision theories and evaluation methodologies are not experienced as discrete events but, because sensory stimuli are received from bounded regions of space, the individual's movement and the movement of entities in and out of the individual's sensory regions (see section 5.6.2) result in marked changes in the stimulation of one or more sensory modalities. An individual in a dark room, for example, will perceive sudden illumination of the room as an event. Similarly, an encounter with a person may be perceived as an event because the person's presence results in distinctive patterns of sensory stimuli. The discrimination of events from the durée of experience depends on both the physical location of the individual and his or her attention to sensory stimuli and interpretation of them with respect to his or her knowledge.

Attribute	**Range**	**Design Implications**
Mechanoreception	Sensitive to intensity, velocity (used for touch to determine texture, etc.), and acceleration (i.e. vibration); indentations of as little as 10um produce sensation, spatial resolution about 1-3mm on hand, greatest sensitivity (as low as 1um) to vibration in 150-300Hz range.	Parameters should be used to guide design of input devices, such as the keyboard and mouse.
Proprioception	Stimuli primarily from the joints, muscles and tendons facilitate awareness of the positioning of the body, direction and speed of movement of body, resistance to movement and judging the amount of force to apply to move objects.	Parameters should be used to guide design of input devices, such as the keyboard and mouse.
Nociceptors	Located at joints, nociceptors are excited by joint movements approaching working limits and, therefore, threatening physical damage.	Should be taken into account as health and safety considerations during ICT artefact development.
Sensory integration	Proprioception integrated with information about position of the head in a gravitational field and signals from motor system to give perception of body in space-time	Studies of sensory integration will contribute to the development of ICT artefacts for interacting through.

Table 5.3: The Characteristics of Mechanoreception and Proprioception and their Design Implications

Drawing upon the models of human memory (section 3.4), perception can be regarded as involving the activation of one or more functions of the brain by patterns of stimuli. What kinds of patterns are activated in memory will depend upon the nature of the stimuli and the knowledge held in long-term memory. The association of stimuli with patterns in memory has a number of implications. Similar stimuli will lead to similar activations in memory. Consequently, many similar stimuli will lead to the frequent activation of common components of the different stimuli. Given that frequent activation increases the probability of future activation, a central tendency effect results that provides a form of implicit categorization or prototyping. Particular instances will be remembered insofar as they have attributes that are distinctive from the central tendency. In this regard, distinguishing attributes and contextual factors play a significant role in recognition and recall. 'Categories' will be activated where partial stimuli are activated and no highly individuating stimuli are available.

The activation of patterns in long-term memory may occur via the working memory or as a direct result of the interaction between sensory and long-term memories. When encoding takes place via working memory, it is usually the case that the individual will be paying conscious attention to the stimuli. Where stimuli are subject to conscious attention, additional processing may take place, leading to a richer encoding of the memories than would otherwise occur. How rich this encoding will be depends upon the time available to the individual, which is affected by: (a) the rate of change of stimuli; (b) other task demands and interruptions; and (c) the number of sensory modalities that can be integrated and need to be integrated for immediate task performance. Moreover, semantic encoding requires more processing and is effective for a wider range of future retrieval situations than phonological encoding, which is generally more effective than visual encoding (Baddeley 1990). Visual stimuli can be translated into semantic representations directly or via phonological encoding, with the latter capability acquired with the development of language skills. Which type(s) of encoding are used for particular memories will be a significant determinant of how the encoded knowledge can be used for constructing representations in working memory. It is also likely that encoding mechanisms previously associated with a type of stimulus will be used again for encoding future experiences of the same type.

5.4.3 Representation

The discussion presented in chapter 3 indicates two possibly distinct forms of representation: the active representations used to guide actions; and the representations stored in long-term memory. Considerable debate has taken place regarding the nature of both types of representation (see sections 3.4 and 3.5.2). Investigating how memories are stored in long-term memory is extremely difficult because such representations are only accessible when they are 'active.' This is not a particular obstacle to developing a model of the individual for studying communication, however, because the focus of interest is the representation active in memory. Long-term memory functions are only of interest insofar as they relate to perception and contribute to the constitution of active representations. Considering active representations, two explanations have been proposed. Kosslyn (1980) and Marr (1982), for example, suggest that representations in memory are analogues of attributes of the environmental stimuli. A rival account, advocated by Pylyshyn (1981), maintains that representations are stored using an amodal form of encoding.

The distinction between analogue and propositional representations has often made reference to imagery and other issues. Of most interest in terms of communication and social interaction, however, are the implications of the accounts for the accuracy of representations and their use in reasoning and guiding action. As Pylyshyn (1981, p.26) points out, when a person manipulates a mental representation, the changes to the represented objects are "a function of what principles one believes govern the events in question." This claim is supported by evidence of the development of cognitive skills (see, for example, Smyth et al. 1994). Many children, for example, imagine that a ball rolling off the end of a table will continue in a straight line for a short distance and then drop vertically, whereas most adults (correctly) imagine a ball descending to the ground on a curved trajectory. The rules applied are a result of the interpretation of observations of, and interactions with, the world, supplemented by the explanations communicated by others.

When constructing a communication artefact, an individual must draw upon discursive and, at least to some degree, intersubjective knowledge when formulating a message. It is widely accepted, however, that one of the distinguishing characteristics of tacit knowledge is that it is, at best, only partially describable (Pylyshyn 1981, Anderson 1990). Indeed, Anderson (1990) proposed that tacit and discursive knowledge are stored in memory in different ways (see section 3.4.1). The dependency upon tacit, experientially acquired knowledge for constructing representations is likely to be highly significant in terms of understanding the impacts of ICT artefacts on reasoning and problem solving, although it has not been systematically explored. The use of virtual reality simulations in training, for example, may lead to the development of skills that exhibit systematic 'errors' (when the skills are applied out of the simulation environment) insofar as the explicit encoding of the simulation deviates from subjective experience. Another practical implication is the quite different representations that are constructed from determinate and indeterminate messages (Mani and Johnson-Laird 1982).

5.4.4 Synchronic Relations

In addition to the distinction between tacit and discursive knowledge, the distinction between knowledge of synchronic and diachronic relations also contributes to explaining social activity. Synchronic relations are the associations between the attributes and relations of objects in the environment that constitute the individual's knowledge about objects and object types. The development of such knowledge is achieved by the encoding of patterns into long-term memory, as outlined in section 5.4.2. In addition to storing details of the spatial properties of objects, other attributes are also encoded, including functional attributes and motivations active at the time when the object was experienced. The encoding of functional attributes is critical to the development of technology as a process, providing associations between types of tasks and objects, particularly constructed artefacts, located in the environment. Given that artefacts are constructed for particular functions, however, the functional specificity of the artefact also plays a role. Staplers, for example, were developed to address a specific function and this function usually determines how they are used as a technology (i.e. they are typically used for stapling papers together). Less conventional associations, however, lead individuals to use staplers as paperweights and pencils as pointing devices. Thus, when analyzing information systems, consideration should not be restricted to the functional attributes implied by the construction of technology artefacts. More importantly, perhaps, it is essential at the design process to avoid assuming that the functions motivating the design and construction of technology artefacts will be apparent to their prospective users, as illustrated by Norman's studies of everyday technology artefacts (Norman 1988).

5.4.5 Social Schemata

Although the basic cognitive processes involved in constructing knowledge about persons are, no doubt, essentially the same as those used for inanimate objects (section 5.4.4), there are a number of reasons why knowledge relating to persons should exhibit some significant differences. Some evidence of such differences was discussed in section 3.9.2. The first major difference is the subject's awareness that, like themselves, other persons act with intention. Consequently, a social event, such as a person smiling and saying 'hello,' will not simply lead to the causal attribution that seeing a person will generally cause another person to smile and say 'hello.' Instead, motives or personality traits will be inferred as reasons for intentional actions. When one person greets another, this may be, for example, because he or she is friendly (a dispositional trait) or knows the other person (i.e. he or she is motivated to maintain a valuable social relationship).

Imputing intentionality to the actions of individuals complicates the basic associative model presented above, as illustrated by the research of Kelley (1967) (see section 3.9.1). Task demands may require interaction with a known individual or group, or an unknown individual or group. The different demands can be broadly accommodated by the associative model because known individuals will be associated with knowledge from previous encounters, whereas available cues from unknown individuals will usually activate the 'central tendencies' common to many past social encounters, unless the individual has any particularly distinguishing attributes that activate more specific memories. When interacting with groups, there is evidence to suggest that general knowledge will again be used or salient cues from individual members may be associated with the entire group (Fiske and Taylor 1991).

A further complication of social interaction is that a single person, unlike a single inanimate object, will have different salient attributes depending upon the situation in which he or she is encountered (individuals will act differently at work and at home, for example). Thus, multiple schemata may need to be constructed for a single individual to account for different situations. It seems likely that the appropriate schemata will be activated by the combination of cues from the presence of the individual and contextual cues from the environment. It is interesting to note, however, that actions tend to be attributed to personality traits rather than environmental factors (see section 3.9.1). Given that dispositional cues are derived from the actions of the actor to whom they are attributed, whereas environmental cues are derived from comparatively static surroundings, it seems likely that the relative salience of dispositional and environmental factors accounts for at least some of this difference (see section 3.9.1).

5.4.6 Diachronic Relations and Causal Attribution

Whereas spatial information leads to the perception of synchronic relations, spatiotemporal information leads to the perception of diachronic relations and, ultimately, to the attribution of causal relations between events. As noted in section 5.4.2, changes in the environment result in marked changes in sensory stimuli and, consequently, enable the subject to break-up the durée of experience, somewhat arbitrarily, into distinct events. Clearly, such discrimination forms an essential component of the development of diachronic relations and causal attributions, which involve making associations between events. It should be noted, however, that knowledge of the continuity of activity in the environment and the continuity of experience play a role in forming and testing causal attributions. Consider, for example, a person who observes a moving object colliding with a stationary object. The entire sequence is continuous but a marked change occurs at the moment of collision: the velocities of one or both objects will change. Thus, the continuous experience can be partitioned into two events: the event of the ball moving up until collision and then the subsequent movements. Discrimination of the two events enables the subject to develop the spatiotemporal associations that will underpin causal attributions, such as the conservation of momentum (although the attribution will probably not be a precisely formulated physical law). It is important to note that the discrimination of the two events does not leave the subject unaware of the continuity of experience. As illustrated by Anderson (1990), perceived discontinuities greatly reduce the likelihood that a subject will make causal attributions. Considering ICT artefacts, animation on television, cinema and computer displays must operate within the confines of the flicker fusion frequency (see table 5.2) if observers are reliably to perceive events as being causally related and not simply a series of discrete and unrelated objects (i.e. images).

As noted in section 3.9.1, causal attribution involves the discrimination of a range of cues, including (White 1988): (a) cause precedes effect; (b) temporal contiguity; (c) spatial contiguity; (d) perceptual salience; and (e) covariation. The first three of these points are addressed by the combination of continuous experience and the discrimination of discrete events. Perceptual salience is addressed by basing this account on explanations of sensory experience and sensory physiology. Covariation can be accounted for by the encoding of spatiotemporal information into an associative network, as suggested in section 5.4.2. As an individual experiences many similar encounters, central tendencies will be reflected in the strengths of associations between the attributes of events and the relations between them, resulting in representations of diachronic relations that bear some resemblance to the analysis of variance. As noted in section 3.9.1, human performance does resemble, but does not exactly mimic the statistical process of analysis of variance.

5.4.7 Regularization

Just as attributions to individuals differ from simple synchronic relations, so social activities differ in some ways from simple diachronic relations. The differences do not just arise from the greater complexity of what is being observed. When social events are experienced - either by observation or participation - the experiences affect future performance and, hence, the ways in which future social events will take place. Thus, the extension of causal attribution to social activity is something of a self-fulfilling prophecy, ultimately leading to the emergence and evolution of regularized social practices.

The cognitive basis of social regularization occurs in its most simple form as the automaticity of response resulting from conditions. As a person repeatedly performs a kind of action in response to a particular kind of stimulus, the behaviour becomes more rapid, more accurate and can be performed with less cognitive effort (Baddeley 1990). As demonstrated by behaviourist research (Gregory 1987, pp.71-74), automaticity is improved by the presence of rewards and sanctions. At the level of social interaction, no deliberate conditioning takes place but there is a general tendency for actions that have improved the likelihood of one's goals being achieved being repeated in the future. As actions become more and more practiced, they become less conscious and may, subsequently, be performed in response to appropriate cues, even though they are not directly motivated. Thus, social actions become regularized and, because other social participants experience the success or failure of others' actions as well as their own, the actions of most individuals within a given social organization will tend to become similar for different kinds of activities.

Although specific goals may explain why some social activities are reproduced, it does not obviously account for the extent and complexity of regularization observed in modern societies. Within the information systems literature, for example, the socio-technical approach was motivated in part by evidence of poor job design and the use of inappropriate technologies (Mumford and Weir 1979, Mumford 1983). More generally, rituals and customs continue in organizations without any obvious direct benefits to the actors involved (Bate 1994). As noted in section 4.3.1, Giddens assumes the existence of "a general motivational commitment to the integration of habitual practices across space and time." (Giddens 1984, p.64). This motivational commitment is not explained in detail, however. Going back to the discussion of self perception (section 3.9.6), Giddens's assumption seems questionable. Repeated performance increases the probability of recall and, hence, the probability of future performance. Thus, regularization arises from the interplay between cognitive functions and situation specific goals. No general motivation seems to be required. Furthermore, as pointed out by Deci and Ryan (1987) and Taylor and Brown (1988), flexibility of behaviour leads to increased motivation and interest. A general motivation to form habits and a motivation for flexibility and control are antagonistic and, consequently, will most likely result in increased stress and anxiety on the part of the individual. Given the considerable evidence supporting the existence of goal-directed behaviours considered in chapter 3 and the role of negotiation in social organization (see section 4.5.3.1), Giddens's motivational claim must be rejected in favour of task-relevant goals and a desire for flexibility. As discussed in section 5.4.9, however, individuals formulate their motivations and goals in terms of self-concepts and there seems to be a general motivation to maintain consistency between the multiple self-concepts that an individual holds (see section 3.9.6). Thus, the regularization of activities may, as Giddens suggests, have some anxiety-reducing effect.

5.4.8 Self-Concepts

Just as individuals construct schemata to represent the individuals with whom they interact, they similarly maintain self-concepts reflecting their own personality attributes, social roles and experiences. As noted below, they also construct idealized self-concepts to facilitate the satisfaction of motivations and goals (see section 5.4.9). Fiske and Taylor (1991, p.182) distinguish between the chronic and working self-concepts (see section 3.9.6). The chronic self-concept refers to the entire self-concept, which includes self-attributed personality traits, social roles, experiences and goals. The individual is never aware of the entire chronic self-concept but, like other parts of long-term memory, different parts of the chronic self-concept are activated by associated stimuli. The part of the chronic self-concept active at any given time is called the working self-concept (Fiske and Taylor 1991). Whereas social cognition researchers have tended to describe the self-concept in terms of numerous schemata, it is maintained here that, like categories in long-term memory, the different schemata are only implicitly defined by central tendencies in the patterns encoded in memory. The working self-concept, which may include motivations and goals as idealized components, however, can be regarded as relatively discrete, as explained in section 5.4.9.

5.4.9 Motivations and Values as Idealized Self-Concepts

Whereas cognition is concerned with the knowledge underpinning behaviours, motivations and values explain why subjects' behaviours occur at all. In terms of human communication, there are two primary concerns:

- what it means for a person to be motivated or to hold a particular value; and
- how a person infers the motivations and values of others.

Considering the first point, it has already been noted that some cognitive models assume that all behaviour is goal-directed, despite evidence that highly learned behaviours become more automatic, implying a reduced need for conscious motivation for the behaviour to occur (Fiske and Taylor 1991). The process is not one way, however. As Bargh (1994) points out, subjects can sometimes moderate or eliminate automatic behaviours if they are motivated to do so. For this reason, it is useful to consider motivations and values as part of larger mental constructions: idealized self-concepts that broadly represent to an individual how, in certain respects, he or she would like to become.

The working self-concept is highly changeable, with different parts of the chronic self-concept becoming more or less active depending upon environmental and internal (to the individual) stimuli. Given that the working self-concept is, therefore, available to the working memory function, it follows that it is the part of the self-concept that is utilized in reasoning and action. The problem of explaining how motivations and goals are brought into the reasoning process is resolved by recognizing that they are parts of the self-concept. Personal experiences are encoded into long-term memory as part of the chronic self-concept. These memories will include sensory stimuli and affective information (Fiske and Taylor 1991). Thus, a working self-concept will recall associated emotions as well as information derived from the sensory stimuli. Thus, a person who has previously experienced rewards or sanctions in association with events of a particular kind will recall the emotional responses to those rewards or sanctions when in a similar environmental context, involved in similar types of social event and/or with the person who meted out the rewards and sanctions.

Adopting this model of motivation, it can be expected that a person will act in a particular way if: (a) the person is aware of one or more of his or her capacities that will increase the likelihood of realizing part or all of an idealized self-concept; and (b) the likely benefits of any action will outweigh the expected requirements in terms of personal effort and resources. Included in the 'costs and benefits' of acting in a certain way will be the maintenance of valued aspects of the self-concept. An action may be highly desirable, for example, because it will increase wealth and the actor has constructed an idealized self-concept in which this is desirable. Attributes of the action may, however, be inconsistent with important aspects of the chronic self-concept, such as perceiving oneself as being honest, and, consequently, will create a conflict between the benefits of realizing the ideal self-concept and contradicting a valued attribute of the chronic self-concept. In such situations, anxiety, indecisiveness and other signs of stress are likely to occur (Fiske and Taylor 1991).

5.5 Social Organization

Having established a broad but integrated account of how individuals acquire knowledge and utilize it to guide their actions, it is now possible to reconsider how social organization develops through the interactions between individuals and consider the implications for developing information and communication technologies and information systems. This section begins by considering how social organization is likely to emerge, given the socio-cognitive model of the individual presented above, before going on to assess present techniques for the analysis of social organization to aid the development of ICTs and information systems.

5.5.1 The Emergence of Social Organization

A useful starting point for analyzing social organization is the individual's limited capacities for action. An individual develops capacities through experience but the number of capacities developed is fundamentally limited by the time taken to acquire them. Furthermore, some tasks cannot be performed by individuals owing to prohibitive factors, such as the physical activity involved being beyond the capability of a single person. Thus, even when considering the pursuit of purely personal motivations and goals, there are good reasons to expect social activity to occur.

The need to draw upon the capacities of others raises the question of whose capacities an individual should draw upon. For an individual to succeed in achieving a goal, he or she must be able to discover which individuals have the required capacities for action. One possible solution is to communicate with every individual one encounters until an appropriate person is found. Clearly, this is inefficient and does not guarantee success. If, however, one does identify a person with a particular ability, then that knowledge will be retained in the form of an association between a motivation, that person and his or her knowledge. In the future, such knowledge will direct one's search because it will be activated by similar motivations. Moreover, such knowledge will be acquired by every actor, so it becomes possible for search to proceed by drawing upon one's knowledge and other actors' knowledge about the locations of different capacities. Provided that all persons are willing to share their knowledge, it is clear that personal experience of pursuing personal goals is sufficient for an informal information system to emerge.

An information system of this sort will not be efficient because of its contingency upon the motivations, capacities and knowledge acquisition of each actor. Owing to the regularization of social activity, however, it is likely that social interaction will be far less 'random' than the account so far suggests. When a person performs an activity for another, he or she inadvertently increases other actors' associations between their motivations and his or her capacities for action. Consequently, when a person requires such a capacity, he or she will be more likely either to recall that person or to be informed about that person when searching by communicating with others. This, in turn, increases the likelihood of future utilization of an actor's capacity. Over time, the repeated associations between motivations and actions and the repeated utilization of capacities by actors will lead to the emergence of a social organization without the need for any explicit attempt by any actor to systematize knowledge of the environment or to systematize the environment itself.

An interesting consequence of this account of social organization is the emergence of a division of labour. Unlike the conventional economic account of this phenomenon (see sections 4.2 and 4.4.1), this division does not result from capital or from exploitative relationships, but from the physiological and cognitive limitations of self-serving individuals. Furthermore, it is knowledge derived from experience that is of value in this account, making the division of labour as much a cognitive as a physical phenomenon.

Where the satisfaction of a person's motivations requires the involvement of several persons, coordination will be one aspect of the required capacities. Two possibilities arise: (a) the actors perform their activities, exchanging information with one another to ensure that each individual performs his or her actions in the required way; or (b) one or more persons acquires particular capacities for providing coordination and, thereby, become involved in the collective action in the capacity of a coordinator of others' skilled actions. Thus, task-contingent teams are likely to arise. Of course, the effective performance of many actions requires the use of physical objects and the concept of resource, therefore, acquires significance in this account of social organization.

5.5.2 Resources

The notion of an information system, as informally described above, illustrates a fundamental trait of much human communication: it takes place for the purpose of exchanging evaluative information. In the case of capacities, the main interest is in selecting the person whose capacities are most likely to enable one's motivations and goals to be satisfied.

The introduction of resources into the model of social organization only serves to reinforce the evaluative role of much communication.

