SOCIO-COGNITIVE ENGINEERING: A METHODOLOGY FOR THE DESIGN OF HUMAN-CENTRED TECHNOLOGY

M. Sharples[1], N. Jeffery[3], J.B.H. du Boulay[3], D. Teather[2], B. Teather[2], G.H. du Boulay[2,3]

[1]School of Electronic and Electrical Engineering, University of Birmingham, UK
[2]Department of Medical Statistics, De Montfort University, Leicester, UK
[3]School of Cognitive and Computing Sciences, University of Sussex, UK
[4]Institute of Neurology, London, UK
Email: rad@cogs.susx.ac.uk

Abstract

We describe a general methodology, socio-cognitive engineering, for the design of human-centred technology. It integrates software, task, knowledge and organizational engineering and has been refined and tested through a series of projects to develop computer systems to support training and professional work. In this paper we describe the methodology and illustrate its use through a project to develop a computer-based training system for neuro-radiology.

Keywords: human-centred technology; software engineering; organizational engineering; human-computer interaction; computer-based training; radiology.

1 Introduction

The benefits of designing for usability have been widely promoted (see, for example, www.useit.com). We contend that usability is a necessary but not sufficient condition for the design of good human-centred technology. Technology should also be useful, elegant and desirable (people should want to use it, rather than being compelled to do so as a condition of their learning or work).

In this paper we concentrate on the criteria of usefulness and elegance, and present a coherent methodology for the design of systems that enhance human knowledge working, decision making and learning. The paper describes the methodology of socio-cognitive engineering which aims to analyze the complex interactions between people and computer-based technology and then transform this analysis into usable, useful and elegant socio-technical systems (technology and its social context). The methodology has been developed over some twenty years, through a series of projects including: a telewriting system (McConnell & Sharples, 1983), a Writer's Assistant (Sharples, Goodlet, & Pemberton, 1992), a system to support requirements capture and design of electronics equipment, a training and decision support system for radiology (Sharples, Jeffery, Teather, Teather, & du Boulay, 1997), and technology to support personal mobile learning (Sharples, 2000). We illustrate the application of socio-cognitive engineering in a project, named MEDIATE, to develop a computer-based training and decision support system for neuroradiology (the MR Tutor). Although this paper discusses a specific application in medical computing, the methodology has been successfully applied to the design of a broad range of human centred technologies, including computer-based learning, decision support and work augmentation systems.

2 Background

Norman proposed the term 'cognitive engineering' to denote "a science of user-centred design", with the aim of understanding "the fundamental principles behind human action and performance that are relevant for the engineering principles of design" and then applying this understanding "to devise systems that are pleasant to use" (Norman, 1986). The original conception of cognitive engineering was as a means of improving the quality of interaction between an individual and technology and was thus couched in terms of cognitive psychology. Since then, computers have become increasingly important as a means of mediating communication and social interaction. There is now a growing recognition of the need for design of interactive computer systems to embrace the social and organizational aspects of human-technology interaction.

Our approach to system design builds on previous work in user-centred system design, soft systems (Checkland & Scholes, 1990), socio-technical and cooperative design (Catterall, Taylor, & Galer, 1991; Greenbaum & Kyng, 1991; Sachs, 1995) and the application of ethnography to system design (see Rogers and Bellotti (1997) for a review). The process includes some existing methods of knowledge engineering, task analysis and object-oriented design, but extends and integrates these into a coherent methodology that places equal emphasis on software, task, knowledge and organizational engineering.

Socio-cognitive engineering has a dialectical relationship to user-centred design. Users are important sources of design information and may be partners in the design process. Interviews with users can illuminate problems and breakdowns in their current work and technology, as well as mismatches between different viewpoints, such as teacher and student, or manager and employee. Users are often good at expressing preferences and choosing between competing products. They may often, but not always, be able to articulate their methods of working, the basis for decision making and the ways in which they structure and deploy knowledge and skill. They may also provide a guide to language and terminology.

But users are not always reliable informants. They may idealize their methods, describing the ways in which they would like to or have been told to work, rather than their actual practices. Although users may be able to describe their own styles and strategies, they may not be aware of how other people can perform the task differently and possibly more effectively. And basing design on a survey of user preferences can result in new technology that is simply an accumulation of features, rather than an integrated system.