As noted in section 4.5, capital plays a fundamental role in the conventional account of economic power. Underlying these explanations is the concept of a resource. Resources are of most interest in explaining economic organization when they are scarce. Where an individual's capital provides them with access to less of an object than they require to satisfy their motivations and goals, choices have to be made. In such cases, evaluative information is critical. Reviewing the discussion of social organization above, it is clear that the properties of social organization arising from, and reproduced by, the scarcity of capacities are equally applicable to scarce physical objects. There are, however, several important differences between capacities and resources:

- capacities are limited in availability by time, whereas resources are limited by quantity;
- capacities are potentials for transforming, whereas resources are potentials for being transformed; and
- capacities can be reproduced through social interactions, whereas resources can only be transformed, with an increase in one leading to a reduction in the availability of others.

These points combined provide a basis for explaining the power relations that arise in social organizations and how they relate to the availability of capacities and resources.

5.5.3 Power and Legitimation

The discussion so far has identified three key concerns relating to the emergence of social (especially economic) organization: (a) capacities for action; (b) resources; and (c) individuals' knowledge insofar as it is constitutive of a social system for providing evaluative information to actors. The analysis has been quite simplistic, however, in assuming that communication and the sharing of capacities are unconstrained. Furthermore, no account of resource ownership and control has been given.

The discussion of organizational change theories in section 4.6.1 provides some insight into the way in which power is associated with capacity. Barnard (1938), for example, points out that - unless physically compelled to do so - actors will only utilize their capacities when they accept as legitimate the authority of the person asking them to perform. Legitimate authority arises from either:

- a mutual exchange of capacities and/or resources either on a one-off basis or over a prolonged period (for example, a labour contract); or
- the provision of capacities and/or resources to others in order to satisfy a motivation or goal (as in the case of voluntary work).

With respect to resources, power is fundamentally derived from possession. The power that can be derived from the information system is more complex, arising from the various options that an individual has for engaging in social activity:

- an individual with knowledge of the location of a resource or capacity may decline to provide information to another or may exchange the information for other information, capacities or resources;
- an individual may utilize resources and capacities to inform others in the information system about his or her capacities or resources, thereby increasing the relative likelihood of his or her capacities or resources being sought by other actors;
- an individual may utilize resources and capacities to obstruct the information activities of others or to provide misinformation about the capacities and resources of other actors;

- an individual may utilize resources and capacities to inform others in the information system of his or her need for capacities or resources and the exchanges that will be entered into for them; and
- an individual may utilize resources and capacities to acquire information about capacities, resources, the information system or the actors participating in it.

In general, the leveraging of the power derived from capacities, resources and information systems increases an individual's ability to satisfy his or her motivations and goals by engaging in exchanges that place goal-relevant capacities, resources and information sources under his or her legitimate authority. Within the bounds of this authority, the individual is able to alter the structural properties of social systems that arise from the interactions described in the previous section.

	No Institutionalized Practices	**Enabling Institutionalized Practices**	**Constraining Institutionalized Practices**
Aligned Goals	**Coalition** Cooperate insofar as joint gains can be achieved.	**Cooperation** Cooperate for joint gains within organizing framework for interaction accepted by all parties.	**Contravention** Cooperative relationship aims for joint gains that conflict with precedent and past agreements.
Unaligned Goals	**Competition** Conflicting goals lead both parties to avoid interacting or only to interact to further own goals.	**Contradiction** Conflicting goals to be resolved within previously accepted organizing framework.	**Conflict** Conflicting goals not resolvable within previously agreed organizing framework.

Table 5.4: General Types of Social Interaction

In cases where shared or aligned goals exist, individuals can go beyond the exchange relationships discussed above by entering networks of negotiations, whereby the individuals with aligned goals agree how they will utilize their capacities, resources and knowledge to contribute to an outcome that will satisfy the motivations and goals of all cooperating actors. In many cases, cooperation of this nature will be shortlived, coming to an end when the cooperative activity is completed. It is possible, particularly where legitimate authority provides a basis for repeated social encounters, that the cooperative activities will become regularized. In such cases, the interactions between individuals may take on a more permanent form, with individuals joining and leaving the cooperative organization depending upon their capacities, resources, knowledge and the extent to which membership satisfies their motivations and goals. It is where social organization achieves this level of stability that roles (i.e. sets of capacities and knowledge requirements) are filled by various individuals that the structural properties of social activity become institutionalized and the cooperative social interactions come to be regarded as an organizational form. Within the organizational form, the interplay between personal motivations and goals and the regularized nature of social activities results in a range of different types of negotiated activity (see section 4.5.3.1), which can be broadly classified as shown in table 5.4.

5.6 Information Systems Design

In section 5.5.3, information systems were shown to be a means of satisfying one's motivations and goals and a potential source of power. An information system contributes to satisfying motivations and goals by providing evaluative information to the actor. The derivation of power from an information system, in contrast, involves the use of capacities and resources to act upon it. The capacities for affecting information systems are referred to here as *information and communication technologies*. The resources used to support information and communication technologies (ICTs) are referred to as *ICT artefacts*. The intentional development of ICTs and ICT artefacts in order to alter the information systems component of a social system constitutes the process of *information systems development*. Drawing upon the theory presented so far, ICTs, ICT artefacts and information systems development are considered in turn.

5.6.1 Information and Communication Technologies

Central to the description of human communication is the ability for each participant to draw upon mental organizations that have some similar components. Communication takes place by encoding as a communication artefact an organization that can be interpreted with respect to these intersubjective components (see section 5.3). Thus, communication involves the following knowledge and skills:

- mental organizations (i.e. knowledge) containing some intersubjective components that serve as the referents of the message;
- intersubjective knowledge that constitutes a convention for translating the organization of a message into a physical organization and the skill to use this knowledge to construct (or deconstruct) a communication artefact; and
- the knowledge and skill to transmit or store, and receive or search for communication artefacts using communication media.

The general characteristics of personal knowledge and how intersubjective knowledge arises have been considered above. The development of knowledge specifically relating to the construction and use of communication artefacts (i.e. information and communication technologies) requires more detailed consideration.

Messages are encoded into communication artefacts through the organization of their attributes. It follows that communication artefacts can most usefully be described in terms of the *modes of organization* that they enable. Many communication artefacts are largely static, enabling only a single, fixed message to be encoded. After encoding, the message cannot readily be changed, or can only be changed in limited ways. A message printed on paper is an example of such a static organization. Such artefacts usually require ICT artefacts, such as pens or printing machines and paper, to encode the message and, thereby, construct a communication artefact from a piece of paper. Other communication artefacts are more dynamic, particularly those communication technologies that exploit propagation (for example, telephone and television) or are used in co-presence, such as an abacus. This communication artefact is the representation of numbers by the arrangement of beads on wires (see figure 5.2). Each row of beads is part of the abacus, which is the ICT artefact.

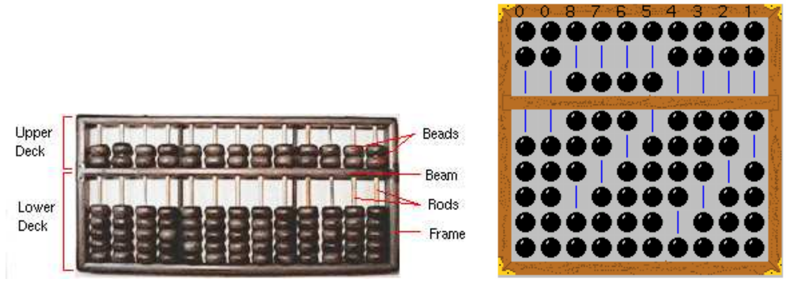


Figure 5.2: The Representation of the Number 87,654,321 Using an Abacus (Source: http://www.ee.ryerson.ca:8080/ elf/abacus/intro.html, Accessed May 1999)

Whilst dynamic in its representation of numbers, the abacus is much more limited than many static forms of representation because it supports only a single mode of organization. In contrast, computer-based ICT artefacts provide users with enormous flexibility for specifying different modes of organization in terms of general rules and procedures for representation (i.e. computer programs). Once specified, modes of organization can be used to construct dynamic representations. The combination of scope for specifying modes of organization and for changing the message represented as a communication artefact on a video display or some other device makes the use of computer-based ICT artefacts potentially the most flexible form of communication currently available.

Like many ICT artefacts, those based around computers are typically able to store representations. In the case of computer-based ICT artefacts, however, the organization used for storage and presentation are different. The stored message must, therefore, be transformed by rules to provide communication artefacts and *vice versa*. As illustrated in chapter 2, the HCI community has focused on the development of computer-based representations as they appear to users, whilst the software engineering community focuses more upon the stored representation. It should be noted, however, that both the presented and stored organizations and the rules for translating between them are constitutive of the message as interpreted by the user. Consider, for example, a relational database. The presentation on screen of a form displaying the contents of a particular database record cannot be accurately interpreted by the user without some knowledge of the semantic con-

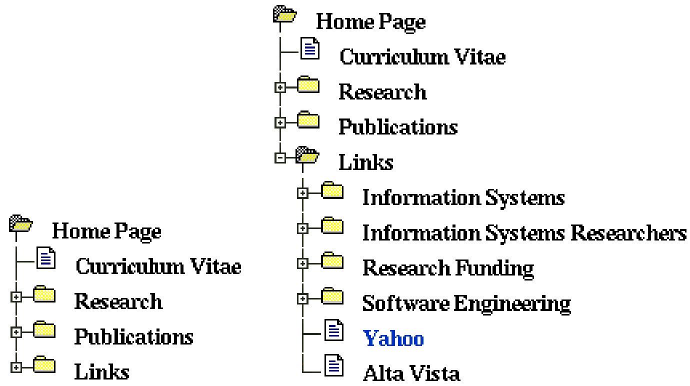


Figure 5.3: A Dynamic Tree Structure as a Mode of Navigation

structs underpinning the organization of the database. Furthermore, there are possible arrangements of a forms-based interface, for example, that violate the rules for encoding into a relational database.

The ability to use some ICT artefacts to reconstruct messages (or components of large messages) from stored representations indicates two other properties of communication artefacts that have significant implications for human communication. First, *modes of selection* refer to the ability to filter information prior to its presentation. In the case of a database, for example, it is possible to specify criteria for restricting the displayed representation to only task relevant parts of the stored representation (for example, retrieving all records where the surname is 'Bloggs'). Second, *modes of navigation* refer to the mechanisms by which the user can interact with the modes of organization and selection of ICTs that support the dynamic construction of representations. An example of a dynamic mode of navigation is shown in figure 5.3. The combination of selection and navigation with the separation of storage from presentation makes it possible to store enormous messages that can be selectively reconstructed by ICT users. The relationships between storage, presentation and interpretation must, however, be carefully considered in the design and construction of such technologies.

5.6.1.1 Modes of Organization

Modes of organization can be described in terms of (a) dimensionality and (b) granularity and abstraction. A brief consideration of the model of the individual presented above identifies several aspects of an individual's knowledge that might be represented as dimensions. Firstly, there are basic parameters of each sensory modality. With vision, for example, there are the three spatial dimensions, the temporal dimension and the properties of colour. In terms of higher level cognition, entities and their attributes and values are likely to feature prominently in representations. Abstract concepts could also be the subject of communication, although the intimate connection between subjective experience, the sense of self and the body's physiology suggest that these will be communicated by reference to concrete phenomena, including the body. A frequently cited example of such communication is the use of body language, where the subject's body serves as a communication artefact.

Whereas the dimensions a person might want to communicate are numerous, information artefacts are quite restricted in the number of dimensions they can represent. As illustrated by Tufte (1990), this has led to the development of many technologies for information design, which enable the number of encoded dimensions to be increased. New information and communication technologies are providing an increasing number of dimensions that can be exploited by the information designer. The electronic display, for example, provides a temporal dimension, through the ability to animate, that was very restricted on paper. Virtual reality displays extend this even further by providing more realistic three dimensional displays.

A key contribution to effective information processes is the development of a clear correspondence between the dimensions to be represented and the dimensions of the communication artefact. In the case of developing a visual analogue, such as a schematic diagram, some correspondences are obvious (spatial dimensions are usually best represented using the spatial dimensions of the medium, for example). Other dimensions must be represented using established social conventions (for example, temperature can be represented by a colour scale with blue representing cold and red representing hot) or an arbitrary but consistently applied convention must be used. In the case of conventions that are likely to be unfamiliar to the recipient community, other properties of the communication artefact, or other communication artefacts, must be exploited to clarify the dimensions being represented. It is in this way that similarities between individuals' mental organizations can emerge through social interaction, confirming the second point raised in section 5.3.

A second information design issue is the need to balance the level of detail available and the level of detail that will convey the intended message most effectively. The criticality of this aspect of organization is illustrated by phenomena such as 'information overload.' In terms of the knowledge being used to construct the message, the key constraints are the *granularity* of the relevant data. Granularity is determined by the basic elements used to construct the message. A message about medical conditions, for example, might refer to specific conditions, such as asthma, or to classes of illness, such as respiratory conditions.

When encoding a message, attention must be paid both to selecting the appropriate level of granularity that will convey the message accurately and to ensuring that the properties of the medium are exploited to preserve the accuracy of the message when it is interpreted by a recipient. Thus, several characteristics of information artefacts must be borne in mind: (a) precision; (b) supported levels of analysis; and (c) layering and separation. In general, the level of precision must be sufficient to accommodate the level of granularity chosen by the sender. If this is not possible with the available information artefact, then a method of data compression must be used, which will allow a complete representation to be constructed, albeit with a consequent reduction in accuracy. As illustrated by the discussion of the visual system, a person is able to experience aspects of the environment at different levels of analysis. When interpreting an information artefact, this ability can be exploited by paying attention to the message design at multiple levels. An effective graph, for example, will permit the user to interpret individual data points and to identify patterns in clusters of points or general trends across the entire data set. Layering and separation, as discussed by Tufte (1990), can be exploited to increase the number of dimensions of a message that can be encoded into the limited dimensionality of an information artefact. Using different coloured lines or different shapes of data points, it is possible to represent several dimensions on a two-dimensional graph. The creation of distinct layers by, for example, using different colours, permits the recipient to shift attention between different dimensions and, perhaps, to make comparisons between dimensions and across the basic elements of the message represented by the single data items. It is important to recognise when designing information artefacts and constructing messages that the relation between the artefact and the body is critical to the successful interpretation of messages. A person can ease the distinction between levels of analysis on a paper graph, for example, by altering the distance between the paper and his or her eyes. Such manipulation is more restricted with a computer display, however, because - although the graph is visually very similar - changes in distance result in considerable changes in light intensity because electronic displays emit light, rather than reflecting it. Furthermore, the spatial resolution of modern displays is limited. As explained below, limitations to information organization can sometimes be overcome by dynamic information and communication technologies (such as computer displays) through the provision of modes of selection and navigation.

5.6.1.2 Modes of Selection

The greater availability of information and communication technologies has led to an increased flow of messages in many organizations and, in many cases, has enabled messages to be created that are much denser in content. Although additional content can sometimes be valuable (for example, a photograph of a person in place of a verbal description), it can also lead to 'information overload,' where the relevant parts of a message effectively become hidden from the recipient (for example, providing a full transaction log in place of an exception report). Generally speaking, it is desirable to include - or, at least, emphasize - those organizations of data most relevant to the sender's or recipient's goals. This may be relatively straightforward for communication between a sender and a single, well-known recipient, whose information needs (or goals) are understood. The situation is greatly complicated, however, by the existence of multiple recipients, uncertainty regarding information needs and a multiplicity of recipient goals. The effective use of modes of selection can greatly increase the value of communication in all cases. By implication, the careful introduction of selectivity into an information artefact's design can reduce communication costs because interpretation time (and, consequently, decision time) is reduced and single artefacts can readily convey multiple, related messages from which the recipient can extract the relevant components.

Fundamentally, modes of selection embellish the organization contained by an information artefact so that is exploits the characteristics of human perceptual systems to help the recipient draw out meaningful (especially goal relevant) portions of the message with minimal cognitive effort. The precise relationships between techniques for communication artefact design and information interpretation are, however, quite poorly understood and require detailed empirical study (see section 7.4).

5.6.1.3 Modes of Navigation

Like modes of selection, modes of navigation enable senders to transmit large, complex messages and recipients to extract from them only those aspects of a message that they require. Whereas selection relies upon properties of the perceptual systems, navigation (like organization) is dependent upon the ability of the individual to construct and manipulate mental representations. The similarity of modes of organization and navigation stem from the fact that navigation effectively arises from an 'organization of organizations.'

The simplest type of navigation is a point of reference. Such references can be compiled to provide indices of the contents of single or multiple communication artefacts. Graph-like structures, such as indices, are extremely flexible, allowing the sender to make any number of relations between any message components. Graph structures are not amenable to efficient mental representation, however, because the relations are arbitrary. Frequently used relations will eventually be learned by artefact users, but no simplified navigation rules can be inferred.

The introduction of more structured modes of navigation can both facilitate learning and reduce cognitive load. Trees, for example, are a specific class of graph structure in which relations between nodes are restricted to a hierarchical pattern (see figure 5.3). Tree structures are more readily learned, particularly by users who have a good knowledge of the domain to which the message refers. Taxonomies are exemplars of the use of trees to enable rapid navigation of large bodies of knowledge. The main limitation of tree structures is that the end-goal needs to be defined in terms of attributes or node identity for successful navigation to be guaranteed. For example, navigating a biological taxonomy to find a particular species requires the user to be able to *recognize* (recall is not essential) either the membership of kingdom, phylum, class, order, family and genus to which the species belongs or have sufficient knowledge to diagnose membership at each level of the hierarchy (for example, matching a species' description with knowledge that the target has grey fur, brown eyes, four legs, and so on). Users will often have only vague knowledge of their end-goal or its defining properties. In such cases, navigation will need to be more flexible, providing facilities based upon partial criteria and/or browsing through the message. Again, systematic evidence of the relationship between modes of navigation and users' abilities to interpret information and perform tasks is very limited.

5.6.2 Interacting Through ICTs

As noted in section 2.4, it is instructive when considering ICTs to distinguish between interacting with and interaction through ICT artefacts. Interacting *through* ICTs involves the development of skills and supporting artefacts for (a) enhancing interpersonal communication in co-presence or (b) providing distant interpersonal communication mediated by some ICT artefact (sometimes referred to as establishing a 'virtual presence'). Common artefacts for enhancing communication in co-presence include presentation aids, such as overhead projectors. Artefacts for mediating distant interpersonal communication include telephony and television. Whether an ICT is to support co-presence or mediated communication, its main impacts will be changes to the usual practices of interpersonal communication. These practices derive from a combination of the physiology of the actors and the regularized social practices with which they are familiar. Thus, the primary objective can be considered to be the accurate transmission of sensory stimuli. Although an understanding of sensory physiology is extremely important in this respect, social factors arise through the regularization of aspects of interpersonal communication, such as body language and conventions for turn-taking in conversations. The concept of regionalization provides a starting point for analyzing these factors. Structuration theory provides a valuable account of how regionalization affects social interaction and, thereby, contributes to the emergence of social institutions (see section 4.3.2). Structuration theory's account of regionalization, however, is not sufficiently detailed to help explain how ICTs affect social interaction. By considering both the physical and mental aspects of psychosocial phenomena, the description of the individual presented in the previous sections can be extended to support such explanation. The location of the body and its perceptual extent defines a complex of regions in which possible interactions between the individual and the environment can take place. This region is potentially very large, depending upon the magnitude of stimuli (some astronomical observations, for example, can be made with the naked eye). Social interactions in co-presence, however, take place within a much narrower area. As shown in figure 5.4, boundaries can be defined by:

- **combined properties of the sensory modalities and concrete objects** - as shown in figure 5.4 by the conical region of visual attention bounded by the walls of a room;
- **the physical dimensions of the body** - as shown in figure 5.4 by the area of a person's reach; and
- **social factors** - as shown in figure 5.4 by the region in which face-to-face conversations typically take place.

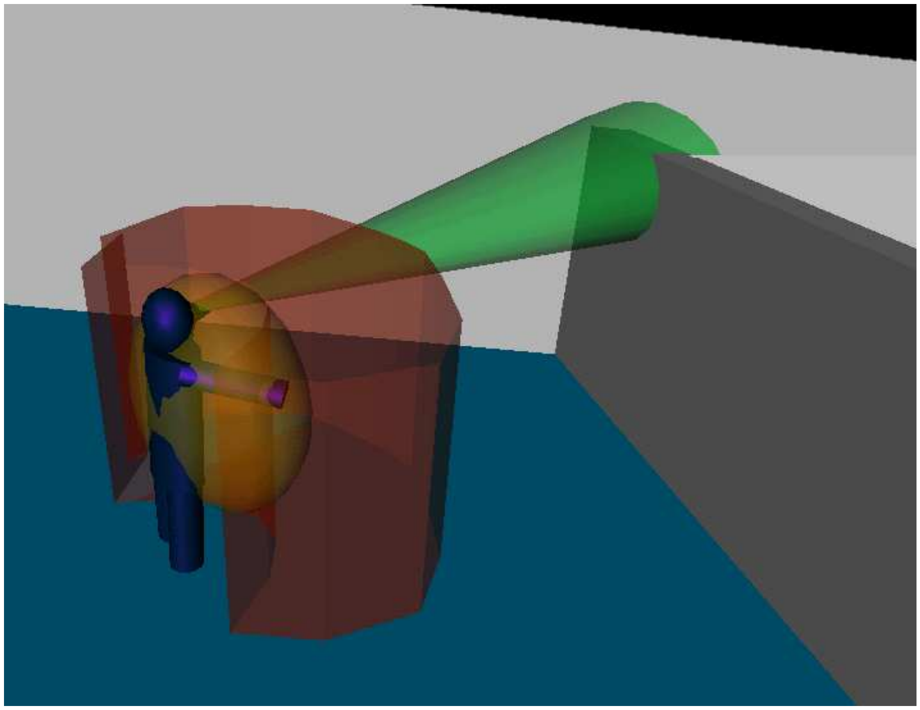


Figure 5.4: An illustration of how: the use of the hands and arms to manipulate objects (yellow); the distance at which conversations usually occur (red); and the area of greatest visual acuity (green) combine with the structure of the environment, resulting in the regionalization of space

Given that regionalization has been explained in terms of the body situated in a physical environment, a major impact of ICT artefacts is their mediation between the environment of interest and the body. In explaining these impacts, it is useful to exploit the distinction made in section 2.4 between interacting *with* and interacting *through* ICT artefacts.

As illustrated by figure 5.5, interacting through ICT artefacts implicates tacit knowledge gained from real world experience. Thus, the actor has a strong expectation (at least for unfamiliar technologies) that typical patterns of regionalization will arise. Thus, for a given distance from a screen, for example, an actor engaging in mediated face-to-face communication will expect visual cues, such as apparent distance judged by visual angle, to be consistent with the analogous situation in co-presence. Actual performance depends, however, upon the size of the display, focal length of the camera lens and distance of the actor from the camera. It is for this reason that interaction problems arise with desktop

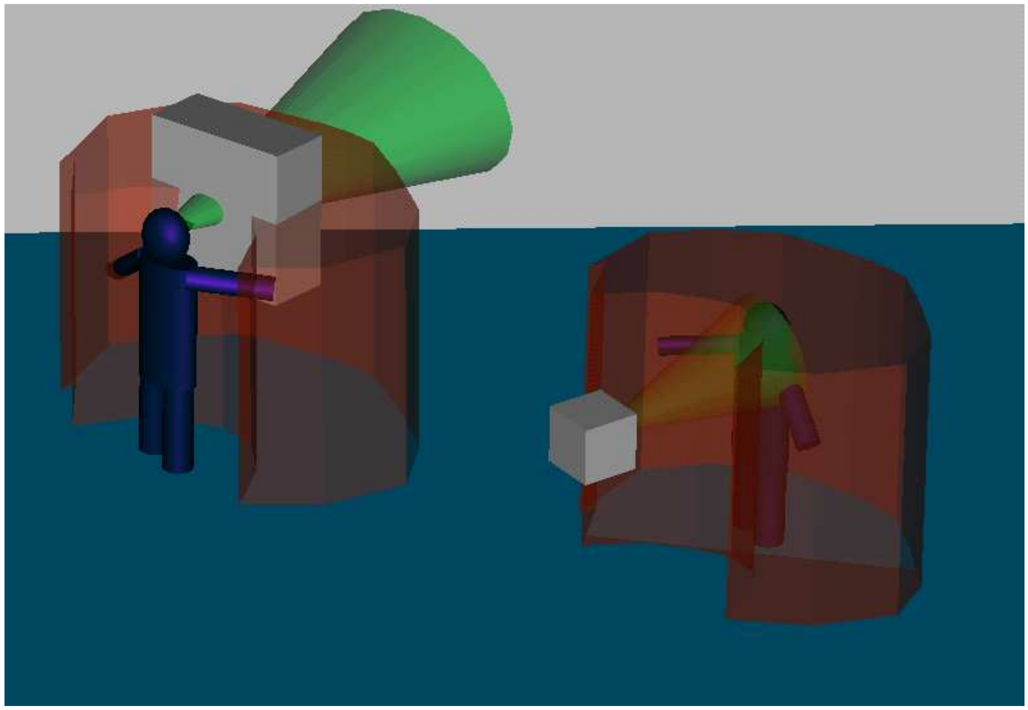


Figure 5.5: An illustration of how a socially defined region is dependent upon the design and arrangement of ICT artefacts (see text for explanation)

videoconferencing. If person *a* sits too close to the camera, his or her face will appear too large on the screen, leading person *b* to move further away from his or her workstation. Consequently, *b* will appear to be distant from *a* who may, consequently, move still closer to the screen or, perhaps, make personality inferences as though *b* was being distant and unfriendly.

Although regionalization is not sufficient to explain how individuals interact in co-presence or through ICTs, it is an essential starting point. Understanding turn-taking and interactions with objects in the environment, for example, first requires an understanding of how individuals position themselves in space as a result of physiological, cognitive and social factors. When examining a small object, for example, both the tendency to pick it up and manipulate it with one's hands and the norms relating to social positioning will determine whether individuals will examine the object together or take turns.

5.6.3 Interacting With ICTs

The conception of an information system presented above has illustrated the fundamental role of both content and context to message interpretation. It has been argued that content and context are encoded into human memory as patterns of association and that similar experiences lead to similar and, therefore, related patterns. Knowledge can be recalled in more or less specific form and used as the basis for actions, including communication. Constructed messages are necessarily somewhat vague for two reasons: (a) the referents of the message are not precisely defined but stored in terms of patterns and tendencies; and (b) intersubjective knowledge is constituted by similarities between individuals' mental organizations and the degree of similarity will vary considerably. Much of the ambiguity arising from the vague nature of messages is resolved by the existence of a shared context of communication, which results from the actual or mediated ('virtual') co-presence. Where large organizations are constructed and stored (as in the case of computer-based information systems), co-presence is not necessary to the communication of messages and, hence, situations are likely to arise where the context of communication does not arise as an accidental property of the communication process itself. Consequently, a major factor determining how successful stored organizations are in terms of informing actors is how well the context of communication can be conveyed.

It was noticed in section 5.6.1 that modes of navigation rely on the 'organization of organization.' In other words, stored representations can make use of structural properties that are abstracted away from the specific content and context of experiences. In this way, central tendencies can be used to define categories, attributes and relations to guide the location of task-relevant components of the stored organization. The categories, attributions and relations are not (at least using current technologies) stored in the informal associative form that characterizes human memory. In the case of figure 5.3, for example, the user has encoded 'information systems' and 'software engineering' as separate catogories. It is likely, however, that there is substantial overlap between the two categories. In human memory this will be reflected in overlapping cases being stored as patterns that are associated with some aspects of the central tendencies defining both terms. In stored organizations, however, the individual encoding details of instances must decide whether to encode boundary cases as members of a single category or as a member of both. In order to make the modes of navigation amenable to other individuals and to allow many individuals to contribute to the development of the stored organization, such organizing principles must either be readily inferred from the storage and communication media, or be stated explicitly.

For organizing principles to be readily inferrable, the presentation of the communication artefact must indicate existing intersubjective knowledge that can be drawn upon by users to manipulate and interpret 'messages' derived from the stored organization. Perceptual cues may be utilized, such as representing information in tables so that categories, attributes or relations are implied by the rows, columns or cells. Because horizontal and vertical lines are perceptually salient, users will tend to interpret an organization of this sort as meaningful and attempt to use it to direct their search. Alternatively, social conventions may be used to indicate how the representation of the stored organization can be manipulated. In the case of figure 5.3, for example, folders are used to indicate elements that contain other elements, whilst pages are used to indicate where messages (rather than further organizational information) is presented. This mode of navigation exploits an analogy with organizing principles used for concrete communication artefacts, even though the computer-based organization actually enables more layers of hierarchy than a physical system would readily allow.

The explicit statement of organizing principles is fundamentally limited by the extent to which discursive knowledge can be used to communicate tacit knowledge. The manipulation of communication and ICT artefacts involves practical skills, for example, that are most readily transferred from one person to another in situations of co-presence. Although such transfer can be achieved by training (i.e. changing socialization activities for new and existing actors in a social system), it is sometimes only possible to provide instruction via the ICT artefact itself. In such cases, the effectiveness of the ICT will be constrained by the ease with which encoded instructions can be interpreted by users and used to develop the tacit skills required by practice (i.e. trial and error learning).

In terms of developing the skills to interact with technologies, it is useful to refer back to the four economic issues relating to technology development, presented in section 4.5.4:

- **Information and Communication Skills** - How can individuals and groups best develop skills for translating, for example, their goals, political and environmental pressures into articulations of the organizational form's technological requirements?
- **Technological Skills** - How can individuals and groups best develop cognitive and practical skills for designing and developing networks of technological activities to satisfy their technological requirements?
- **Technological Artefact Evaluation** - How can individuals and groups best be guided in their selections of appropriate technological artefacts, given that they must often acquire such artefacts with very little knowledge of their future technological requirements?
- **Technology Skills** - How can individuals and groups develop skills for incorporating and adapting available technological artefacts effectively to realize their technologies?

The conventional approaches to developing computer-based information systems, discussed in chapter 2, have tended to place most emphasis on developing ICT artefacts, rather than the skills to use them effectively. Where user issues have been addressed, as in the information systems and HCI literature, more emphasis has been placed on eliciting user needs and attempting to satisfy them through the development of ICT artefacts or through organizational change. Whilst these efforts are by no means wasted, it would seem that more attention to developing the above types of skills in the context of users' task characteristics ought to be a major component of information systems development. This will require a substantial change to the way in which information systems development is approached.

5.7 Implications for Information Systems

The theory presented in this chapter is intended to provide a basis for understanding information systems and how they develop. By paying particular attention to the relationship between individual actor and social organization (of which an information system is an intrinsic component), a somewhat different conception of information systems development and the issues relevant to the design of ICT artefacts has been proposed. Chapter 6 illustrates how the theory can be utilized to reconsider practical information systems issues. First an alternative model of the information systems development lifecycle is proposed. Second, and within the context of this lifecycle model, the evaluation of potential information systems developments is explored.

Chapter 6

Initial Applications of the Socio-Cognitive Theory of Information Systems

6.1 Introduction

The general aim of information and communication technology (ICT) investment evaluation can be broadly stated as follows:

> to provide a process by which (a) the costs and benefits of alternative investments can be estimated; (b) an alternative can be selected; (c) the development of the ICT system and accompanying organizational changes can be monitored to control costs and realize benefits; and (d) the long-term payoff of the investment can be reviewed.

Unfortunately, this general statement does not translate into practice because the above statement belies two well-recognized problems that have beset all attempts to evaluate

ICT-based systems developments:

1. the identification of costs and benefits and the demonstration that they are incurred by, or arise from, a specific course of action; and
2. the incurring and perception of different costs and benefits by different stakeholders.

The thesis of this chapter is that these two problems and the common failure to realize the benefits of ICT investment (Ward, Taylor and Bond 1996) can be overcome by modelling information systems development in terms of social transformation. Section 6.2 places the above claims in context by briefly reviewing current 'best practice' as described in the academic and practitioner literature. Section 6.3 utilizes the theory presented in chapter 5 to develop a model of the information systems development lifecycle. Section 6.4 then utilizes this lifecycle model to develop a methodology for appraising working practices and aggregating the resulting analysis to address the problems of evaluating ICT investment payoff. Section 6.5 highlights the practical implications of the methodology and draws conclusions.

6.2 Current Evaluation 'Best Practice'

6.2.1 Evaluation as a Process

In addition to developing numerous evaluation techniques to address the specific needs of ICT investment, the academic literature has emphasised the importance of the organizational process in which the techniques are applied. An exemplar of this research is the analysis presented by Symons (1991), which is based upon Pettigrew's distinctions between content, context and process (Pettigrew 1985). Content refers to the application of an evaluation technique to a chosen set of criteria. Context and process are more ambiguously defined and difficult to distinguish between in practice. Context includes, for example, organization structure and internal politics, whereas process "refers to the actions, reactions and interactions of interested parties" (Symons 1991, p.210).

The process and context categories are essentially attempts to distinguish between aspects of the organization that directly and indirectly affect the use of evaluation techniques. Although this type of analysis may be a useful conceptual aid for an observer, it is less useful as a basis for guiding evaluation because process and context classification has little bearing upon the actions of stakeholders. Consider, for example, a chief information officer (CIO) who submits an investment proposal to the Board of an enterprise. The nature of that proposal and his or her approach to gaining its acceptance are determined by the CIO's experiences of a 'culture' and an 'organizational structure' and not his or her perception of them as entities.

Whilst the argument above questions the practicability of one conceptual model of process, it is not intended to deny the value of considering social and political factors affecting investment evaluation. Currie's study of investments in computer-aided design (Currie 1989), for example, illustrates that there are documents to be produced and negotiations to be undertaken during evaluation. As argued above, it needs only to be recognized that the artefacts and events that constitute evaluation are the results of actors who have different personal experiences to draw upon and different perceptions of the same evaluation problem and process. This suggests that evaluation can be facilitated throughout systems development by encouraging stakeholders to model their perceptions of the factors they believe most influence their own and others' actions. Such information provides a basis for understanding the complex relations that bear upon the application of evaluation techniques.

6.2.2 Costs and Benefits

The overriding concern of evaluation must be obtaining the information needed to guide development towards achieving its objectives. The identification of points in a process at which information should be collected and what should be recorded at these points has, however, received less attention than techniques for manipulating the results. Thus, a small but significant step forward for evaluation methodology is the clear formulation of principles for constructing evaluation measures. The following analysis provides a justification for adopting the following four principles:

- metrics must be causally relevant;
- metrics must be reliable;
- metrics must directly relate to concrete phenomena; and
- metrics must be relevant to users' actions.

Considering the first aspect of causal relevance, a uniform process, such as a semi-automated production line, can be subject to precise evaluation because it is highly uniform. Each stage in the process is formally defined and has readily identified inputs and outputs. The basis of scientific management (Taylor 1913) was the development of such processes and the use of precise measures to evaluate their efficiency. When studying more varied processes, where workers have considerable discretion over the organization of their activities, identifying a causal chain to evaluate is much more difficult and inevitably less accurate. In general terms, the accuracy of metrics is proportional to the degree of organization of the process being measured. Thus, a clear understanding of the various configurations of causal factors is necessary for the construction of sound metrics.

The second principle - that metrics should be reliable - refers to a number of trade-offs to be made when constructing metrics, particularly when the activities are subject to change. Specifically, establishing a reliable metric involves balancing at least the following: (a) the provision of similar results in similar circumstances; (b) ensuring the variation in the measure corresponds with variations in the phenomena being measured; and (c) ensuring the measures are stable across evaluators. The validity of the first point is clear: when measuring a specific attribute of a class of phenomena, it is desirable for the value obtained to reflect similarities and differences between different instances in that class. When evaluating the colour of apples, for example, one wants to ensure that - even though apples are not of a uniform colour - predominantly red apples are recorded as being 'red' and predominantly green apples are labelled 'green.' Accuracy is of more importance in this respect than precision. In many cases, an ordinal scale or a task-specific categorization will be sufficient. In cases where precision is desirable, a trade-off between the accuracy of the metric's scale/classification and the cost of evaluation efforts must still be made. The correspondence of changes in the system and changes to a metric's value, point (b) above, is another requirement. If document flow is being incorporated into the metrics for insurance claims processing, for example, it is essential that the measurement system tracks the movement of documents to an agreed level of accuracy. If, for example, a goal of the system is to process 80% of claims through a 10 stage process in 20 days and 99% of claims in 30 days, metrics can only contribute to improving the process if they record claims document movement to an accuracy 1 day or better. For evaluating this goal, it is of little value, for example, to know that all documents stay at stage 5 for up to 3 days, but valuable to discover that 60% of claims pass this stage in 1.5 days, 20% in 2.0 days and 20% in 2.8 days. To record with such precision, the measurement system must be sufficiently sensitive to changes in the activity system. Point (c) above recognizes the differences between individuals' task performances and the recording of evaluative information. Where metrics are being related to individual performance, it is critical to management that the measures accurately reflect differences in the performance of individuals over time and variations in the performance of tasks by different individuals. It is both ethical and politically astute to ensure that metrics reflect and are seen by workers to reflect variations in performance in a fair and consistent manner, particularly if they are used as the basis for rewards and sanctions.

The third principle for constructing metrics underpins the previous two and, perhaps, for this reason is most frequently overlooked. A great deal of research has been conducted into developing metrics for estimating the intangible costs and benefits of IT investments (see, for example, Parker et al. 1988 and Farbeyet al. 1995). Most of the approaches attempt to make investment evaluation more complete by supplementing readily quantified metrics with notional values where attributes cannot be measured (see, for example, Coleman 1995). The limitations of such an approach are numerous (see Hemingway 1997a for a detailed critique), and include:

- financial values mislead decision makers because they do not reflect actual exchanges under market conditions (Self 1970);
- the different values cannot be shown to be independent (Hemingway 1997a) and the evaluation, thus, risks double counting costs and benefits; and
- the definitions of the values are open to very different interpretations by stakeholders and the estimates are, consequently, unlikely to satisfy the reliability conditions described above (Guba and Lincoln 1989).

The difficulties of notional values cannot be overcome because the values have little or no basis in fact. For this reason, it is asserted here that metrics should only be developed where concrete evidence can support them. Where it cannot, the judgement and skill of stakeholders should be relied upon instead.

The final principle asserts that metrics must not only be causally relevant, they must have some bearing upon the actions of the individuals participating in the activity system. Of particular concern in this respect is the use of abstract metrics, such as Return on Management (Strassmann 1990). Adopting a high level of abstraction reduces the association between metrics and verifiable facts. In terms of guiding action - the only way in which evaluation can itself add value to a development process - a sound basis in fact is essential. Thus, when evaluating the impacts of technologies on the performance of workers, metrics must directly relate to their core activities.

6.2.3 Monitoring Activity Systems

During development, changes to an activity system necessarily change the meanings as well as the values of metrics. For evaluation to be useful at this stage, the complex of changes must somehow be delineated to determine whether the development is leading towards its objectives. Most, if not all, current evaluation methodologies provide no support for this process, assuming that meanings remain constant and the initial values of metrics, thus, only need to be kept up-to-date to guide decision making effectively. To illustrate how the meaning of a measure can change, consider a system that allows consultants to remotely examine patients with the support of an on-site nurse. Such a system would change the nature of some examinations and treatments conducted by nurses (i.e. they may perform some treatments formerly given by the examining consultant). If one continued to use the metrics from prior to the system development, both consultant and nurse examination times, for example, may appear to increase and treatment times decrease. Such changes would be difficult to understand (or believe), however, unless one is also aware of the increased cooperation between staff and the reduction in time spent by nurses attending to patients until a consultant is available. Such understanding can only be developed if a model of the activities is available to demonstrate how the actions that constitute the system interrelate. However, evaluation techniques rarely centre around the use of models, but provide frameworks or classifications for metric choice that are largely independent of the activity systems being analysed. This is in contrast with the metrics used to evaluate artefacts, such as telecommunications products and software. Although it is impossible to achieve a comparable level of precision when evaluating social organizations, the principle of using evidence-based models to construct metrics is equally valid.

6.2.4 Evaluation and Decision Making

Whilst the continuous monitoring of systems provides a basis for managing gradual change, there are some occasions where significant decisions must be made. By far the most significant are choices between alternative investments and decisions to abandon projects. A decision to invest ought to be based upon knowledge of the current activity systems in an organization, how well they are performing and the extent to which they can be improved. Such information is provided by monitoring within the organizational form. The choice involves consideration of two distinct issues: (a) which activities are most in need of improvement (or which activities need to be established to accommodate new demands); and (b) which available change options are preferred.

Problems are caused by failing to recognize both the need for change and the merits of alternative change options. Consider, for example, the use of return on investment (ROI) and hurdle rates. This approach to allocating funding to capital investments considers only the relative cost efficiency of alternative options. No consideration is given to the relative importance of the activity systems being changed. Thus, the automation of back-office administration is placed on an equal footing with production and sales, and projects for all organizational activities are discriminated on the basis of short-term cost savings or financial gains. The budgeting process implicitly recognizes the need to consider the relative importance of different activities but it is an essentially political process that is not well-related to investment evaluation techniques. Strategic approaches to evaluation (for example, Clemons 1991 and Earl 1996) pay more attention to identifying critical activities and ensuring that they are supported.

6.2.5 Long-Term Payoff

Farbey et al. (1993) argue that post-implementation review has three purposes: (a) ensuring benefits delivery; (b) identifying unexpected costs and benefits; and (c) providing feedback and, thereby, facilitating learning. Each point suggests that evaluation is best carried out for some time after implementation so that the activity system can be seen to settle following the technological and other changes. Evidence suggests, however, that, although many organizations conduct some form of post-implementation review, most evaluations are used to verify the technical conformance of the system and thereby signify the end of the development project (Ward et al. 1996).

Farbey et al. (1993) identify three types of evaluation that are most common at the post-implementation stage: (a) the comparison of expected and actual performance; (b) the assessment of user satisfaction; and (c) the monitoring of system usage. None of these provides any clear evidence for comparison with the equivalent pre-implementation working practices. Consequently, current practice is limited in addressing the above three purposes of post-implementation review.

6.2.6 Summary

As the above discussion shows, there is a need for evaluation methodologies that support the process of acquiring evidence regarding the impacts of changes to working practices resulting from IS developments. The methodologies developed should also support users in interpreting their metrics because the meanings of metrics change as a result of the development process. The discussion has further illustrated the importance of regarding evaluation as an intrinsic part of the development process. Consequently, a prerequisite to developing an evaluation methodology based around the understanding of changes to working practices is an information systems development lifecycle model that is sensitive to such changes. For this reason, several existing lifecycle models are briefly reviewed in section 6.3.1 and an alternative lifecycle model that is more suited to supporting the evaluation of working practices is developed in section 6.3.2. The lifecycle model is then used in section 6.4 to support the development of an evaluation methodology that addresses some of the issues identified above.