Thus, socio-cognitive engineering draws on the knowledge of potential users and involves them in the design process. But it is critical of the reliability of user reports and it extends beyond individual users to give an analytic account of the cognitive processes and social interactions, the styles and strategies of working, and the language and patterns of communication, so a to form a composite picture of the human knowledge and activity. We shall use the term "actor" rather than "user" to indicate that the design may involve people who are affected by the new technology but are not direct users of it (for example, patients, surgeons, general practitioners, hospital administrators, and technicians in a system for radiologists)

Socio-cognitive engineering has similarities to contextual design (Beyer & Holtzblatt, 1998) in its approach of studying human activities and work practices in order to support them with new technology. Contextual design is aimed at defining customer-centred systems that are based on a sound understanding of how people work. Our aim is to define human-centred systems that are based a sound understanding of how people think, learn, perceive, work and interact.

3 Overview of the Approach

Figure 1 gives a picture of the flow and main products of the design process. It is in two main stages: a stage of activity analysis that sets constraints on the design and analyses how people work and interact with their current tools and technologies; and a stage of design of new technology.

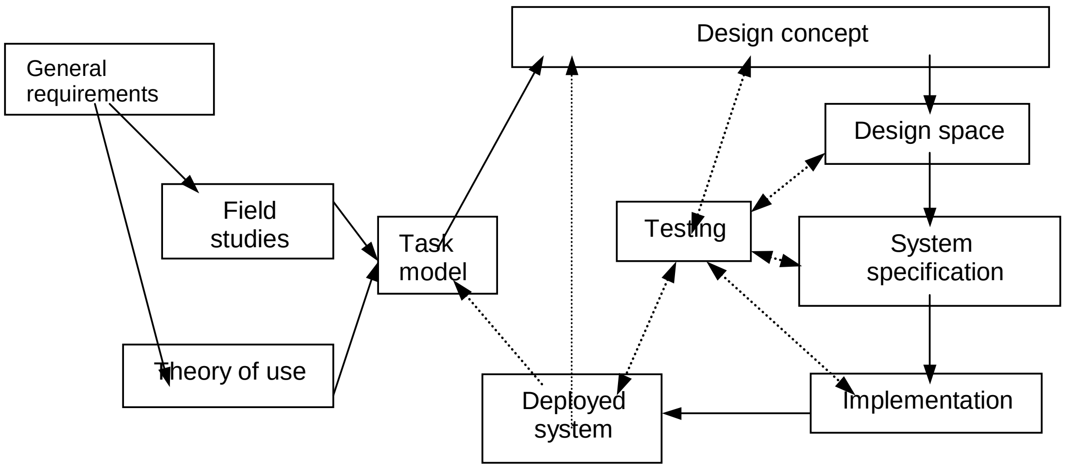

Figure 1. Overview of the flow and main products of the design process.

The process starts by specifying the general requirements and constraints for the system to be designed. This sets out the type of activities to be supported by the new technology (such as learning and knowledge management), the general domain (such as radiology) and any general constraints (such as time and budget available for the system design). This leads to two parallel studies, an investigation into how the activities are performed in their normal contexts, and a more theory-based study of the underlying cognitive and social processes. The outcomes of these two studies are synthesized into a task model. The aim of the task model is to provide a coherent account of how the activities are performed, the people involved, their contexts, the tools and technologies they employ, the structure of the tasks and an account of their cognitive processes, management of knowledge, and social interactions.

The main purpose of the task model is not to create a hierarchical analysis of the task structure or to model of the mental states and operations of the principal actors (though it might include both of these), but to describe the interactions between the people and their tools and resources, and to analyze how people externalize their work, through representations such as notes and diagrams, the rules and conventions that influence the activity, and the terminology and patterns of discourse.

The task model provides the bridge to a cycle of iterative design that includes: specifying a design concept; generating a space of possible system designs; specifying the functional and non-functional aspects of the system; implementing and deploying the system. Testing is an integral part of the design process, with the results of the evaluation being fed forwards to provide an understanding of how to deploy and implement the system, and backwards to assist in fixing bugs and improving the design choices. Although this stage is based on a conventional process of interactive systems design (see (Newman & Lamming, 1995) for an overview), it gives equal emphasis to cognitive and organizational factors as well as task and software specifications.

The result of the process is a new socio-technical system consisting of new technology and its associated documentation and proposed methods of use. When this is deployed, in the workplace, home, or other location it will not only produce bugs and limitations that need to be addressed, but also engender new patterns of work and social and organizational structures. These become contexts for further analysis and design.

The flow of design activity in Figure 1 is an aid to project planning, but it is not sufficiently detailed to show all the design activities nor does it make clear that to construct a successful integrated system requires the designers to integrate software engineering with design for human cognition and social interaction. The “building block” diagram in Table 1 gives a more detailed picture of the design process.