6.3 Social Transformation and the Information Systems Development Lifecycle

Owing to the overriding need to manage the complexity of software systems, conventional approaches to software engineering used technology-centred methods for design (see, for example, Yourdon 1972 and DeMarco 1979). These approaches embody the 'waterfall' model of the development lifecycle (Royce 1970), viewing analysis as the identification of the correct structure to encode the problem domain. Avison and Fitzgerald (1995) suggest that the software development community recognizes the need for a more user-centred approach, although development methods still leave many key design decisions under the effective control of technical experts. The empowerment of users during the development lifecycle has received attention by the Information Systems community for some years, recognizing the need for user participation in systems development and the importance of managing end-user computing. In light of progress in supporting user participation in major development efforts, Mumford (1983) proposed the following typology:

1. **Consultative participation** - Analysts discuss the system's requirements with users, but the technical experts perform design.
2. **Representative participation** - Representative users work on the design team and are, thereby, involved in decision making.
3. **Consensus participation** - Users drive the design process, making all key decisions.

An obvious trend in this list is the transfer of authority and responsibility for design decisions from technical experts to users. It has rarely been noted, however, that this transfer does not empower users unless they also have the capacity to act accordingly (see chapter 5 for a definition of this term). Furthermore, the education of users to improve their capacities for developing and using ICT-based information systems is poorly understood. It was shown in chapter 2 that, although existing approaches to systems development have considered authority, responsibility and capacity, their use of these concepts as the basis for guiding the development process has been limited. Furthermore, analyses of capacity are not acted upon during systems development to ensure that users have the information handling skills that the systems design implies. To address these weaknesses, this section sets out a case for the adoption of a user-centred information systems lifecycle. A lifecycle model based upon the socio-cognitive theory is then developed and the implications for information systems is considered.

6.3.1 Current Perspectives on the Information Systems and Software Development Lifecycles

Software engineering differs from other engineering disciplines in two key respects:

- the physical environment is viewed in terms of peripheral constraints rather than as the basic medium for development; and
- engineers are primarily concerned with systems and pay comparatively little attention to the development of general systems components.

The consequence of these differences is that software requires a fundamentally different approach to engineering. Numerous methods have been proposed but they all (with the exception of formal methods) typically result in software systems that have numerous design flaws. Quality management is being improved through the application of quality management techniques, such as the Capability Maturity Model (Herbsleb, Zubrow, Goldenson, Hayes and Paulk 1997, Ferguson and Sheard 1998), and standards, such as ISO9001 and ISO15504. Nevertheless, the systemic approach to design and the absence of architectural principles for developing software systems limits the quality and reliability that can be achieved and contrasts sharply with, for example, electronic engineering, where applications are constructed from a very small number of precisely engineered modules (for example, standard amplifier circuits).

The most widely known model of the software development lifecycle is the conventional 'waterfall' model (see section 2.3.1). Although this model is now regarded as grossly oversimplified, its influence on more recent lifecycle models is readily discerned. Two serious and well-documented deficiencies of such lifecycle models are considered here. First, the customer (who defines the system's requirements) and the users are excluded from much of the decision process. Consequently, many decisions are based upon the developer's interpretations of what the customer requires, supplemented by his or her understanding of what constitutes a user-friendly system. This limitation has been cited as a key reason for low levels of software acceptance (see, for example, Norman and Draper 1986). Second, despite the strong technical focus, the process fails to satisfy the technical criteria applied to other branches of engineering. Several alternatives to the waterfall model have been proposed as solutions to these and other development problems, the most widely used of which are incremental development, evolutionary development using prototypes, and rapid application development (RAD).

Whilst incremental development represents a simple extension to the waterfall model, the use of prototyping in evolutionary development is a substantial change (see section 2.3.1). A significant problem for many customers and users is that, although they can state their problem, they have insufficient experience of possible solutions to state their preferences between them and their contribution to design is, therefore, limited. Prototyping helps overcome this by providing users with experience of alternative software solutions and a mechanism for improving the dialogue between analyst, customer and user. Prototyping must be carefully managed, however, to avoid slowing down development or unduly increasing costs and is only useful for studying some aspects of systems (Olle et al. 1988). A critical decision when using prototyping is whether to throw away the prototype or develop it into an operational system. Throwaway prototyping raises the prospect of long development times, as with the waterfall lifecycle, whereas developing the prototype risks poor software quality and increased maintenance costs.

Rapid application development extends prototyping tools to enable the production of workable applications for certain problem domains. RAD is particularly well-suited to transaction-based processes with well-defined inputs and relatively simple representations, such as tables, for output. From the customer's and users' perspectives, RAD has several potential drawbacks that must be carefully managed. Crucially, the focus on developing database systems that can readily be modified may encourage a short-term perspective and, thereby, shorten the time between revisions of the system. This has three potential consequences: (a) systems changes are made in response to user requests without full consideration of the strategic/organizational implications; (b) the coherence and integrity of the system may deteriorate more rapidly; and (c) savings at the initial development stage are offset by increased maintenance/redevelopment activity.

As illustrated above, software engineering has had limited success in improving the quantity and quality of customer and user participation. The HCI discipline, however, regards its primary goal as the integration of user-centred techniques into software engineering (Dix et al. 1993, Sutcliffe 1995). As illustrated by the example below, HCI methods tend to augment the software engineering process with user-centred tools and techniques, rather than integrate them at the methodological level. Consequently, development remains document or risk driven, albeit with an increased awareness of customer and user needs. The "psychological and organizational tools" developed by Clegg et al. (1996) illustrate the limitations of augmenting software engineering approaches (see section 2.4.1). With the waterfall model, the tools can provide only a list of tasks to be computerised and a method for filtering out unacceptable solutions after the development process. An evolutionary prototyping methodology would resolve some of these problems by allowing usability evaluation to feed back into the task allocation and job design processes, but this would require some revision of the tools and how they are collectively applied.

In comparison with software engineering and HCI, the information systems community has proposed more radical approaches to user involvement in systems development. Lyytinen (1987), for example, presents a taxonomy of information systems development methodologies based upon three contexts: technology, organization and language. The taxonomy is used to illustrate how the different premises of methodologies lead to different perceptions of the development process. A central point in Lyytinen's analysis is the demonstration that methodologies are partial in terms of their coverage of the three contexts. The socio-technical approaches, for example, emphasise the organization and technology contexts relative to the language context. The language context relates to the capacities of users to manipulate symbolic representations, which are the skills constitutive of information systems. The technology context refers to the competence for developing and manipulating technology artefacts in order to gain access to symbolic representations and to supplement language skills. The organizational context refers to the social relations that bear upon the access, control and change of communication artefacts. A lifecycle model that uses social transformation as its central precept provides the potential for making methodologies contingent in their balancing of the three contexts proposed by Lyytinen. Given that systems development implies organizational change, a contingent approach is preferable to the use of a taxonomy to guide the selection of a methodology that is fixed in its treatment of social organization.

Iivari (Iivari 1990a, Iivari 1990b) develops a lifecycle model from a similar perspective to that adopted by Lyytinen. Extending Boehm's (Boehm 1988) spiral model of the lifecycle and drawing upon the PIOCO development methodology, Iivari presents a three-phase evolutionary lifecycle. The three phases - organizational, technical and conceptual - are essentially synonymous with the three contexts proposed by Lyytinen. By incorporating these factors into a lifecycle model, Iivari addresses several of the problems identified by Lyytinen's taxonomy, resulting in a feasible, although somewhat complex, lifecycle model that is sensitive to social change. As illustrated above, Lyytinen's contexts and, hence, Iivari's three-phase approach focus on developing software rather than user skills. Thus, although Iivari recognizes that the evolutionary development involves a considerable amount of learning on the part of the developers, no explicit mechanism is proposed for ensuring that users participating in the development process have the requisite technical and conceptual skills. Furthermore, the lifecycle retains a focus on software and does not provide any specific process to support the development of users' information handling abilities.

Iivari, Hirschheim and Klein (1998) identify and compare five approaches to IS development that are significantly different to conventional software engineering. In a similar vein to Lyytinen's taxonomy, the analysis clarifies and contrasts the basic premises underlying the approaches. Whilst this framework provides some useful insights, its analysis of ontological and epistemological positions has a number of weaknesses, the most significant being the failure to make a useful distinction between ontological and epistemological anti-realism (see Hacking 1983) and the narrow range of ontological commitments considered. The predominance of philosophical concerns in the analysis suggests that it is of academic, rather than practical, interest. Indeed, Iivari et al. (1998) conclude with a number of comments about IS as an academic discipline. If extended to address the above weaknesses, however, the framework may prove a useful starting point for analysing existing methodologies and considering how they might address issues of social transformation more thoroughly. Prerequisites for such methodological development are systems and software development lifecycles that regard social transformation as central to effective information systems development. The following section describe such a lifecycle model based upon the socio-cognitive theory of information systems.

6.3.2 Incorporating Social Transformation into the Information Systems Development Lifecycle

Social transformation is an intentional change to a social organization. In the context of information systems, such transformations must address six types of issues, which form the basis for any intentional information systems development:

1. the development of actors' capacities to represent, communicate and interpret messages and use the resulting knowledge to guide action;

2. the development of communication artefacts to supplement the capacities identified in point 1;
3. the development of users' capacities to exploit ICT artefacts;
4. the provision of sufficient access to, and control over, communication artefacts for the effective performance of tasks;
5. the alignment of power over and responsibility for information sources; and
6. the identification and resolution of deficiencies in the information available within the organizational form.

Information systems development begins when at least one stakeholder participating in an activity system perceives a deficiency in the system and subsequent analysis reveals that the deficiency can be resolved, at least in part, through changes to communication processes and information usage. Thus, as illustrated in figure 6.1, the first two stages of the lifecycle are stakeholders' perceived dissatisfaction initiating the lifecycle, followed by the analysis of the activity system to determine whether the deficiency can suitably be addressed in information systems terms.

Where the contribution of IS developments is clearly identified, the next stage of analysis is to gain a detailed understanding of the stakeholders' perceived deficiencies and to classify these according to the six types listed above. The results of this analysis of the activity system is the identification of activities and communication artefacts that need to be modified. Having identified what needs to be changed, the analysis goes on to consider who will be affected by the changes. It should be noted at this stage that past and present power relations and information needs will have led to the emergence of stable information flows and standardised communication artefacts within the activity system. Many such flows will be components of regularised working practices. Thus, unless widespread dissatisfaction is apparent, it can generally be assumed that the status quo serves the participants in the activity system reasonably well or, at least, satisfies those participants with most power. Consequently, a key issue to be addressed when identifying stakeholders affected by the proposed change is the assessment of stakeholders' power over the affected information flows and communication artefacts and their satisfaction with the existing organization. The balance of satisfaction and power is a key factor determining the feasibility of any changes to the activity system. During this analysis, it may be useful to distinguish between the following stakeholder categories:

- Current Information Users - Those participants in the information system who are presently able to access or control a communication artefact and/or its content.
- Potential Information Users - Those participants who have legitimate access to communication artefacts, but presently do not have the capacity to manipulate the artefacts and/or interpret the messages that they contain.
- Secondary Information Users - Those participants in the activity system who have power to direct the activities of any current information users with respect to the use of the information flows and communication artefacts in question; and
- Activity Relevant Stakeholders - Those persons in the activity system who will be directly or indirectly affected by changes to the information flows and communication artefacts being considered.

Given that stakeholders may themselves lack the capacities to make the requisite changes to the system, the following points should also be considered at this stage:

- Do stakeholders with the capacity and authority (i.e. legitimate power) to change communication artefacts perceive a need for change?
- Do stakeholders with the authority to make the required changes perceive the need for changes but lack the capacity to make them?
- Do stakeholders with the capacity to make the changes lack the authority to enact them?

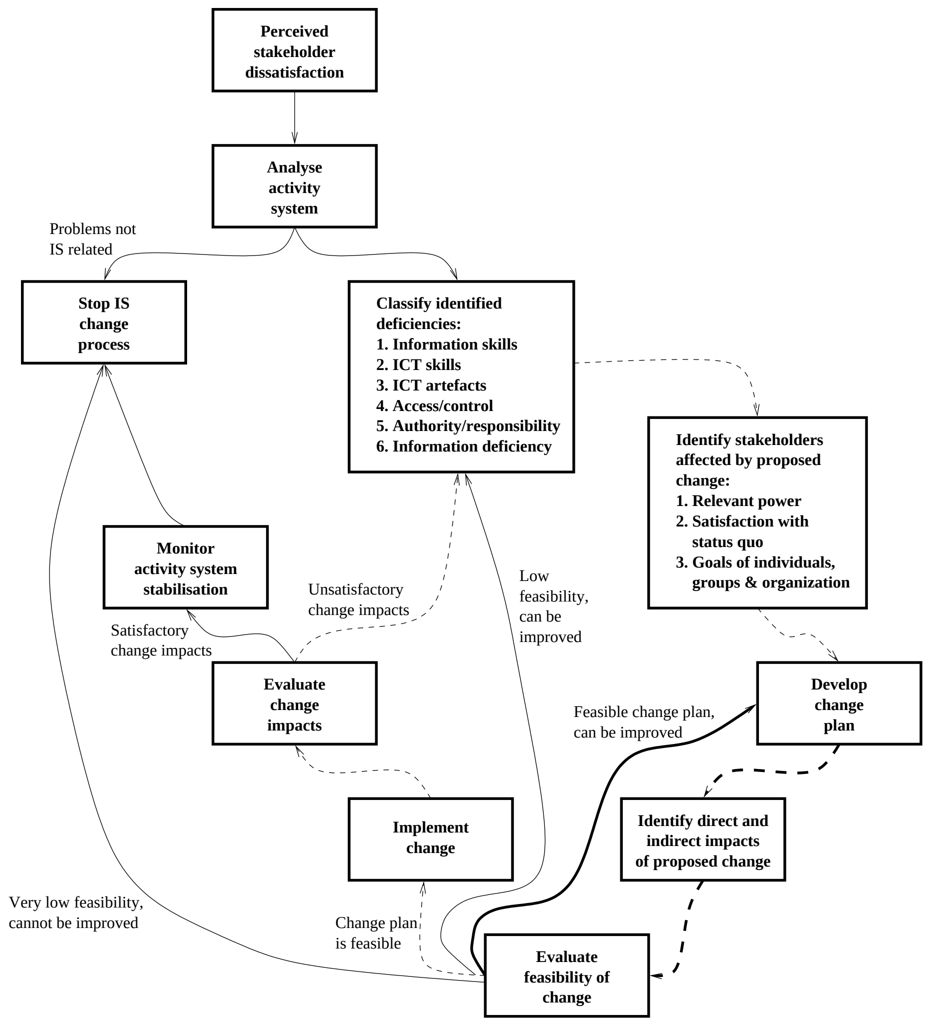


Figure 6.1: An Information Systems Lifecycle Based Upon the Socio-Cognitive Theory of Information Systems

The analysis of stakeholders in the above terms will provide a good indication of which system changes are likely to be successfully initiated. A further consideration at this stage is the extent to which the changes affect the goals of stakeholders. The analysis of goals is one of the most difficult aspects of information systems analysis because the goals of individuals, groups and organizations must be considered, yet, as illustrated by the ontology of the socio-cognitive theory of information systems, all such goals ultimately reflect the motivations of, and interactions between, individual stakeholders.

Having determined what changes are to be made and who is likely to effect, affect and be affected by the changes, the next phase of development is to plan the change process. Owing to the diversity in goals and the complex network of power relations, the planning process takes the form of a negotiation cycle, illustrated by the bold lines in figure 6.1. Accounting for the findings of the two analysis stages, an initial change plan is formulated. The direct and indirect impacts of the plan are then identified, perhaps including some consideration of the likely extent of 'ripple effects' resulting from the changes made to the activity system. The anticipated impacts are then compared with the findings of the analysis to evaluate the feasibility of the planned changes. Whilst an entire spectrum of findings may result from the evaluation, four broad types of outcome may result:

- the proposed changes are not implemented because the likelihood of success is low, the benefits are low and/or the associated risks are high;
- a second analysis phase is conducted because the proposed changes are found to be of low feasibility but are perceived to be feasible in the context of a change process that is essentially different in scope (i.e. it is realized at some stage that the perceived problem that triggered the lifecycle is not actually the central problem that needs to be addressed);
- the change plan is modified as a result of improvements identified during the evaluation of the initial plan; or
- the change plan is executed.

As illustrated by the negotiated implementation cycle (shown as a dashed line in figure 6.1), it is not possible in practice to distinguish between the negotiation of the change plan and the implementation of change. The reason for this is that the precise nature of the change and its consequences cannot be completely predicted and, consequently, any actual change to an activity system must itself take place through a negotiation process. The negotiation of implementation is shown to include further analysis precisely because the impacts of the changes need to be understood and responded to if change is to be successful (analysis at this stage is not, of course, as detailed as during the initial stages of IS development). During the process of changing the activity system, new communication artefacts may need to be introduced or existing artefacts redesigned. A guide for considering social transformation when developing communication artefacts is presented in section 6.3.3. Where the communication artefacts are ICT-based, a software development lifecycle will constitute part of the implementation cycle (Hemingway and Gough 1999a).

Following the negotiated changes to communication artefacts and information usage, the activity system will eventually stabilise and negotiations will significantly reduce, or even cease. Provided that no other significant changes to the system occur, modified processes will eventually become regularised and accepted as social norms. It should be noted that the lifecycle described above has various iterative phases. Furthermore, numerous change processes can co-occur and interact, with the initial change process sometimes stimulating other changes. Such knock-on effects can even occur at the planning stage because the discussion of systems developments can affect the satisfaction of stakeholders with other aspects of the activity system. Although the change plan has effectively been completed at this stage, it is essential that monitoring of the changes takes place whilst the system settles into a new routine. If it becomes apparent that unexpected side-effects of the change are becoming problematic, a new information systems development lifecycle may be initiated.

6.3.3 Social Transformation and the Development of Communication Artefacts

A key component of many information systems developments is the creation or modification of communication artefacts. To provide a basis for the user-centred design of such artefacts, which includes software systems, table 6.1 relates the key concepts of social transformation to three modalities that can be used to describe communication artefacts.

Communication Artefact	**Individual**	**Social**
Organization	Memory	Regularisation and standardisation
Selection	Interpretation	Communication
Navigation	Capacity	Power (Access, control, authority and responsibility)

Table 6.1: Socio-Cognitive Concepts Used in the Design of Communication Artefacts

Each row of table 6.1 permits the characterisation of communication artefacts in their present and planned future states in terms of the social transformations they will support. Considering the first row, a communication artefact can be regarded as located on a continuum. At one extreme, the artefacts are standardised to reflect well-established practices within the activity system (this is common in, for example, accounting systems). At the other extreme are flexible artefacts that can encode highly individualised messages. Typical examples include 'notes' fields in structured databases and an employee's personal records about a client. Messages encoded in unstructured artefacts are subject to broader interpretation than standardised messages and are not always well-understood by individuals outside their immediate context of use.

Considering the second row of table 6.1, communication artefacts can be located on a continuum representing the encoding and extraction of messages. The possible means for encoding and extracting information are determined by the possible modes of organization. An artefact that encodes the details of numerous customers, for example, may be sorted in a particular order (or permit certain methods of sorting). The resulting organization affects the ways in which parts of the encoded message can be encoded and extracted, and the relative efficiency of encoding and extracting messages. Of most significance, however, is whether users interact through the communication artefact or interact with it. The distinction between organization and selection can be subtle because, particularly with static artefacts, such as paper forms, the modes are closely coupled. Organization refers to what can be encoded by a communication artefact and how. Selection refers to the means available to a user for manipulating a message during its interpretation.

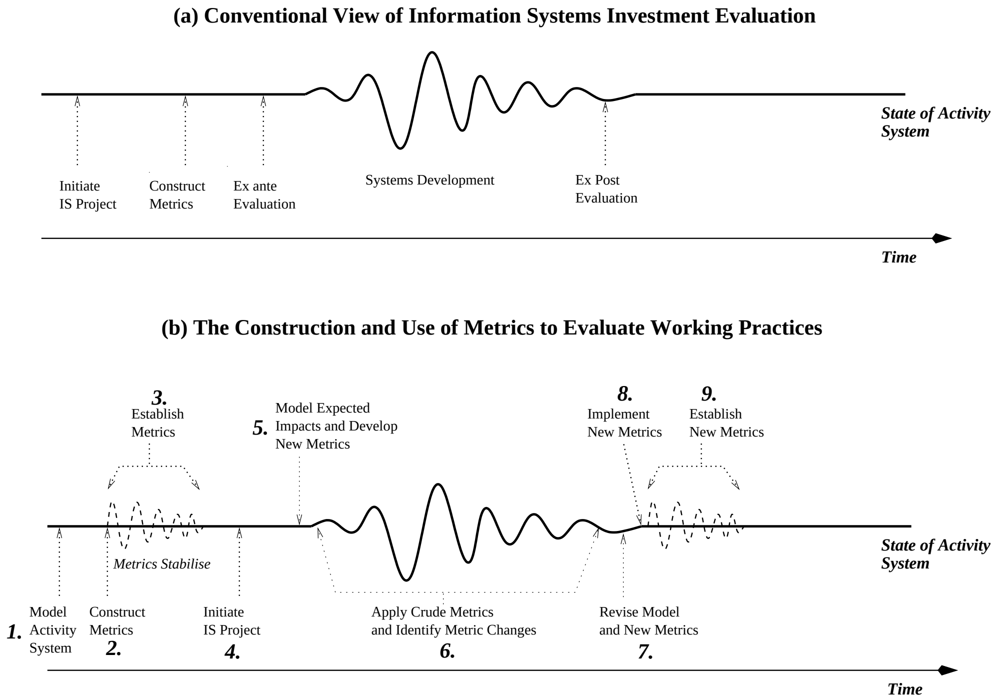


Figure 6.2: The Conventional Approach to IT Investment Evaluation Compared with the Evaluation of Working Practices

The third row of table 6.1 does not represent a continuum, but a number of factors that need to be considered to ensure the stability and effectiveness of the information system. Access to communication artefacts relevant to a task is often critical to the effective performance of the task. Control refers to the ability of the user to modify the message content of a communication artefact in order to communicate information to other users. Authority refers to the legitimate means by which a stakeholder may limit access to, and control over, communication artefacts and their message contents. Responsibility for communication artefacts and their message contents identifies the stakeholders who will be rewarded or sanctioned to reflect the quality of the communication artefact and its message contents. Much of the analysis relevant to balancing these issues in a systems context is addressed by the IS development lifecycle. In terms of developing artefacts, the main concern is providing modes of navigation by which information access and control can be managed and, particularly in terms of software systems, authority and responsibility implemented.

6.4 A Methodology for Appraising the Impacts of Information and Communication Technologies on Working Practices

A conventional view of evaluation considers two key points at which evaluations are made. Firstly, during the early stages of a project, metrics will be constructed to guide decisions concerning the type of technical change to be made. Existing benchmarks may also be used to set targets for the development. This metric construction and *ex ante* evaluation are illustrated in figure 6.2(a). Following development, the same metrics are applied again, with technical and other decision criteria used to check conformance of the implemented system with the agreed specification and business benchmarks used to assess the likely payback from the system. As noted by Ward et al. (1996), this activity also signifies the end of the development project. This activity is shown as *ex post* evaluation in figure 6.2(a). The limitations of this conventional approach have already been discussed in section 6.2. The methodology described here evaluates working practices according to the stages presented in figure 6.2(b). The correspondence between the stages of evaluation and the information systems development lifecycle (see section 6.3 and Hemingway and Gough 1999b) are shown in figure 6.3.

6.4.1 Stage 1 - Modelling the Activity System

The modelling of social organization in causal terms is useful for gaining an appreciation of what leads individuals to act in the ways they do. As noted earlier, the modelling of social activity is most reliable where specific types of activity have become regularised and discretion is limited by accepted norms, by actors' capacities and by power and authority relations. In an organizational form, it is usually possible to identify many institutionalised practices of this sort, which are referred to here as working practices. Working practices are the basic elements in systems design and development and, for the purpose of constructing metrics, are analysed here using two types of model: episodic models; and causality models.

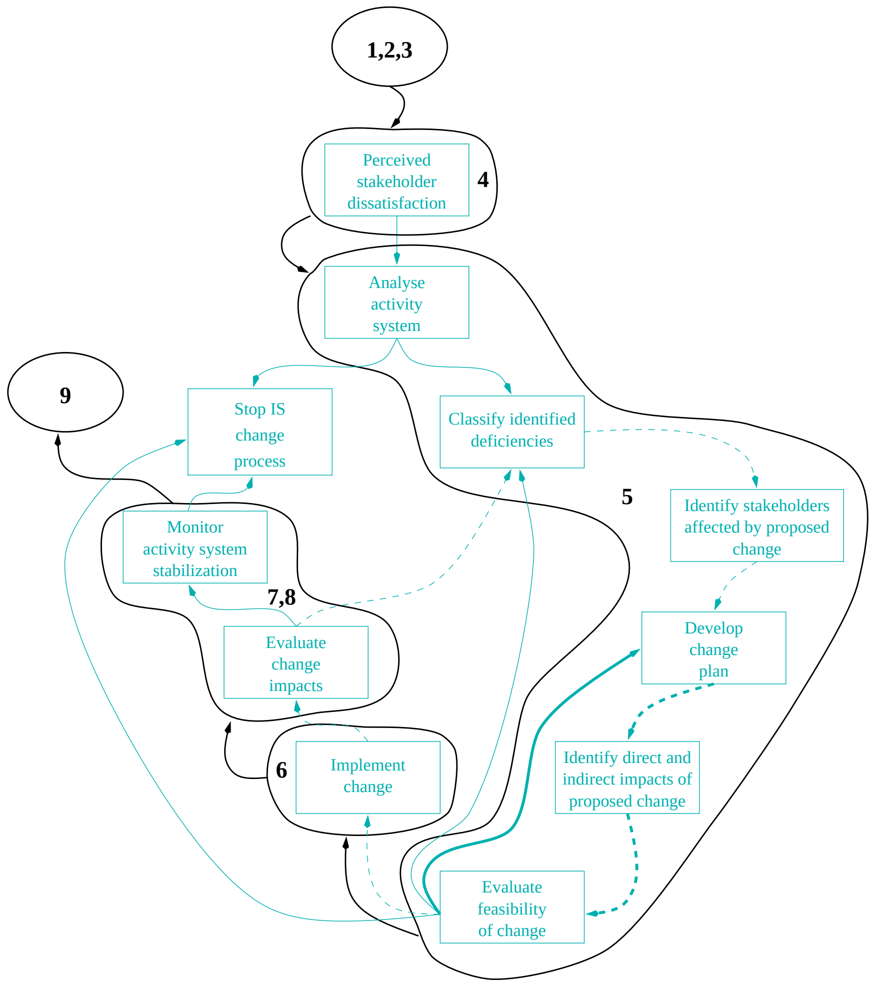


Figure 6.3: How the Evaluation of Working Practices Relates to Hemingway and Gough's Information Systems Development Lifecycle

Both types of model use the set of primitives presented in figure 6.4. Episodic models identify the key actions and information flows over time, linking specific episode types together to illustrate an overall process. In doing so, they help identify the points at which metrics can best be located. Figure 6.5 provides a simple example of episodes that might describe the nurses' and consultants' (both classified as 'carers') perceptions of treating patients in a hospital ward. Causality models identify the key causal factors that affect individuals' actions and, thereby, encourage understanding of what the metrics mean, how metrics are likely to impact upon the activities they measure and how metrics will change in meaning when the activity system is changed.

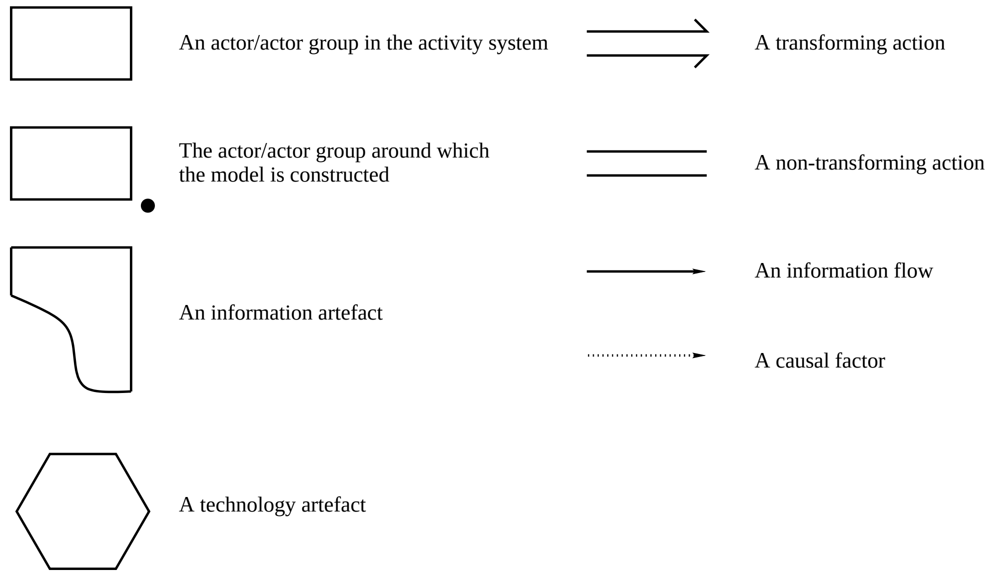


Figure 6.4: Primitives for Modelling Activity Systems to Locate and Construct Evaluation Metrics

The episodic model presented in figure 6.5 is a high-level description of two key activities in a hospital ward: examining and treating patients. Although these activities will be conducted many times in practice, the example in the diagram does not show iteration. The overall process of treating patients involves key entity types - patients, carers and patient records - and is drawn from the carers' perspective (a high-level model of this sort can often be agreed upon by a broad class of stakeholders). Both types of activity occur as brief episodes and, thus, occurrences of these activities are shown to span a short period of time. In contrast, the 'activity' internal to the patient (i.e. the cause of the illness and the response of the immune system) is shown to be continually present. The examination activity has been represented as a non-transforming activity. The purpose of a non-transforming activity is to provide a description of the current state of an entity or to provide indirect access to information about a transforming action. In this case, the examination process provides both, although the emphasis in figure 6.5 is upon the information acquired about the patient's developing condition. Treatment is a transforming action and, as indicated by the direction of the arrow, has a change impact upon the patient. As illustrated by the curved arrow, a transforming activity is an information source insofar as

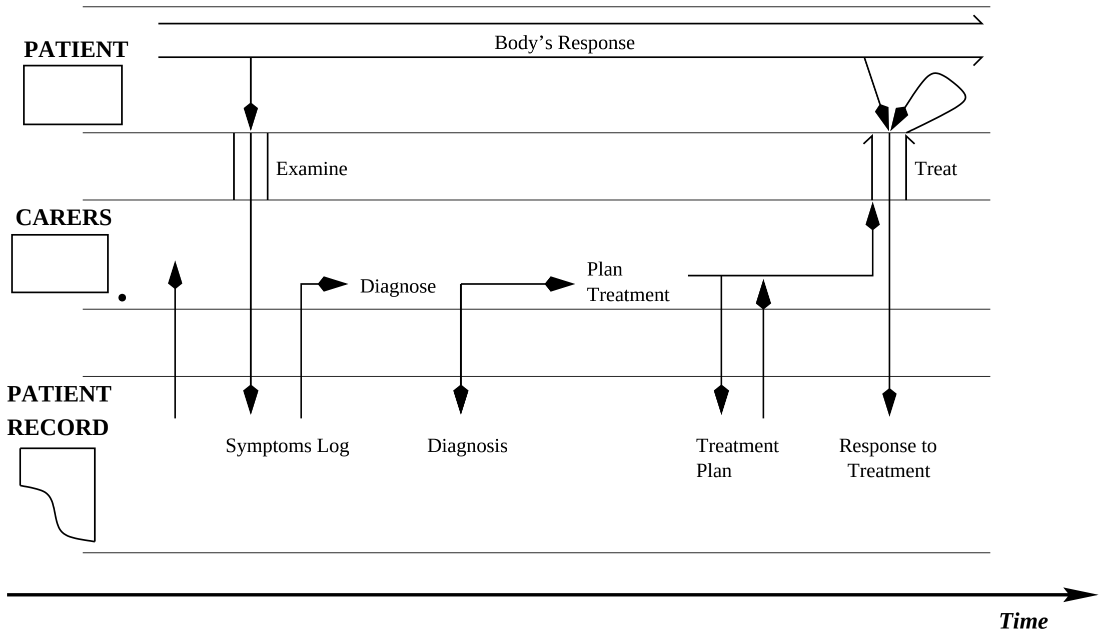


Figure 6.5: An Episodic Model of Key Working Practices in a Hospital Ward

change implies that there is something new to be known. Two information flows provide feedback from the patient to the carers in the form of a combined information flow. This indicates that carers cannot readily distinguish changes in the patient's condition resulting from the treatment and those resulting from internal changes, such as immune responses. Where information flows integrate, uncertainty necessarily ensues regarding causal relations. The remainder of the diagram documents the information flows between the carers and the main communication artefact, the patient record.

The view adopted for the causal model, shown in figure 6.6, places more emphasis on the causal factors that nurses perceive to stimulate the performance of examination and treatment activities. Figure 6.6(a) shows that different types of carer may perform examinations independently or, possibly, together. The diagram indicates a number of information flows that may lead to an examination or that may facilitate communication between actors as well as the logging of patient details on the patient record. It is important to note that the causal factors are placed within the actors to indicate that the information flows themselves do not cause the carers to examine a patient, but the interpretation of

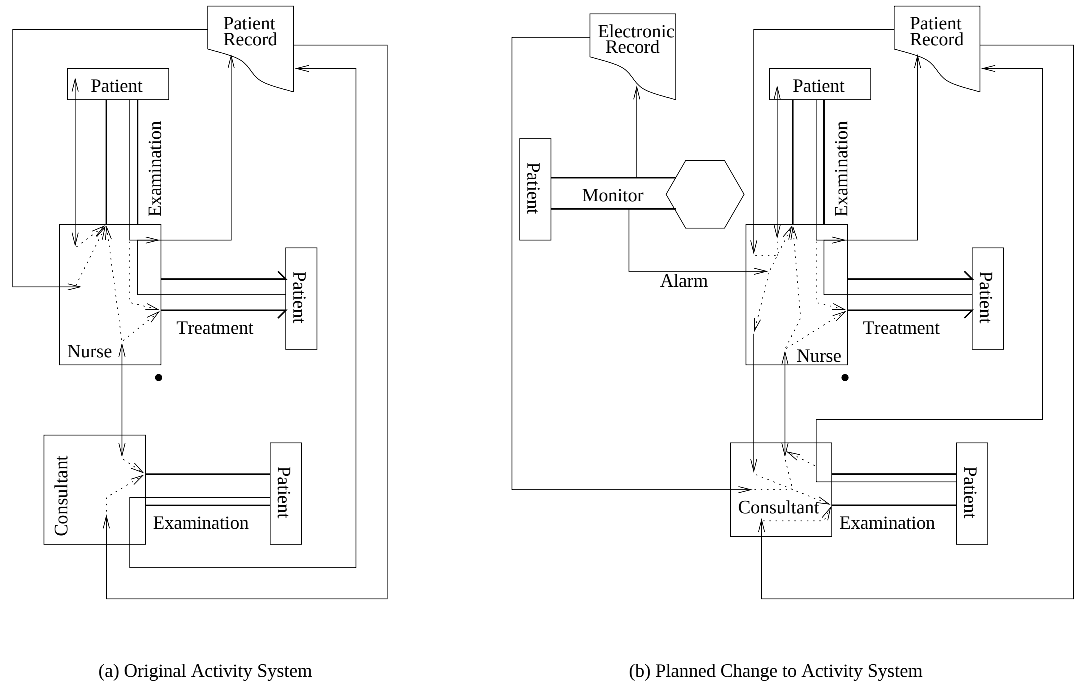


Figure 6.6: A Causal Model of Key Working Practices in a Hospital Ward

the information flows may motivate a carer to examine a patient. Figure 6.6(b) presents a possible development of the system using technology artefacts to monitor patients and provide information about exceptional conditions to the nurses. It is clear from figure 6.6 that the automated monitoring equipment results in both the nurses and consultants being required to integrate information from more information sources, even though the monitors only directly affect the nurses' conduct. The corresponding increase in causal factors has implications for the meaning and construction of metrics, which are considered below.

The value of causal models of this sort becomes most apparent when they are supplemented by further information relating to the issues identified by Hemingway and Gough's lifecycle (see section 6.3 and Hemingway and Gough 1999b):

1. Information skills

2. ICT skills

Activity: Nurse Examination	
Information Skills	Recognition of symptoms; interpretation of measurements; association of symptoms with conditions; association of conditions with treatments.
ICT Skills	Operation of monitoring equipment, especially to manipulate display; identify and correct faults on monitoring equipment.
ICT Artefacts	Patient record; monitoring equipment.
Access/control	
Authority/responsibility	Consultant has power to instruct nurse to examine patient; nurse has authority to conduct examination; nurse must respond to monitor; patient has authority to allow or prevent any examination.
Information deficiency	Patient may not communicate symptoms; failure to use a monitor to collect required information; other activities prevent routine examination.

Table 6.2: A Causal Matrix for Nurses' Examinations of Patients

3. ICT artefacts
4. Access/control
5. Authority/responsibility
6. Information deficiency

Information for each of these categories can be compiled into a causal matrix, as illustrated in table 6.2, which provides details for figure 6.6(b).

6.4.2 Stage 2 - Constructing Metrics

Although most *ex ante* evaluation methodologies begin their analyses by identifying evaluation criteria immediately prior to the change activity, this is in fact not a good time to begin evaluation. Metrics are most reliable when they are based around stable working practices and must be in place for some time for the metric users to learn how they behave and how best to act in response to evaluation results. Only then can the current activity system be effectively assessed and the impacts of possible changes reliably estimated. Thus, metric construction can be considered in three stages: (a) identifying the points in a model at which measures can best be located; (b) deciding what measurements to take; and (c) implementing and establishing metrics.

Considering the first stage, it is generally accepted that problems are best dealt with as close to their sources as possible. It follows that systems changes ought to be measured as close to the start and end of activities as possible so that problems with them are directly apparent and the causal factors leading to problems are most reliably identified. For this reason, it is generally advisable to locate measures at either side of transforming activities (accommodating the time elapsing between the intervention episode and effects becoming observable). In the case of the hospital example presented in figure 6.3, the most obvious locations for measures are before and after treatment. As in many systems involving human activity, evaluation is fundamentally limited by the presence of sources of change that are beyond both direct control and direct measurement. In this case, the body's response is a continual source of change that can be evaluated only indirectly and, moreover, continues to apply to the patient during the treatment activity. Consequently, even when measures are taken very close to the start and end of treatment episodes, there will always be some uncertainty regarding the underlying causes of observed effects. The longer the duration between start and end measurements, the greater this uncertainty will be. The only apparent way of dealing with this uncertainty is through the use of measurements during pre-treatment examinations, which describe only the effects of the body's response. Measurements taken at this point can be used to help discriminate between the body's response to the patient's condition and the actual impacts of any treatment. From an information systems development perspective, a possible development of the activity system might be to introduce such measures and develop a system for drawing upon knowledge of previous cases to support the discrimination of the changes internal to the patient from those directly resulting from the treatment.

Having identified where to locate measures in order to assess the effectiveness of change, the third stage in constructing metrics is to consider what needs to be known in order to maintain or improve the effectiveness of the system. The focus is not upon the transforming activities themselves, but upon the causal factors leading to the performance of the activity and determining the nature of its performance. The causal model presented in figure 6.6(a) is of particular value here. This model shows that the nurses perceive three key causal factors as affecting the likelihood that they will examine patients: (a) communication with the patient suggests the need for an examination; (b) the consultant communicates the need for an examination to the nurse; and (c) the nurse identifies the need for examination from his/her interpretation of the patient record. Two causal factors affect the likelihood that nurses will treat patients: (a) the results of an examination; and (b) consultants communicating the need for a particular course of treatment.

6.4.3 Stage 3 - Establishing Metrics

Metrics are only valuable if they serve as guides to action. Like any complex cues, the associations between particular evaluation results and effective courses of action take time to learn. The use of metrics as the basis for action also implies some change to the organization of the system being measured, which, in turn, requires adjustments to be made to the metrics or their underlying measures.

6.4.4 Stage 4 - Initiating an Information Systems Development Project

As indicated by the information systems development lifecycle, a development project is initiated when one or more stakeholders are dissatisfied with some aspect of the activity system in its present form. Dissatisfaction at a personal level is highly subjective and can only be acted upon if the dissatisfied person(s) has/have the power to change the system or can communicate their dissatisfaction to another. Dissatisfaction in terms of a failure to achieve "organizational goals," although nominally the responsibility of senior managers, cannot be detected in such a way because it is not directly perceived by any individual. For this reason, some means of detecting problems is required. Metrics designed to reflect the structure and purpose of an activity system are most suited to this task. Thus, by establishing metrics as intrinsic parts of an activity system, a mechanism is put in place by which managers and other actors can detect problems in achieving the system's goals and initiate developments accordingly.

6.4.5 Stage 5 - Modelling the Expected Impacts of an Information Systems Development

Ex ante evaluation has two aims: to choose between development options and to anticipate the impacts of the chosen course of action. The activity system models proposed above can be used at this stage. Different systems and working practices in an organization can be examined to inform decision makers of the areas of activity that have most scope for improvement. Expected impacts can also be modelled to help inform the choice between development options. This second use is critical as a stimulus to debate, providing a focal point for discussions about impacts on individuals/groups as well as organizational implications. Such debate will clarify the resultant change plan and help to ensure that all parties are working towards the same intended outcomes. Modelling the impacts of the chosen development option will also provide some indication of how existing measures will change in meaning and how metrics will need to be revised.

6.4.6 Stage 6 - Evaluation During the Development Process

During information systems development, the activity system is in a state of flux and existing measures cannot be relied upon to behave in the same ways that they did after their use had been established (stage 3). Furthermore, new measures and metrics cannot be effectively introduced because, as already noted, learning how to interpret and act upon metrics requires a relatively stable learning period. Thus, at this stage, metrics are of very limited value for guiding action. This is not a particular problem, however, because the change plan developed at stage 5 indicates what actions should be taken to transform the current activity system. Crude metrics may be of some value for indicating the general direction of change. In terms of evaluation, the primary objective at this stage is to make significant changes to, and then stabilise, the system.

6.4.7 Stage 7 - Revising Metrics After Implementation

As the system stabilises following a major change, the metrics documented in the change plan can be introduced and used to establish the differences between actual and intended outcomes. If the difference appears to be insufficient or adverse, a second change plan may be contemplated, although the impacts on the commitment and morale of the actors in the system and other change agents must be considered. Where the discrepancy between outcomes is acceptable, its implications for the measures and metrics proposed in the change plan must be assessed. If necessary, the measures and metrics can be revised before they are introduced into the changed activity system. The issues identified at stage 2 are relevant here because the task is essentially the same.