The four ‘pillars’ indicate the main processes of software, task, knowledge and organizational engineering. Each ‘brick’ in the diagram shows one design activity, but it is not necessary to build systematically from the bottom up. A design team may work on one ‘pillar’, such as knowledge engineering, up to the stage of system requirements, or they may develop an early prototype based on a detailed task analysis but without a systematic approach to software engineering. How each activity is carried out depends on the particular subject domain, actors and contexts of use.

	Software Engineering	*Task Engineering*	*Knowledge Engineering*	*Organizational Engineering*
Maintain	Installed system	New task structure	Augmented knowledge	New organizational structure
Evaluate	Debugging	Usability	Conceptual change, skill development	Organizational change
Integrate	Prototype System			
Implement	Prototypes, Documentation	Interfaces, Cognitive tools	Knowledge representation	Communications, Network resources
Design	Algorithms and heuristics	Human-computer interaction	Domain map, user model	Socio-technical system
Interpret	Task Model			
Analyze	Requirements	Tasks: goals, objects, methods	Knowledge: concepts, skills	Workplace: practices, interactions
Survey	Existing systems	Conventional task structures and processes	Domain knowledge	Organizational structures and schedules
Propose	General Requirements			

Table 1. A “building block” framework for socio-cognitive system design.

The design activities are modular, allowing the designer to select one or more methods of conducting the activity, according to the problem and domain. For example, the usability evaluation could include an appropriate selection of methods for assessing usability, such as heuristic evaluation or cognitive walkthrough, or it could include an evaluation designed for the particular domain. It should be noted that although the diagram shows a distinct phase of system evaluation, there will also be a continual process of testing, to verify and validate the design.

One aspect of the diagram that should be emphasized is that the blocks are not fixed entities. As each level of the system is developed and deployed is will affect the levels below, (for example, building a prototype system may lead to revising the documentation or re-evaluating the human-computer interaction; deploying the system will create new workplace practices). These changes need to be analyzed and supported through a combination of new technology and new work practices. Thus, the building blocks must be revisited both individually to analyze and update the technology in use, and through a larger process of iterative re-design.

4 Application: Computer-based Training in Neuroradiology

The domain in which we demonstrate the approach of socio-cognitive engineering is the design of a knowledge-based training system for neuroradiology. MEDIATE is a joint project between the University of Sussex, De Montfort University, Leicester, the University of Birmingham, and the Institute of Neurology, London. Its aim is to address a need identified by the radiology specialty for a more structured approach to reporting and training. One outcome of the project is the MR Tutor, a training system to assist radiologists in interpreting Magnetic Resonance (MR) images of the brain, particularly images presenting diseases that are acknowledged to be difficult to differentiate. A prototype system has been developed in the Java programming language.

The MR Tutor combines an active tutor with a responsive support system. It is based on a constructivist approach to learning, in which the trainee is helped to acquire a well-structured method for describing abnormal features of images by engaging in an active process of case-based image description and diagnosis. The MEDIATE team has devised a structured image description language (IDL) for MR neuro-images. The language is image-based, covering the visual appearance and anatomical locations of the image abnormalities. An archive of around 1200 cases, fully described using terms of the IDL and accompanied by clinical information and confirmed diagnoses, forms the knowledge base of the Tutor. At present some 40 of these cases have been included in the prototype system and we have recently developed a semi-automated method of transferring further cases and their descriptions.

The MEDIATE team has followed the socio-cognitive engineering methodology to develop a prototype training system. This paper gives an overview of the design process as it has been applied to the MR Tutor.

5 An Overview of the Design process for the MR Tutor

5.1 General Requirements

The general requirements for the MR Tutor are to develop a computer-based system that can support trainee specialists in neuro-radiology to develop a structured and reliable approach to radiological description and diagnosis. The need for a more systematic approach to the teaching of radiology is widely acknowledged (Royal College of Radiologists, 1995; Towle, 1991), to replace the traditional mixture of ad hoc apprenticeship and formal lectures with a curriculum based on a combination of structured tuition and case-based experiential learning. Computer-enhanced reporting, whereby radiologists describe images by means of a structured notation for abnormal image features supplemented by a computer-generated diagnosis, has been shown (for mammography) to improve the diagnostic accuracy of general radiologists to equal that of specialists (Getty, Pickett, D'Orsi, & Swets, 1988).