6.4.8 Stage 8 - Implementing New Metrics

This stage is essentially the same as stage 3, although the working practices will not be well-established. Thus, in addition to considering refining metrics, small adjustments to the activities being evaluated may be considered.

6.4.9 Stage 9 - Establishing New Metrics for Long-Term Appraisal

When considering the longer-term value of a systems development, the emphasis shifts from 'keeping the system going' whilst improving its performance, to looking at the contribution of the change to the organization's goals, which are often financial. As noted in section 6.1, this has typically been addressed by using notional costs, which raise as many problems as they appear to solve. Financial values are only meaningful in the context of exchanges under market conditions. This is demonstrated by artefacts, such as works of art, where the physical attributes have no meaningful relationship to the amount paid for them. When evaluating an activity system, the relationship is even less direct, referring to the contribution that the system makes to the value differential between inputs from, and outputs to, the market. Unless an activity system is a direct and complete route of transformation, its value cannot be directly established. The best that can be achieved is to develop a high-level causal model indicating how systems in the organizational form link together. As already noted, the accuracy of such models is fundamentally determined by the level of organization of the organizational form.

6.5 Summary

This chapter has identified two key weaknesses of current approaches to information systems development and evaluation: (a) the failure to establish causal relations between costs and benefits and (b) the incurring and perception of different costs and benefits by different stakeholders. An information systems development lifecycle based around the notion of social transformation was developed and used as the basis for an evaluation methodology to address these issues. The methodology places particular emphasis on the use of causal relations in modelling activity systems so that the meanings of the measures used are explicit. In recognition of the changes in the meanings of measures and the time taken for metrics to become established and serve as useful guides to action, the methodology advocates the development of metrics during periods of organizational stability so that preparations for managing change are always in place.

A defining characteristic of the methodology described in this chapter is the recognition that investment evaluation should utilise measures that provide descriptions of everyday activity. This is a departure from more conventional approaches to developing and utilizing evaluation measures during the change process. Only by understanding how an activity system operates on a daily basis is it possible to appreciate and measure the implications of change.

Chapter 7

Conclusions and Future Work

7.1 Reflections on the Research Approach

As explained in section 1.5, studies of human knowledge and action raise fundamental concerns about the validity of the assumptions underpinning empirical research and theory development. These philosophical difficulties are being increasingly recognized by the IS community and numerous attempts have been made to address them. A particularly significant contribution in this respect is interpretive research, which has drawn upon, for example, phenomenology (Boland 1985) and critical theory (Hirschheim 1992). This research has greatly enriched the body of empirical knowledge upon which the IS community can draw to develop its theories. The process of theory development has not, however, been as well considered and the development of substantive and well-integrated theories of information systems has been limited. In particular, there has been a tendency towards the proliferation of new terms and concepts as a consequence of the lack of progress made in clarifying the meanings and appropriate uses of basic terms. Unfortunately, the increasingly complex and varied language of IS has often increased the fragmentation of the IS community and made it more difficult to communicate or interpret research findings with precision.

In recognition of the need to clarify basic concepts and to provide a consistent and coherent theoretical basis for information systems, this research has utilized a research method based upon a materialist form of dialectic. Dialectic, at least of the sort used here, was not chosen as a research method because it is a well established method but emerged as a method from the research process itself. The analysis of IS evaluation methodologies led to the conclusion that an integrated theory was required and it was only through initial attempts at conceptual integration that the epistemological and ontological problems of explaining human knowledge and action became apparent. It was at this stage that a period of philosophical study was undertaken and dialectic was considered as a possible basis for supporting the theory development process.

As illustrated by the discussion in section 1.5, dialectic has taken a number of different forms, which have mainly been used in the humanities to address problems quite different to those motivating this research. Furthermore, many uses of dialectic are based on an idealism that is inconsistent with scientific and much psychological and social scientific research. Coming to appreciate these issues and, thereby, to develop a research method that genuinely supported the development of an integrated socio-cognitive theory was an activity that continued throughout the entire research project. Much work remains to be done in developing this type of dialectic as a research method, particularly in terms of its use to support theory development by means of empirical study. A crucial aspect of this future work will be the demonstration of how dialectic can integrate the empirical results of a range of research methods, including experiment, longitudinal case study and action research (see figure 1.2). Possible ways forward in this regard are considered in section 7.4.

An inevitable consequence of both developing and using a research method at the same time is that initial theory development is limited. In the case of this research, the most obvious limitation was the focus on conceptual analysis and synthesis and the lack of empirical study. This limitation has been beneficial in leading to a better appreciation of the problems of theory development in the IS discipline and how they can be addressed.

It is sometimes assumed that, because theory development is an inductive process, it cannot be methodical. By building upon a method for philosophical analysis, this research has shown that theory development can benefit from a methodical approach. Of equal significance, it has also been shown that the use of a dialectical approach does not prevent theory development from being a creative activity.

The dialectic, as described here and elsewhere, emphasizes the need both to rely upon and to depart from existing theory when developing new, integrated concepts and explanations. In utilizing dialectic, this research has shown that theory development requires a combination of creativity and the systematic analysis of empirical and theoretical research in the context of the researchers' historical and institutional background. In so doing, the research has indicated that the development of our knowledge of IS phenomena requires a greater understanding of the history of the many reference disciplines of Information Systems and how they have come to converge owing to the rapid uptake of technologies that are bringing into question our understanding of what knowledge is and how it can be developed and communicated. This is clearly a research issue of both scientific and historical interest, which is discussed in more detail in section 7.4.

7.2 Reflections on the Socio-Cognitive Theory of Information Systems

The socio-cognitive theory of information systems, as presented in this thesis, represents an initial attempt to integrate a broad range of concepts to explain the varied and complex phenomena of interest to the IS community. Although much more research is required before the theory will provide a detailed and rigorous basis for understanding and developing information systems, a useful start has been made in determining the scope of the theory, identifying relevant bodies of existing theory to draw upon, providing a method for integrating these theories and beginning the process of integration. This attempt at developing an integrated theory makes clear the need for IS to pay more attention to the conceptual analysis and development component of research. When conducted as a dialectic process, integrated theory development complements the diverse and rich empirical results of the scientific and interpretive research methods by facilitating dialogue about the commonalities and contradictions of multiple perspectives.

As explained in section 1.5, dialectic research demands a balanced programme of theoretical and empirical research. It also benefits from the availability of a range of existing theory that can be subject to critical analysis. Clearly, as a piece of research based around a dialectic relationship between theoretical and empirical study, the theory development in this thesis has focused more upon conceptual development. The reasons for this stem from the original motivation for this research and the time constraints placed upon it. The research originally set out to develop and test an evaluation methodology. This development stage, however, raised a number of conceptual issues that demanded attention. As explained in section 1.4, these issues concerned conflicts between different theoretical viewpoints and between concepts and existing evidence. Having identified some conceptual issues that needed to be resolved, the theorizing part of dialectic was a natural starting point. This stage progressed to yield the initial applications of the theory, presented in chapter 6. Owing to time constraints, empirical testing of the theory and initial applications are planned as future work (see section 7.4). It should be noted that the decision to apply the theory to evaluation problems was a result of two factors. First, having begun this research by studying evaluation, the researcher had sufficient familiarity and expertise in this area. Second, applying the theory to evaluation problems would help to develop its value concepts, which were considered to be weaker that most other parts of the theory. Clearly, the focus on evaluation may be reflected in a bias in the theory as presently stated. Thus, it can be expected that the theory will change in response to a broader range of study of IS issues. Such change should be regarded as a welcome contribution to the development of one general, integrated theory of information systems.

For the reasons provided in section 1.4, a key aim of this research was to develop an ontology of IS phenomena. This aim is basic to any materialist account and will remain central to any future development of the theory. It became clear from the study of many psychological concepts, such as value, that the current state of knowledge in psychology and the social sciences is not sufficient to justify many concepts that are essential to explaining information systems phenomena. These concepts were, therefore, introduced into the theory through the same process of analysis and integration/synthesis but without strict application of the materialist requirement. Although some of these concepts may never be open to convincing empirical demonstration, it is envisaged that future developments in the reference disciplines will make it possible to improve upon the existing statement of the theory presented here. It is instructive to note that the analysis of these concepts was facilitated by the dialectic method and that little discussion of alternative approaches has been found in the literature.

The presentation of the theory was based around the communication process. The impact of this approach appears to have been largely beneficial, leading to a clear conception of information systems as constituted by regularized social behaviours and intentionally constructed technology artefacts. One limitation of the process viewpoint is that it is not entirely consistent with the dialectical process described in the research method. The reasons for this are twofold. First, much of the research in the IS literature described IS phenomena in processual terms. Given that the concepts in the socio-cognitive theory are, in part, derived from existing theory, some influence of predominant themes in the literature is both inevitable and - given that they are likely to reflect empirical evidence - desirable. Second, as already noted, the consideration of dialectic emerged over the duration of the research and, consequently, did not become apparent as a processual perspective from which to consider information systems until the theory development had progressed significantly. It should be noted, however, that the consideration of information systems in processual terms is broadly consistent with the dialectic and the theory would retain approximately the same character if given the more specific interpretation. Developing such an interpretation is an interesting avenue for exploration as future research (see section 7.4).

In terms of future development of the theory, a combination of increased detail and precision is required. The theory so far is quite general in nature; a result of the integrative approach. In order to make the theory more substantive and better justified, however, it needs to be made more specific so that empirical research can be conducted with precision. This will be an evolutionary process, with different concepts in the theory addressed in turn through a combination of empirical study and conceptual analysis. Having established a general theory, a basis is in place for considering more specific theory developments in the context of explaining information systems phenomena generally. Some initial progress in extending and refining the concepts relating to information skills and the development of representations has already been made (Hemingway and Gough 1999b).

7.3 Research Contribution

The principle contribution of this research is the formulation of an integrated theory of information systems, which benefits the IS community in several ways. The attention paid to formulating an ontology and the attempts at clarifying the nature of basic IS phenomena provides an initial basis for explaining a wide range of phenomena in a coherent and consistent manner. This attempt at integration and the discussion of dialectic method shows that theory development is valuable in establishing commonalities between multiple theoretical perspectives. This is particularly important in IS and other social sciences because of the lack of methodological orthodoxy. Without persistent attempts at theoretical integration, such diverse disciplines risk becoming too fragmented and undermining progress by confusing the meanings of basic terms.

The socio-cognitive theory's emphasis on the embodied nature of cognition and the situated nature of social interaction strengthens and further develops the arguments formulated by Mingers (1996). By doing so, the theory contributes to developing an understanding of the psychological and social impacts of ICTs when used to support social activities by enabling users to interact with or through them. Without an emphasis on the physical aspects of cognition and action, the distinction between interacting with and interacting through ICTs cannot be fully understood. Further developing the theory to reflect the dialectical nature of human action and interaction will strengthen this aspect of the theory (see section 7.4).

A secondary contribution of this research is the initial application of the theory to propose a lifefycle model and evaluation methodology. These applications represent first steps in using the theory to address practical IS concerns. This sort of application and the conduct of empirical study to test the proposed theory are essential to its successful development, as indicated by the emphasis that the dialectic described in section 1.5 places on the usefulness of theory as a basis for action. Although the lifecycle model and evaluation methodology have not yet been used in practice, their potential for improving present understanding in these areas is suggested by the significant differences between them and existing approaches in IS. The lifecycle model, for example, places information and technology handling skills at the centre of the development process, rather than the development of ICT artefacts. This appears to be a sensible substitution in terms of encouraging a user-centred approach to information systems development. Although likely to change in response to its practical application, the evaluation methodology is useful in emphasizing the relationship between the meanings that metrics take on and the way in which social interaction becomes regularized. The methodology is also somewhat novel in its emphasis on using evidence more typical of systems analysis techniques than investment appraisal. Again, it is suggested that such an approach (although not necessarily in this form) may prove useful in overcoming the persistent difficulties in evaluating the 'intangible' costs and benefits of IS development (Farbey et al. 1995).

A final contribution of this research is the development of a method to support theory development. By using a form of dialectic to support this research, it has been shown that theory development and conceptual analysis are essential parts of the knowledge development process, particularly where approaches to empirical study are diverse. The application of the dialectic method was somewhat limited in this research by the need to develop the method during its application. Nevertheless, it has proven valuable in enabling theory development to take advantage of a combination of evidence presented by researchers using conventional scientific and interpretive research methods. A further limitation on the use of the dialectic was the lack of time for a detailed exploration of the various reference disciplines, their relationship to IS and the problems they all share in understanding information systems as psychosocial phenomena.

7.4 Future Work

The research presented in this thesis has been broad in scope, covering issues in philosophy, psychology, sociology and economics, as well as information systems. Moreover, the dialectical approach to research raises conceptual problems, which it cannot guarantee satisfactorily to solve. The timespan of this research has meant that a number of long-term research issues have been raised that require further study. These are now briefly considered.

The dialectic method developed during this research has proven to be valuable in supporting theoretical work in a discipline with a range of empirical research methods and that draws upon several other disciplines. The method requires further development, however, particularly in its application to support empirical research. Future research in this area would need to demonstrate that the use of dialectic to support a 'dialogue' between theory development and empirical study would make better use of the diverse empirical results in IS and not encourage the emergence of an unjustified methodological orthodoxy. A possible way forward is to use the dialectic process to underpin multidisciplinary research. This approach has the advantage of encouraging collaboration with researchers in reference disciplines, such as psychology and sociology, who have expertise in the use of established research methods and who can help to make comparisons between an integrated theory of information systems and the explanatory power of existing theory in the reference disciplines. Such a research approach is currently being explored by the author; Tom Gough of the School of Computer Studies, University of Leeds; Peter Gardner of the School of Psychology, University of Leeds; and Janice McLaughlin of the Department of Sociology and Social Policy, University of Leeds.

In order to use dialectic successfully, a sound appreciation of the history of theory development is required. In the case of information systems, this means understanding the emergence of IS as a discipline and its overlap with its various reference disciplines. In particular, such study needs to consider the relationship between the development and use of ICTs and academic study. The cognitive sciences, for example, grew out of the analogy between the brain and an information processing machine. Although there has been commentary on the IS discipline, such as Baskerville and Lee (1999) and other work of IFIP Working Group 8.2, the use of historical research methods in information systems is not widespread. This represents another avenue for future research.

The socio-cognitive theory of information systems represents only an initial step in integrated theory development. Much remains to be done in developing this theory. Given that the method for development was itself devised during the statement of the theory presented here, an initial step must be to review the theory in its present form. There also remains a great deal in the literatures of the IS and reference disciplines that could contribute to the development of the detail of the theory. An obvious next step is to test and refine particular concepts through a dialectic involving both empirical and conceptual study. Given that any empirical results are best interpreted in the light of an established research method, collaboration with researchers from the various reference disciplines would greatly strengthen such empirical testing (as the multidisciplinary research mentioned above is intended to demonstrate).

A final means of testing the theory is to examine its practical contribution through its application to current information systems problems. The lifecycle model and evaluation methodology are two such applications, which could be applied and further developed. The results of such practical application, particularly if conducted in a research context (using action research, for example), could be used in combination with conventional research results to test specific concepts and explanations provided by the theory.

A potential practical application of the theory, which the author intends to pursue, is its use to explain the interrelationships between data models, human-computer interface and users' mental models (see section 5.6.1). Using the socio-cognitive theory as a guide, it may be possible to exploit advances in analyzing the informational content of interfaces (Rollinson and Roberts 1998) to create tools that enable end-users to design and develop their own information systems with minimal technical knowledge. Such tools would advance end-user computing considerably. Studies of the organizational impacts of such tools would be particularly interesting because such technology could facilitate the construction of personal information models, which can be associated to develop intersubjective constructs. In a sense, such a system would make explicit the ontologies and beliefs of individuals in particular domains and, over time, would provide an indication of how differences in meaning are resolved. Such study, although extremely challenging in terms of research methods and the volume of evidence required, would provide insights of value to information systems, computer science, psychology, sociology and philosophy.

Bibliography

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction, *Journal of Personality and Social Psychology* **35**: 303–314.

Alchian, A. and Demsetz, H. (1972). Production, information costs and economic organization, *in* L. Putterman (ed.), *The Economic Nature of the Firm: A Reader*, Cambridge University Press, Cambridge, chapter 9.

Alt, F. (1971). On the measurability of utility, *in* J. S. Chapman et al. (eds), *Preferences, Utility and Demand: A Minnesota Symposium*, Harcourt Brace Jovanovich, New York.

American Accounting Association (1966). *A Statement of Basic Accounting Theory*, American Accounting Association, Illinois.

Anderson, J. A. and Hinton, G. E. (1989). Models of information processing in the brain, *in* G. E. Hinton and J. A. Anderson (eds), *Parallel Models of Associative Memory*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Anderson, J. A. and Mozer, M. C. (1989). Categorization and selective neurons, *in* G. E. Hinton and J. A. Anderson (eds), *Parallel Models of Associative Memory*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Anderson, J. R. (1990). *The Adaptive Character of Thought*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Anderson, J. R. (1993). *Rules of the Mind*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Anderson, J. R., Matessa, M. and Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention, *Human-Computer Interaction* **12**: 439–462.

Angell, I. O. and Smithson, S. (1990). Managing information technology: A crisis of confidence, *European Management Journal*.

Arkes, H. R. and Harkness, A. R. (1980). Effect of making a diagnosis on subsequent recognition of symptoms, *Journal of Experimental Psychology: Human Learning and Memory* **6**: 568–575.

Asch, S. E. (1946). Forming impressions of personality, *Journal of Abnormal and Social Psychology*.

Avgerou, C. (1995). Evaluating information systems by consultation and negotiation, *International Journal of Information Management* **15**(6): 427–436.

Avison, D. E. and Fitzgerald, G. (1995). *Information Systems Development: Methodologies, Techniques and Tools*, 2 edn, Blackwell Scientific Publications, Oxford.

Avison, D. E. and Wood-Harper, A. T. (1990). *Multiview: An Exploration in Information Systems Development*, Alfred Waller Ltd., Henley on Thames.

Avison, D. E., Wood-Harper, A. T., Vidgen, R. T. and Wood, J. R. G. (1998). A further exploration into information systems development: The evolution of Multiview 2, *Information Technology and People* **11**(2): 124–139.

Baber, C. (1993). Usability is useless, *Interfaces* **22**(3): 4.

Backman, C. W. (1988). The Self: A dialectical approach, *Advances in Experimental Sociology* **21**: 229–260.

Baddeley, A. D. (1986). *Working Memory*, Oxford University Press, Oxford.

Baddeley, A. D. (1990). *Human Memory: Theory and Practice*, Lawrence Erlbaum Associates, London.

Baddeley, A. D. (1999). *Essentials of Human Memory*, Psychology Press, Hove, UK.

Baddeley, A. D. and Hitch, G. (1974). Working memory, *in* G. A. Bower (ed.), *Recent Advances in Learning and Motivation*, Vol. 8, Academic Press, New York.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned at school, *Journal of Experimental Psychology: General* **113**: 1–29.

Ballantine, J. A., Galliers, R. D. and Stray, S. J. (1994). Information systems/technology investment decisions: The use of capital investment appraisal techniques in organizations, *in* A. Brown and D. S. J. Remenyi (eds), *Proceedings of the First Eurpean Conference on IT Investment Evaluation*, Operational Research Society.

Banville, C. and Landry, M. (1992). Can the field of MIS be disciplined?, *in* R. D. Galliers (ed.), *Information Systems Research: Issues, Methods and Practical Guidelines*, Blackwell Scientific Publications, Oxford, chapter 4.

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments, *Acta Psychologica* **44**: 211–233.

Bar-Hillel, M. (1982). Studies of representativeness, *in* D. Kahneman and A. Tversky (eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge, chapter 5.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency and control in social cognition, *in* R. S. Wyer and T. K. Srull (eds), *Handbook of Social Cognition, Volume 1: Basic Processes*, 2 edn, Lawrence Erlbaum Associates, Hillsdale, NJ, chapter 1.

Bargh, J. A. and Tota, M. E. (1988). Context-dependent automatic processing in depression: Accessibility of negative constructs with regard to self but not others, *Journal of Personality and Social Psychology* **54**: 925–939.

Barnard, C. (1938). *The Functions of the Executive*, Harvard University Press, Massachusetts.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories, *Journal of Experimental Psychology: Learning, Memory and Cognition* **11**: 629–654.

Baskerville, R. and Lee, A. S. (1999). Distinctions among different types of generalizing in information systems research, *in* O. Ngwenyama, L. D. Introna, M. D. Myers and J. I. DeGross (eds), *New Information Technologies in Organizational Processes: Field Studies and Theoretical Reflections on the Future of Work*, Kluwer Academic Publishers, Boston, pp. 49–65.

Bate, P. (1994). *Strategies for Cultural Change*, Butterworth Heinemann, Oxford.

Beiser, F. C. (1993). Introduction: Hegel and the problem of metaphysics, *The Cambridge Companion to Hegel*, Cambridge University Press, Cambridge, pp. 1–24.

Bennatan, E. M. (1992). *Software Project Management: A Practitioner's Approach*, McGraw-Hill, London.

Besner, D. and Davelaar, E. (1982). Basic processes in reading: Two phonological codes, *Canadian Journal of Psychology* **36**: 701–711.

Beynon-Davies, P. (1995). Information systems 'failure': The case of the London Ambulance Service's computer aided despatch project, *European Journal of Information Systems* **4**(3): 171–184.

Boehm, B. W. (1981). *Software Engineering Economics*, Prentice-Hall, chapter 33.

Boehm, B. W. (1988). A spiral model of software development and enhancement, *Computer* **21**(5): 61–72.

Boehm, B. W., Egyed, A., Kwan, J., Port, D., Shah, A. and Madachy, R. (1998). Using the WinWin Spiral Model: A case study, *Computer* **31**(7): 33–44.

Boland, R. (1985). Phenomenology: A preferred approach to research on information systems, *in* E. Mumford et al. (eds), *Research Methods in Information Systems*, North Holland, Amsterdam.

Borgida, E. and DeBono, K. G. (1989). Social hypothesis-testing and the role of expertise, *Personality and Social Psychology Bulletin* **15**: 212–221.

Boulange, B. (1998). *Software reuse*, Springer-Verlag, London.

Bower, G. H., Clark, M. C., Lesgold, A. M. and Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists, *Journal of Verbal Learning and Verbal Behavior* **8**: 323–343.

Broadbent, D. E. (1958). *Perception and Communication*, Pergamon Press, London.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress, *Academy of Management Review* **17**(1): 39–61.

Brown, J. (1968). Reciprocal facilitation and impairment of free recall, *Psychonomic Science* **10**: 41–42.

Bryson, C. and Currie, W. (1995). IT strategy: Formal rational orthodoxy or contingent adhocracy?, *Omega, International Journal of Management Science* **23**(6): 677–689.

Butler, D., Barker, S. and Levitin, L. (1996). Internet growth - Boom or bust?, *Nature* **380**(6573): 378.

Cantor, N. and Mischel, W. (1977). Traits as prototypes: Effects on recognition memory, *Journal of Personality and Social Psychology* **35**: 38–48.

Card, S. K., Moran, T. P. and Newell, A. (1983). *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Cartwright, N. (1983). *How the Laws of Physics Lie*, Clarendon Press, Oxford.

Catchpole, K., McKeown, J. D. and Withington, D. (1998). Alerting, informing, and localizable auditory warnings, *Proceedings of the Second International Conference on Engineering Psychology and Cognitive Ergonomics*, Oxford, UK.

Cavaye, A. L. M. and Christiansen, J. K. (1996). Understanding IS implementation by estimating power of subunits, *European Journal of Information Systems* **5**: 222–232.

Chan, K. C. (1998). User performance differences between relational and entity relationship models: A summary review of the literature, *Behaviour and Information Technology* **17**(1): 59–61.

Checkland, P. (1981). *Systems Thinking, Systems Practice*, John Wiley and Sons, Chichester.

Checkland, P. and Holwell, S. (1998). *Information, Systems and Information Systems: Making Sense of the Field*, John Wiley and Sons, Chichester.

Checkland, P. and Scholes, J. (1990). *Soft Systems Methodology in Action*, John Wiley and Sons, Chichester.

Chen, P. P.-S. (1976). The entity-relationship model - Towards a unified view of data, *ACM Transactions on Database Systems* **1**(1): 9–36.

Clegg, C., Coleman, P., Hornby, P., Maclaren, R., Robson, J., Carey, N. and Symon, G. (1996). Tools to incorporate some psychological and organizational issues during the development of computer-based systems, *Ergonomics* **39**(3): 482–511.

Clemons, E. K. (1991). Evaluation of strategic investments in information technology, *Communications of the ACM* **34**(1): 24–36.

Clemons, E. K. (1995). Using scenario analysis to manage the strategic risks of reengineering, *Sloan Management Review* pp. 61–71.

Coase, R. (1937). The nature of the firm, *in* L. Putterman (ed.), *The Economic Nature of the Firm: A Reader*, Cambridge University Press, Cambridge, chapter 5.

Coleman, T. (1995). IT value for money: Going beyond financial analysis, *in* B. Farbey, F. Land and D. Targett (eds), *Hard Money - Soft Outcomes*, Alfred Waller, Henley on Thames, chapter 6.

Computer Economics (1998a). Information systems budgets increase at 54% of sites, Press Release, 3 June 1998.

Computer Economics (1998b). Total worldwide information technology expense to approach $1 trillion (US) in 1999, Press Release, 20 November 1998.

Conrad, R. and Hull, A. J. (1964). Information, acoustic confusion and memory span, *British Journal of Psychology* **55**: 429–432.

Couger, J. D. (1987). Techniques for estimating system benefits, *in* R. D. Galliers (ed.), *Information Analysis*, Blackwell Scientific Publishers, chapter 21.

Craik, F. I. M. and Lockhart, R. S. (1972). Levels of processing: A framework for memory research, *Journal of Verbal Learning and Verbal Behavior* **11**: 671–684.

Craik, F. I. M. and Watkins, M. J. (1973). The role of rehearsal in short-term memory, *Journal of Verbal Learning and Verbal Behavior* **12**: 599–607.

Currie, W. L. (1989). The art of justifying new technology to top management, *Omega: International Journal of Management Science* **17**(5): 409–418.

Cyert, R. M. and March, J. G. (1992). *A Behavioral Theory of the Firm*, 2 edn, Blackwell Publishers, Oxford.

Daft, R. L. (1995). *Organization Theory and Design*, 5 edn, West Publishing Company, Minneapolis.

Damodaran, L., Simpson, A. and Wilson, P. (1980). *Designing Systems for People*, NCC Publications, Manchester.

Daniels, A. and Yeates, D. A. (1971). *Basic Training in Systems Analysis*, 2 edn, Pitman Publishing, Bath.

Davenport, T. H. and Short, J. E. (1990). The new industrial engineering: Information technology and business process redesign, *Sloan Management Review* **31**(4): 11–27.

Deci, E. L. and Ryan, R. M. (1987). The support of autonomy and the control of behavior, *Journal of Personality and Social Psychology* **53**: 1024–1037.

DeLisi, P. S. (1990). Lessons from the steel axe: Culture, technology and organizational change, *Sloan Management Review* **32**(1): 83–93.

DeMarco, T. (1979). *Structured Analysis and System Specification*, Prentice Hall, London.

deVega, M., Intons-Peterson, M. J., Johnson-Laird, P. N., Benis, M. and Marschark, M. (eds) (1996). *Models of Visuospatial Cognition*, Counterpoints: Cognition, Memory and Language, Oxford University Press, New York.

Diaz, M. and Sligo, J. (1997). How software process improvement helped Motorola, *IEEE Software* **14**(5): 75–81.

Dix, A., Finlay, J., Abowd, G. and Beale, R. (1993). *Human-Computer Interaction*, Prentice Hall, New York.

Dos Santos, B. L. (1994). Assessing the value of strategic information technology investments, *in* L. Willcocks (ed.), *Information Management: The evaluation of information systems investments*, Chapman and Hall, London, chapter 7.

Dudel, J. (1986). General sensory physiology, psychophysics, *in* R. F. Schmidt (ed.), *Fundamentals of Sensory Physiology*, Springer-Verlag, Berlin, chapter 1.

Durkheim, E. (1933). *The Division of Labor in Society*, Free Press, Glencoe, Illinois.

Earl, M. J. (1996). Information equity: A model for measuring performance in the information age, *Technical Report 2*, Center for Research in Information Management, London Business School.

Eason, K. (1988). *Information Technology and Organisational Change*, Taylor and Francis, London.

Easteal, C. and Davies, G. (1989). *Software Engineering: Analysis and Design*, McGraw-Hill, London.

Ebbinghaus, H. (1885). *Über das Gedachtris, Translated by H. Ruyer and C.E. Bussenius (1913)*, Teachers College Press, New York.

Edelheit, J. A. and Miller, M. R. (1997). Electronic commerce comes to the 'Net', *BT Technology Journal* **15**(2): 24–31.

Ejiogu, L. O. (1991). TM - A systematic methodology of software metrics, *Sigplan Notices* **26**(1): 124–132.

Ellis, N. C. and Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English, *British Journal of Psychology* **71**: 43–52.

Ernst and Young (1992). *UK GAAP: Generally Accepted Accounting Principles in the United Kingdom*, 3 edn, MacMillan, Chichester.

ESPRIT (1993). ESPRIT Project Summaries, Summary 5429.

Estes, W. K. (1980). Is human memory obsolete?, *American Scientist* **68**: 62–68.

Farbey, B., Land, F. and Targett, D. (1993). *How to Assess Your IT Investment*, Butterworth-Heinemann, Oxford.

Farbey, B., Land, F. and Targett, D. (1994). A taxonomy of evaluation methods, *in* A. Brown and D. Remenyi (eds), *Proceedings of the First European Conference on IT Investment Evaluation*, Operational Research Society, Birmingham, UK, pp. 44–70.

Farbey, B., Land, F. and Targett, D. (1995). *Hard Money, Soft Outcomes*, Alfred Waller, Henley on Thames.

Feldman, J. A. (1989). A connectionist model of memory, *in* G. E. Hinton and J. A. Anderson (eds), *Parallel Models of Associative Memory*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Ferguson, J. and Sheard, S. (1998). Leveraging your CMM efforts for IEEE/EIA 12207, *IEEE Software* **15**(5): 23–28.

Fiske, S. T. (1982). Schema-triggered affect: Applications to social perception, *in* M. S. Clark and S. T. Fiske (eds), *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition*, Vol. 1, Lawrence Erlbaum, Hillsdale, NJ, pp. 55–78.

Fiske, S. T. and Taylor, S. E. (1991). *Social Cognition*, 2 edn, McGraw-Hill, Inc., New York.

Fiske, S. T., Lau, R. R. and Smith, R. A. (1990). On the varieties and utilities of political expertise, *Social Cognition* **8**: 31–48.

Fitzgerald, B. and O'Kane, T. (1999). A longitudinal study of software process improvement, *IEEE Software* **16**(3): 37–45.

Fitzgerald, G. and Willcocks, L. (1994). Relationships in outsourcing: Contracts and partnerships, *in* W. R. J. Baets (ed.), *Proceedings of the Second European Conference on Information Systems*, Nijenrode University Press, Breukelen.

Fletcher, G. J. O. and Ward, C. (1988). Attribution theory and processes: A cross-cultural perspective, *in* M. H. Bond (ed.), *The Cross-Cultural Challenge to Social Psychology*, Sage Publications, Newbury Park, CA.

Fong, G. T., Krantz, D. H. and Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems, *Cognitive Psychology* **18**(3): 253–292.

Forster, M. (1993). Hegel's dialectical method, *in* F. C. Beiser (ed.), *The Cambridge Companion to Hegel*, Cambridge University Press, Cambridge, chapter 5.

Forsterling, F. (1985). Attributional retraining: A review, *Psychological Bulletin* **98**: 495–512.

French, S. (1986). *Decision Theory: An Introduction to the Mathematics of Rationality*, Ellis Horwood Ltd, Chichester.

Gabriel, Y. (1991). On organizational stories and myths: Why it is easier to slay a dragon than to kill a myth, *Interational Sociology* **6**(4): 427–442.

Geiselman, E. (1988). Improving eyewitness memory through mental reinstatement of context, *in* G. M. Davies and D. M. Thomson (eds), *Memory in Context: Context in Memory*, John Wiley, New York.

Giddens, A. (1984). *The Constitution of Society*, Polity Press, Cambridge.

Giddens, A. (1995). *Politics, Sociology and Social Theory*, Polity Press, chapter 5.

Gigerenzer, G. (1994). Why the distinction between single-even probabilities and frequencies is important for psychology (and vice versa), *in* G. Wright and P. Ayton (eds), *Subjective Probability*, Wiley, Chichester.

Gigerenzer, G. and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality, *Psychological Review* **103**(4): 650–669.

Godden, D. and Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water, *British Journal of Psychology* **66**: 325–331.

Goodland, M. and Slater, C. (1995). *SSADM Version 4*, McGraw-Hill, London.

Goodman, D. J. (1997). *Wireless Personal Communication Systems*, Addison-Wesley, Reading, Massachusetts.

Gough, T. G. and Hemingway, C. J. (1999). Appraising the impacts of information and communication technologies on working practices, *Technical Report 1999.08*, School of Computer Studies, University of Leeds.

Gregory, F. H. and Lau, S. P. (1999). Logical soft systems modelling for information source analysis - The case of Hong Kong Telecom, *Journal of the Operational Research Society* **50**(2): 124–147.

Gregory, R. L. (ed.) (1987). *The Oxford Companion to the Mind*, Oxford University Press, Oxford.

Guba, E. and Lincoln, Y. (1989). *Fourth Generation Evaluation*, Sage Publications, London.

Hacking, I. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge University Press, Cambridge.

Hanley, J. R. and Morris, P. E. (1987). The effects of amount of processing on recall and recognition, *Quarterly Journal of Experimental Psychology Section A - Human Experimental Psychology* **39**(3): 431–449.

Harrison, E. F. (1995). *The Managerial Decision-Making Process*, 4 edn, Houghton Mifflin Company, Boston.

Hastie, R. (1984). Causes and effects of causal attribution, *Journal of Personality and Social Psychology* **46**(1): 44–56.

Hatano, G. and Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation, *Cognition* **15**(1-3): 95–110.

Hayek, F. (1945). The use of knowledge in society, *in* L. Putterman (ed.), *The Economic Nature of the Firm: A Reader*, Cambridge University Press, Cambridge, chapter 4.

Hayes-Roth, B. and Hayes-Roth, F. (1977). Concept learning and the recognition and classification of exemplars, *Journal of Verbal Learning and Verbal Behavior* **16**: 321–338.

Haykin, S. (1994). *Communication Systems (3rd edn.)*, John Wiley and Sons, Inc., New York.

Hegel, G. W. F. (1976). *Science of Logic*, London.

Hemingway, C. J. (1997a). Nonsense on crutches: Popular misconceptions in information systems evaluation, *Technical Report 97.14*, School of Computer Studies, University of Leeds.

Hemingway, C. J. (1997b). A universal spanner in the works: Economic issues of evaluating general purpose technologies, *in* E. W. Berghout and D. S. J. Remenyi (eds), *Proceedings of the Fourth European Conference on the Evaluation of Information Technology*, Delft University Press, Delft.

Hemingway, C. J. (1999). Towards a socio-cognitive theory of information systems: An analysis of key philosophical and conceptual issues, *in* T. J. Larsen, L. Levine and J. I. DeGross (eds), *Information Systems: Current Issues and Future Changes*, International Federation for Information Processing, Laxenburg, Austria, pp. 275–286.

Hemingway, C. J. and Gough, T. G. (1999a). Incorporating social transformation into the information systems and software development lifecycles, *in* O. Ngwenyama, L. Introna, M. Myers and J. I. DeGross (eds), *New Information Technologies in Organizational Processes: Field Studies and Theoretical Reflections on the Future of Work*, Kluwer Academic Publishers, pp. 277–290.

Hemingway, C. J. and Gough, T. G. (1999b). Introducing information design skills into user-centred systems development, Under review.

Henry, S. and Kafura, D. (1984). The evaluation of software systems structure using quantitative software metrics, *Software Practice and Experience* **14**(6): 561–573.

Herbsleb, J., Zubrow, D., Goldenson, D., Hayes, W. and Paulk, M. (1997). Software quality and the capability maturity model, *Communications of the ACM* **40**(6): 30–40.

Herrmann, D. J. and Petro, S. J. (1990). Commercial memory aids, *Applied Cognitive Psychology* **4**(6): 439–450.

Hewstone, M. (1989). *Causal Attribution: From Cognitive Processes to Cognitive Beliefs*, Basil Blackwell, Oxford.

Higgins, E. T. (1987). Self discrepancy: A theory relating self and affect, *Psychological Review* **94**(3): 319–340.

Higgins, E. T. (1989). Continuities and discontinuities in self-regulatory and self-evaluative processes: A developmental theory relating self and affect, *Journal of Personality* **57**(2): 407–444.

Hilton, D. J. and Knibbs, C. S. (1988). The knowledge-structure and inductivist strategies in causal attribution, *European Journal of Social Psychology* **18**(1): 79–92.

Hilton, D. J. and Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model, *Psychological Review* **93**(1): 75–88.

Hinsz, V. B., Tindale, R. S., Nagao, D. H., Davis, J. H. and Robertson, B. A. (1988). The influence of the accuracy of individuating information on the use of base rate information in probability judgment, *Journal of Experimental Social Psychology* **24**(2): 127–145.

Hinton, G. E. and Anderson, J. A. (1989). *Parallel models of associative memory*, Lawrence Erlbaum Associates Inc.

Hirschheim, R. A. (1992). Information systems epistemology, *in* R. D. Galliers (ed.), *Information Systems Research: Issues, Methods and Practical Guidelines*, Blackwell Scientific Publications, Oxford.

Hirschheim, R. and Smithson, S. (1987). Information systems evaluation: Myth and reality, *in* R. D. Galliers (ed.), *Information Analysis*, Addison Wesley, Wokingham, chapter 20.

Hitch, G. J. and Halliday, M. S. (1983). Working memory in children, *Philosophical Transactions of the Royal Society London* **302**: 325–340.

Hitchman, S. (1995). Practitioner perceptions on the use of some semantic concepts in the entity-relationship model, *European Journal of Information Systems* **4**(1): 31–40.

Hogarth, R. M. (1986). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics, *in* H. R. Arkes and K. R. Hammond (eds), *Judgment and Decision Making: An Interdisciplinary Reader*, Cambridge University Press, Cambridge.

Horowitz, L. M. and Prytulak, L. S. (1969). Reintegrative memory, *Psychological Review* **76**: 519–531.

Hospers, J. (1956). *An Introduction to Philosophical Analysis*, Routledge and Kegan Paul Ltd., London.

Iivari, J. (1990a). Hierarchical spiral model for information system and software development, Part 1: Theoretical background, *Information Systems and Software Technology* **32**(6): 386–399.

Iivari, J. (1990b). Hierarchical spiral model for information system and software development, Part 2: Design process, *Information and Software Technology* **32**(7): 450–458.

Iivari, J., Hirschheim, R. and Klein, H. K. (1998). A paradigmatic analysis of contrasting information systems development approaches and methodologies, *Information Systems Research* **9**(2): 164–193.

Intons-Peterson, M. J. and Newsome, G. L. (1992). External memory aids: Effects and effectiveness, *in* D. Herrmann, H. Weingarner, A. Searleman and C. McEvoy (eds), *Memory Improvement: Implications for Memory Theory*, Springer-Verlag, New York.

Jacobson, I., Booch, G. and Rumbaugh, J. (1999). *The Unified Software Development Process*, Addison-Wesley, Reading, Mass.

Johnson, G. and Scholes, K. (1993). *Exploring Corporate Strategy*, 3 edn, Prentice Hall, New York.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*, Cambridge University Press, Cambridge.

Johnson-Laird, P. N., Gibbs, G. and deMowbray, J. (1978). Meaning, amount of processing and memory for words, *Memory and Cognition* **6**: 372–375.

Jones, T. and King, S. F. (1998). Flexible systems for changing organizations: Implementing RAD, *European Journal of Information Systems* **7**(1): 61–73.

Kahneman, D. and Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives, *Psychological Review* **93**(2): 136–153.

Kahneman, D. and Tversky, A. (1982a). On the psychology of prediction, *in* D. Kahneman et al. (eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Kahneman, D. and Tversky, A. (1982b). The simulation heuristic, *in* D. Kahneman et al. (eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge, chapter 14.

Kahneman, D. and Tversky, A. (1982c). Subjective probability: A judgment of representativeness, *in* D. Kahneman et al. (eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Kaplan, R. S. and Norton, D. P. (1992). The balanced scorecard - Measures that drive performance, *Harvard Business Review* **70**(1): 71–79.

Keat, R. and Urry, J. (1982). *Social Theory as Science*, Routledge and Kegan Paul, London.

Kelley, H. H. (1967). Attribution theory in social psychology, *in* D. Levine (ed.), *Nebraska Symposium on Motivation*, Vol. 15, pp. 192–240.

Kelley, H. H. (1972). Attribution in social interaction, *in* E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins and B. Weiner (eds), *Attribution: Perceiving the Causes of Behavior*, General Learning Press, Morristown, NJ, pp. 151–174.

Kelly, J. E. (1982). Economic and structural analysis of job redesign, *in* J. E. Kelly and C. Clegg (eds), *Autonomy and Control at the Workplace: Contexts for Job Redesign*, Croom Helm, London, chapter 2.

Kirveennummi, M., Hirvo, H. and Eriksson, I. (1999). Framework for barriers to IS-related change: Development and evaluation of a theoretical model, *in* T. J. Larsen, L. Levine and J. I. DeGross (eds), *Information Systems: Current Issues and Future Changes*, IFIP WG8.2 and WG8.6, International Federation for Information Processing, Laxenburg, Austria, pp. 509–528.

Klinke, R. (1986). Physiology of hearing, *in* R. F. Schmidt (ed.), *Fundamentals of Sensory Physiology*, Springer-Verlag, Berlin, chapter 6.

Kosslyn, S. M. (1980). *Image and Mind*, Harvard University Press, Cambridge, Mass.

Kosslyn, S. M. (1994). *Elements of Graph Design*, W.H. Freeman and Company, New York.

Kroll, N. E. A. (1975). Visual short-term memory, *in* D. Deutsch and J. A. Deutsch (eds), *Short-Term Memory*, Academic Press, New York.

Kuhn, K. M. and Sniezek, J. A. (1996). Confidence and uncertainty in judgmental forecasting: Differential effects of scenario presentation, *Journal of Behavioral Decision Making* **9**(4): 231–247.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*, Chicago University Press, Chicago.

Lacity, M. C. and Hirschheim, R. (1994). *Information Systems Outsourcing: Myths, Metaphors and Realities*, John Wiley and Sons, Chichester.