The prime objective of the MR Tutor is train beginning specialists in the use of a structured language for describing abnormal magnetic resonance images (MRI) of the brain and to support their application of this language to new diagnostic problems. The Image Description Language (IDL) has been developed over fifteen years by a senior neuroradiologist and its power has been demonstrated for key diagnostic tasks (du Boulay, Teather, Teather, Jeffery et al., 1994; du Boulay, Teather, Teather, Santosh, & Best, 1994). It provides a set of some 180 terms (such as "lesion shape: lentiform"; "lesion margin: graded"; "lesion position: pituitary fossa") to describe abnormalities in the appearance MRI images. The terms of the IDL are supported by precise definitions and, for lesion shape, by indicative examples.

The expectation is that implementing the IDL in a computer-based training system will enable trainees to develop an appropriate and consistent method of description, and to discuss findings with colleagues in a standardized and carefully defined terminology.

5.2 Theory of Use

It is essential that a theory of use should not consist of a vague collection of desirable human activities ("constructivist approach to learning"; "culturally situated" etc.) but should provide a clear guide to system design. Thus, it must be relevant to the system domain, be based on well-founded models of human cognition and social interaction and must provide specific requirements in a form that can be interpreted by system designers.

To design a computer-based training system requires an understanding of two different but overlapping sets of activities: the activities of teaching, learning and skill development, and the activities involved in carrying out the task being trained (in this case neuro-radiological description and diagnosis). There is an extensive literature on the cognitive, social and cultural practices of teaching and training, but less understanding of how to train and develop professional skills. The existing literature does, however, provide a consistent and detailed account of factors affecting the development of professional skills (Eraut, 1994) (such as the need to reconcile making rapid decisions based on incomplete information, with the need to account for those decisions to colleagues, clients and the law) and of methods of radiology training (Azevedo & Lajoie, 1995).

There is a similarly useful literature on radiological reporting. Lesgold et al. (1988) indicate that radiologists carry out a multi-stage process of interpretation. On first seeing a film they automatically invoke a mental schema that covers the salient abnormal features, resulting in one or more tentative diagnoses. This triggers a process of active search for other cues in the image along with case data and medical knowledge that might constrain the interpretation. Lastly, they articulate their findings as a verbal report. A more recent study (Rogers, 1996) indicates that perception and problem solving are tightly coupled, with experts creating an active mapping between their perception of the visual image and their evolving hypotheses about competing diagnoses.

5.3 Field studies

The aim of carrying out field studies is to uncover how people interact with technology in their normal contexts of work (Beyer & Holtzblatt, 1998). For the MR Tutor system, a central issue is how the new technology might fit into the daily schedule of trainee radiologists, so as to support their everyday practices of learning and case reporting. The theory of use provides a good indication of the knowledge and skills required for radiological reporting, the cognitive processes that underlie medical image interpretation, and the activities of radiology teaching. There are, however, pieces missing from the picture, particularly in our understanding of how radiology is taught in practice, the perspectives of the different participants (trainee, trainer, curriculum designer), and any differences between the practices of radiology in the United States (where the published studies were carried out) and the United Kingdom. These are mainly concerned with the organizational engineering 'pillar' of the building block framework, involving workplace studies of radiology training and reporting, and this section addresses that aspect.

Questions to be addressed through the workplace studies included: What would be the most appropriate configuration and location for a training system (portable, or based in the reporting room, or library)? To what levels of general or specialist medical training would the system be suited? How does reporting of cases fit into other demands on the time of a trainee radiologist? Is there a standard method of teaching radiological reporting, within a hospital and across hospitals? What do trainees and consultants perceive to be the main issues and problem areas in conventional training? Similar questions face designers of technology for other areas of training and decision support and they often remain unanswered until the system is adopted for use, when the designers and adopters discover unforeseen problems such as a lack of integration with the curriculum or an inappropriate location for the equipment.

During the project we have developed a multi-level approach to workplace studies, based on the work of Plowman et al. (1995), whose purpose is to inform the design and deployment of new technology in the workplace. It is descriptive, but is also concerned with the consequences of intervention (not just how the workplace is, but how it might be). It is designed for situations where there may be limited access to the workplace, to experts, and to potential users.

The role of the fieldworker is both to interpret workplace activity and to assist technology design and organizational change. This addresses the widely recognized problem of ethnographic approaches that, while they can provide an understanding of current work practices, they are not intended to explore the consequences of socio-technical change.

Table 2 shows a multi-level structure for workplace studies, with level 1 consisting of a survey of the existing organizational structures and schedules, levels 2 and 3 providing an analysis of situated practices and interactions of those for whom the technology is intended, and level 4 offering a synthesis of the findings in terms of designs for new socio-technical systems.