Laird, J. E., Newell, A. and Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence, *Artificial Intelligence* **33**(1): 1–64.

Larichev, O. I. (1984). Psychological validation of decision methods, *Journal of Applied Systems Analysis* **11**: 37–46.

Lawrence, P. R. and Lorsch, J. W. (1969). *Organizations and Environment: Managing Differentiation and Integration*, Irwin, Homewood, Illinois.

Lee, M. W., Yun, M. H., Jung, E. S. and Freivalds, A. (1997). High Touch: Ergonomics in a conceptual design process - Case studies of a remote controller and personal telephone, *International Journal of Industrial Ergonomics* **19**(3): 239–248.

Lindgaard, G. (1994). *Usability Testing and System Evaluation*, Chapman and Hall, London.

Linville, P. W. (1987). Self-complexity as a cognitive buffer against stress-related depression and illness, *Journal of Personality and Social Psychology* **52**(4): 663–676.

Ljubic, T. and Stefancic, S. (1994). Problems in the introduction of Rapid Application Development (RAD) Principles put to PRAXIS, *Proceedings of the Fourth International Conference for Information Systems Development - ISD94: Methods and Tools, Theory and Practice*, pp. 233–241.

Locksley, A. and Stangor, C. (1984). Why versus how often? Causal reasoning and the incidence of judgmental bias, *Journal of Experimental Social Psychology* **20**(5): 470–483.

Loftus, E. F. (1979). *Eyewitness Testimony*, Harvard University Press, Cambridge, MA.

Loftus, E. F. and Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory, *Journal of Verbal Learning and Verbal Behavior* **13**: 585–589.

Lord, C. G. (1987). Imagining the Self and others, *Journal of Personality and Social Psychology* **53**(3): 445–450.

Lyytinen, K. (1987). A taxonomic perspective of information systems development: Theoretical constructs and recommendations, *in* R. Boland and R. Hirschheim (eds), *Critical Issues in Information Systems Research*, Wiley, Chichester.

Macleod, M., Bowden, R. and Bevan, N. (1996). The MUSiC performance measurement method, *Technical report*, National Physical Laboratory.

Mandler, G. (1992). Toward a theory of consciousness, *in* H.-G. Geissler, S. W. Link and J. T. Townsend (eds), *Cognition, Information Processing, and Psychophysics: Basic Issues*, Scientific Psychology Series, Lawrence Erlbaum Associates, Hillsdale, NJ, chapter 3.

Mani, K. and Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions, *Memory and Cognition* pp. 181–187.

Marks, G. and Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review, *Psychological Bulletin* **102**(1): 72–90.

Markus, H., Cross, S. and Wurf, E. (1990). The role of the self-system in competence, *in* R. J. Sternberg and J. K. Jr (eds), *Competence Considered*, Yale University Press, New Haven, CT, pp. 205–225.

Markus, M. L. (1983). Power, politics and MIS implementation, *Communications of the ACM*.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W.H. Freeman and Company, New York.

Marx, K. (1867). *Capital, Translation by E. Paul and C. Paul*, Dent, London.

Marx, K. (1974). *Economic and Philosophic Manuscripts of 1844*, Progress Publishers, Moscow.

Massaro, D. W. (1988). Some criticisms of connectionist models of human performance, *Journal of Memory and Language* **27**(2): 213–234.

McArthur, L. Z. (1972). The how and what of why: Some determinants and consequences of causal attribution, *Journal of Personality and Social Psychology* **22**: 171–193.

McClelland, J. L. and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1, An account of basic findings, *Psychological Review* **88**(5): 375–407.

McGuire, W. J., McGuire, C. V. and Cheever, J. (1986). The Self in society: Effects of social contexts on the sense of Self, *British Journal of Social Psychology* **25**(3): 259–270.

Miner, J. B. (1982). *Theories of Organizational Structure and Process*, Dryden Press, Hinsdale, Illinois.

Mingers, J. C. (1996). Embodying information systems, *in* W. J. Orlikowski, G. Walsham, M. Jones and J. I. DeGross (eds), *Information Technology and Changes in Organizational Work*, Chapman and Hall, London, chapter 17.

Mintzberg, H. (1979). *The Structuring of Organizations*, Prentice-Hall, Englewood Cliffs, NJ.

Mintzberg, H. (1991). The effective organization: Forces and forms, *Sloan Management Review* **32**(2): 54–67.

Miyake, A. and Shah, P. (eds) (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, Cambridge University Press, Cambridge.

Moody, D. (1996). Graphical entity relationship models: Towards a more user understandable representation of data, *in* B. Thalheim (ed.), *Proceedings of the 15th International Conference on Conceptual Modelling*, Cottbus, Germany, pp. 227–244.

Moorhead, G. and Griffin, R. W. (1992). *Ogranizational Behavior: Managing People and Organizations*, 3 edn, Houghton Mifflin Company, Boston.

Morris, L. W. and Liebert, R. M. (1977). Relationships of cognitive and emotional components of text anxiety to physiological arousal and academic performance, *Journal of Consulting and Clinical Psychology* **35**: 332–337.

Morris, P. E. (1979). Strategies for learning and recall, *in* M. M. Gruneberg and P. E. Morris (eds), *Applied Problems in Memory*, Academic Press.

Morris, P. E. (1992). Cognition and consciousness, *The Psychologist* **5**: 3–8.

Morris, P. E., Tweedy, M. and Gruneberg, M. M. (1985). Interest, knowledge and the memorizing of soccer scores, *British Journal of Psychology* **76**(11): 365–371.

Mumford, E. (1983). *Designing Participatively: Participative Approach to Computer Systems Design*, Manchester Business School, Manchester.

Mumford, E. (1995). *Effective Systems Design and Requirements Analysis: The ETHICS Approach*, MacMillan, Basingstoke.

Mumford, E. and Weir, M. (1979). *Computer systems in work design - The ETHICS method*, Associated Business Press, London.

Neumann, P. G. (1977). Visual prototype formation with discontinuous representation of dimensions of variability, *Memory and Cognition* **5**: 187–197.

Newell, A. (1990). *Unified Theories of Cognition*, Harvard University Press, Cambridge, Massachusetts.

Newell, A. (1992). Precis of unified theories of cognition, *Behavioral and Brain Sciences* **15**(3): 425–492.

Newell, A. and Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search, *Communications of the ACM* **19**(3): 113–126.

Newell, A., Rosenbloom, P. S. and Laird, J. E. (1989). Symbolic architectures for cognition, *in* M. I. Posner (ed.), *Foundations of Cognitive Science*, MIT Press, Massachusetts, chapter 3.

Ngwenyama, O. K. and Bryson, N. (1999). Eliciting and mapping qualitative preferences to numeric rankings in group decision making, *European Journal of Operational Research* **116**(3): 487–497.

Nickerson, R. S. and Adams, M. J. (1979). Long-term memory for a common object, *Cognitive Psychology* **11**: 287–307.

Nielsen, J. (1993). *Usability Engineering*, Academic Press, Boston.

Nisbett, R. E. and Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, NJ.

Norman, D. A. (1988). *The Psychology of Everyday Things*, Basic Books, New York.

Norman, D. A. and Draper, S. W. (1986). *User centered systems design: New perspectives on human-computer interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition and typicality, *Journal of Experimental Psychology: Learning, Memory and Cognition* **14**(4): 700–708.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory, *Journal of Experimental Psychology: Learning, Memory and Cognition* **17**: 3–27.

Oakhill, J. V., Yuill, N. and Parkin, C. (1988). Memory and inference in skills and less-skilled comprehenders, *in* P. E. Morris and R. N. Sykes (eds), *Practical Aspects of Memory: Current Research and Issues*, John Wiley, Chichester, pp. 315–320.

Olle, T. W., Hagelstein, J., Macdonald, I. G., Rolland, C., Sol, H. G., van Assche, F. J. M. and Verrijn-Stuart, A. A. (1988). *Information Systems Methodologies: A Framework for Understanding*, Addison-Wesley Publishing Company, Wokingham.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations, *Organization Science* **3**(3): 398–427.

Orlikowski, W. J. and Gash, D. C. (1994). Technological frames: Making sense of information technology in organizations, *ACM Transactions on Information Systems*.

Owen, G. E. L. (1986). *Logic, Science and Dialectic: Collected Papers in Greek Philosophy*, Duckworth, London.

Palvia, S. C. and Chervany, N. L. (1995). An experimental investigation of factors influencing predicted success in DSS implementation, *Information and Management* **29**(1): 43–53.

Parker, M. M., Benson, R. J. and Trainor, H. E. (1988). *Information Economics: Linking Business Performance to Information Technology*, Prentice Hall, New Jersey.

Parkinson, J. (1991). *Making CASE work : Strategies for the successful introduction of CASE technology in commercial information systems development*, NCC Blackwell, Oxford.

Paulk, M. C. (1995). How ISO9001 compares with CMM, *IEEE Software* pp. 74–83.

Pearce, J. A. and David, F. R. (1983). A social network approach to organizational design-performance, *Academy of Management Review* pp. 315–332.

Peters, T. J. and Waterman, R. H. (1982). *In Search of Excellence*, Harper and Row, New York.

Pettigrew, A. M. (1985). *The Awakening Giant: Continuity and Change in Imperial Chemical Industries*, Blackwell, Oxford.

Popper, K. R. (1940). What is dialectic?, *Mind*.

Popper, K. R. (1972). *The Logic of Scientific Discovery*, Hutchinson and Co. Ltd., London.

Potts, T. C. (1996). The projection of semantic onto syntactic structures, *Forms of Representation: An Interdisciplinary Theme for Cognitive Science*, Intellect Books, Exeter, pp. 86–100.

Powell, P. (1992). Information technology evaluation: Is it different?, *Journal of the Operational Research Society* **43**(1): 29–42.

Pressman, R. S. (1997). *Software Engineering: A Practitioner's Approach*, 4 edn, McGraw-Hill, New York.

Price Waterhouse (1992). *The Price Waterhouse Information Technology Review 1992/93*, Price Waterhouse United Kingdom, London.

Price Waterhouse (1995). *The Price Waterhouse Information Technology Review 1995/96*, Price Waterhouse United Kingdom, London.

Puuronen, S. and Savolainen, V. (1997). Mobile information systems - Executives' views, *Information Systems Journal* **7**(1): 3–20.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge, *Pyschological Review* **88**(1): 16–45.

Pylyshyn, Z. W. (1989). Computing in cognitive science, *in* M. I. Posner (ed.), *Foundations of Cognitive Science*, The MIT Press, Cambridge, MA, chapter 2.

Raiffa, H. (1982). *The Art and Science of Negotiation*, Belknap Press, Cambridge, Mass.

Redl, S. H., Weber, M. K. and Oliphant, M. W. (1998). *GSM and Personal Communications Handbook*, Artech House.

Rollinson, S. R. and Roberts, S. A. (1998). Formalizing the informational content of database user interfaces, *in* T. W. Ling, S. Ram and M. L. Lee (eds), *Conceptual Modeling - ER'98, Proceedings of the the 17th International Conference on Conceptual Modelling*, Springer-Verlag, pp. 65–77.

Rosenberg, L., Hammer, T. and Shaw, J. (1998). Software metrics and reliability. Presented at the 9th International Symposium on Software Reliability Engineering, Germany.

Royce, W. W. (1970). Managing the development of large software systems, *Proceedings of WESTCON*, San Francisco CA.

Rumelhart, D. E. (1989). The architecture of mind: A connectionist approach, *in* M. I. Posner (ed.), *Foundations of Cognitive Science*, MIT Press, Massachusetts, chapter 4.

Rundus, D. (1973). Negative effects of using list items as recall cues, *Journal of Verbal Learning and Verbal Behavior* **12**: 43–50.

Saiedian, H. and Kuzara, R. (1995). SEI Capability Maturity Model's impact on contractors, *Computer* **28**(1): 16–26.

Salamé, P. and Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory, *Journal of Verbal Learning and Verbal Behavior* **21**: 150–164.

Savage, A. and Mingers, J. (1996). A framework for linking Soft Systems Methodology (SSM) and Jackson System Development (JSD), *Information Systems Journal* **6**(2): 109–129.

Sayers, S. (1985). *Reality and Reason: Dialectic and the Theory of Knowledge*, Blackwell, Oxford.

Schmidt, R. F. (ed.) (1986). *Fundamentals of Sensory Physiology*, 3 edn, Springer-Verlag, Berlin.

Self, P. (1970). Nonsense on stilts: Cost-benefit analysis and the Roskill Commission, *Political Quarterly* **41**(4): 249–260.

Serafeimidis, V. and Smithson, S. (1994). Evaluation of IS/IT investments: Understanding and support, *in* A. Brown and D. S. J. Remenyi (eds), *Proceedings of the First European Conference on IT Investment Evaluation*, Operational Research Society.

Serafeimidis, V. and Smithson, S. (1996). The management of change for information systems evaluation practice: Experience from a case study, *International Journal of Information Management* **16**(3): 205–217.

Shackel, B. (1986). Ergonomics in design for usability, *in* M. D. Harrison and A. F. Monk (eds), *People and Computers: Designing for Usability*, Cambridge University Press, Cambridge.

Shanks, D. R. (1991). Categorization by a connectionist network, *Journal of Experimental Psychology: Learning, Memory and Cognition* **17**: 433–443.

Shannon, C. E. and Weaver, W. (1963). *The Mathematical Theory of Communication*, University of Illinois Press, Urbana.

Shaw, M. E. (1981). *Group Dynamics: The Psychology of Small Group Behavior*, 3 edn, McGraw-Hill, New York.

Shepard, E. N. and Metzler, J. (1971). Mental rotation of three-dimensional objects, *Science* **171**: 701–703.

Simon, H. A. (1951). A formal theory of the employment relationship, *in* H. A. Simon (ed.), *Models of Bounded Rationality*, Vol. 2, MIT Press, Cambridge, MA, pp. 11–23.

Simon, H. A. (1960). *The New Science of Management Decision*, Harper and Row, New York.

Simon, H. A. (1967). The logic of heuristic decision making, *in* N. Rescher (ed.), *The Logic of Decision and Action*, Pittsburgh University Press, Pittsburgh.

Simon, H. A. (1982). A behavioral model of rational choice, *Models of Bounded Rationality: Behavioral Economics and Business Organization*, MIT Press, Cambridge, Mass., chapter 7.2.

Skowronski, J. J. and Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases, *Journal of Personality and Social Psychology* **52**(3): 689–699.

Slovic, P. and Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment, *Organizational Behavior and Human Performance* **6**: 649–744.

Smith, A. (1776). From *the wealth of nations*, *in* L. Putterman (ed.), *The Economic Nature of the Firm: A Reader*, Cambridge University Press, Cambridge, chapter 1.

Smith, E. E. (1989). Concepts and induction, *in* M. I. Posner (ed.), *Foundations of Cognitive Science*, The MIT Press, Cambridge, MA, chapter 13.

Smith, E. R. (1984). Model of social inference processes, *Psychological Review* **91**(3): 392–413.

Smith, S. M. (1988). Environmental context-dependent memory, *in* G. M. Davies and D. M. Thomson (eds), *Memory in Context: Context in Memory*, John Wiley and Sons, New York.

Smith, S. M. (1994). Theoretical principles of context-dependent memory, *in* P. E. Morris and M. Gruneberg (eds), *Theoretical Aspects of Memory*, 2 edn, Routledge, London.

Smyth, M. M., Collins, A. F., Morris, P. E. and Levy, P. (1994). *Cognition in Action*, 2 edn, Lawrence Erlbaum Associates, Hove, UK.

Sommerville, I. (1992). *Software Engineering*, 4 edn, Addison-Wesley, Wokingham.

Strassmann, P. (1990). *The Business Value of Computers*, Information Economics Press, New Canaan, Conn.

Sutcliffe, A. G. (1995). *Human-Computer Interface Design*, 2 edn, MacMillan, Basingstoke.

Symons, V. J. (1990). *Evaluation of Information Systems: Multiple Perspectives*, Phd thesis, St Edmund's College, University of Cambridge.

Symons, V. J. (1991). A review of information systems evaluation: Content, context and process, *European Journal of Information Systems* **1**(3): 205–232.

Symons, V. J. (1994). Evaluation of information systems investments: Towards multiple perspectives, *in* L. Willcocks (ed.), *Information Management: The Evaluation of Information Systems Investments*, Chapman and Hall, London, chapter 13.

Taylor, F. W. (1913). *The Principles of Scientific Management*, Harper, New York.

Taylor, S. E. and Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health, *Psychological Bulletin* **103**(2): 193–210.

Taylor, S. E. and Fiske, S. T. (1978). Salience, attention and attribution: Top of the head phenomena, *in* L. Berkowitz (ed.), *Advances in Experimental Social Psychology*, Vol. 11, Academic Press, New York, pp. 249–288.

Tetlock, P. E. and Boettger, R. (1989). Accountability: A social magnifier of the dilution effect, *Journal of Personality and Social Psychology* **57**(3): 388–398.

Thompson, S. C. (1981). Will it hurt less if I can control it?, *Psychological Bulletin* **90**(1): 89–101.

Travica, B. (1998). Information aspects of new organizational designs: Exploring the non-traditional organization, *Journal of the American Society for Information Science* **49**(13): 1224–1244.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Connecticut.

Tufte, E. R. (1990). *Envisioning Information*, Graphics Press.

Tufte, E. R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*, Graphics Press, Cheshire, Connecticut.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability, *Cognitive Psychology* **4**: 207–232.

Tversky, A. and Kahneman, D. (1981). The framing of decisions and the psychology of choice, *Science* **211**(4481): 453–458.

Tversky, A. and Kahneman, D. (1982a). Belief in the law of small numbers, *in* D. Kahneman et al. (eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Tversky, A. and Kahneman, D. (1982b). Judgments of and by representativeness, *in* D. Kahneman et al. (eds), *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge, chapter 6.

Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment, *Psychological Review* **90**(4): 293–315.

vanOrden, G. C., Pennington, B. F. and Stone, G. O. (1990). Word identification and the promise of subsymbolic psycholinguistics, *Psychological Review* **97**(4): 488–522.

Vincent, D. R. (1990). *The Information-Based Corporation: Stakeholder Economics and the Technology Investment*, Richard D. Irwin, Inc., Homewood, Illinois.

Walsham, G. (1993). *Interpreting information systems in organizations*, Wiley, Chichester.

Ward, J. (1994). A portfolio approach to evaluating information systems investments and setting priorities, *in* L. Willcocks (ed.), *Information Management: The Evaluation of Information Systems Investments*, Chapman and Hall, London, chapter 4.

Ward, J., Taylor, P. and Bond, P. (1996). Evaluation and realisation of IS/IT benefits: An empirical study of current practice, **4**(4): 214–225.

Weber, M. (1930). *The Protestant Ethic and the Spirit of Capitalism*, Allen and Unwin, Translated by Talcott Parsons, London.

Weber, M. (1947). *The theory of social and economic organization*, Oxford University Press, Oxford.

Weber, M. (1949). *The Methodology of the Social Sciences*, Free Press, Glencoe, Illinois.

Weick, K. E. (1985). The significance of corporate culture, *in* P. J. Frost et al. (eds), *Organizational Culture*, Sage, Beverly Hills.

White, P. A. (1988). Causal processing: Origins and development, *Psychological Bulletin* **104**(1): 36–53.

Willcocks, L. and Lester, S. (1994). Evaluating the feasibility of information systems investments: Recent UK evidence and new approaches, *in* L. Willcocks (ed.), *The Evaluation of Information Systems Investments*, Chapman and Hall, London, chapter 3.

Willcocks, L. and Mason, D. (1987). *Computerising Work: People, Systems Design and Workplace Relations*, Paradigm.

Willcocks, L., Lacity, M. and Fitzgerald, G. (1995). Information technology outsourcing in Europe and the USA: Assessment issues, *International Journal of Information Management* **15**(5): 331–351.

Wojtkowski, W. G. (1990). *Applications Software Programming with Fourth-Generation Languages*, Boyd and Fraser, Boston.

Wolff, J. G. (1989). The management of risk in system development: 'Project SP' and the 'New Spiral Model', *Software Engineering Journal* **4**(3): 134–142.

Wright, G. (1984). *Behavioural Decision Theory: An Introduction*, Penguin Books, Harmondsworth.

Wright, P., Holloway, C. M. and Aldrich, A. R. (1974). Attending to visual or auditory information while performing other concurrent tasks, *Quarterly Journal of Experimental Psychology* **26**: 454–463.

Wurf, E. and Markus, J. (1990). Possible selves and the psychology of personal growth, *in* D. Ozer, A. Stewart and J. Healey (eds), *Perspectives on Personality: Theory, Research and Interpersonal Dynamics*, Vol. 3, JAI Press, Greenwich, CT, pp. 39–62.

Wyer, R. S. and Carlston, D. E. (1994). The cognitive representation of persons and events, *in* J. Robert S. Wyer and T. K. Srull (eds), *Handbook of Social Cognition, Volume 1: Basic Processes*, 2 edn, Lawrence Erlbaum Associates, Hillsdale, NJ, chapter 2.

Young, P. (1983). *Power of Speech: A History of Standard Telephones and Cables 1883 - 1983*, Allen and Unwin, London.

Yourdon, E. (1972). *Design of On-Line Computer Systems*, Prentice-Hall, Englewood Cliffs.

Zuboff, S. (1988). *In the Age of the Smart Machine*, Heinemann, Oxford.

Zuckerman, M. and Feldman, L. S. (1984). Actions and occurrences in attribution theory, *Journal of Personality and Social Psychology* **46**: 541–550.