The four levels of the approach give an overview of workplace activity leading to more detailed investigation of particular problem areas, with each level illuminating the situated practices, and also providing a set of issues to be addressed for the next level. These piece together into a composite picture of how people interact with technology in their working lives, the limitations of existing work practices, and ways in which they could be improved by new technology.

Level 1	*Work structures and schedules*
Activity:	study timetables, organizational structures, syllabuses, resources
Purpose:	to discover how workplace activity is supposed to be conducted
Outcome:	description of the existing organizational and workplace structures; identification of significant events
Level 2	*Significant events*
Activity:	observe representative formal and informal meetings and forms of communication
Purpose:	to discover how workplace activity, communication, and social interaction is conducted in practice
Outcome:	a description and analysis of events that might be important to system design; identification of mismatches between how work has been scheduled and how it is has been observed to happen
Level 3	*Conceptions and conflicts*
Activity:	conduct interviews with participants to discuss areas of work needing support, breakdowns, issues, differences in conception
Purpose:	to determine differing conceptions of work; uncover issues of concern in relation to new technology; explore mismatches between what is perceived to happen and what has been observed
Outcome:	issues in working life and interactions with existing technology that could be addressed by new technology and working practices
Level 4	*Determining designs*
Activity:	elicitation of requirements; design space mapping; formative evaluation of prototypes
Purpose:	to develop new system designs
Outcome:	prototype technologies and recommendations for deployment

Table 2. Multi-level approach to workplace studies.

Our own studies of radiological reporting and training have included detailed observations and interviews at a general radiological training institution and a specialist neuroradiological training centre. The studies were carried out over a two year period and involved radiologists of all levels of ability from first year radiology specialist to consultant.

Level 1 consisted of an analysis of training materials and curricula supplemented by interviews with course directors from the main radiology training institutions in the UK, in order to identify aims, standards and significant teaching events;

Level 2 involved observations of representative formal and informal teaching and reporting events at the general hospital (7 visits) and specialist unit (4 visits).

Level 3 involved structured interviews with eight trainee specialist radiologists and a radiology lecturer at the two institutions to clarify issues raised during the Level 2 observations, including the trainees' knowledge of anatomy, the relationship between radiological reporting and patient management, and mismatches between curriculum aims and observed practice.

Level 4 was an evaluation of aspects of an initial prototype training system in terms of the situated practice of radiology training uncovered from the previous studies, and elicitation of requirements for future versions.

In general, our investigations supported previous accounts of radiological reporting as consisting of rapid perceptual analysis followed by hypothesis-driven reasoning. They also confirmed the main teaching events identified by Azevedo and Lajoie (1995), of guiding junior residents systematically through the interpretation process and of scaffolding the interpretation of more senior trainees by offering hints and indications of where to direct attention.

The studies raised issues not previously identified in the literature, concerned with the practice of training and reporting situated in the workplace. These include how computer-based training might be included in the timetable, the embedding of reporting in patient management, and the role of case notes in supporting a diagnosis.

Our initial analysis of training syllabuses and timetables had indicated that self study formed a large part of the specialists' overall training, but the interviews indicated that they have little time for self study. Whereas the timetable allocated one dedicated day per week of self study plus 2 to 3 hours per night, the trainees in practice undertook just 1 or 2 hours a week of self-directed learning.

An important issue raised by the observations is the need to consider the process of patient management in the design of radiological training. Knowing how to interpret and describe a case is only part of the skill of radiology. A radiologist works within the context of uses to which the report will be put, and the inter-relationship between referring clinicians, radiologists and surgeons has a significant effect on the nature of the reporting. This has implications for the design of technology to support and train radiologists. If the task of radiology is embedded within a process of patient management, then can a system that presents decontextualized cases provide an appropriate training? If not, then can it be designed to provide sufficient richness of context?

In summary, our investigations have indicated that a computer system to support the training of radiologists should:

– base the training on a large library of cases representative of radiological practice;
– provide a means of making rapid comparisons between cases by similarity of diagnostically relevant features;
– expose the trainee to cases in an order that promotes understanding and retention;
– help the trainee to make rapid, accurate initial judgements;
– help the trainee to integrate fragmentary knowledge into more general structural schemata;
– help the trainee to reflect on experience gained and to integrate general and situated knowledge;
– be implemented on a portable computer, for use as part of self study at home or work.

5.4 Task Model

A task model draws together the theory of use and the field studies into a composite picture of the task performance, cognitive processes and workplace activities of the people who will interact, directly and indirectly, with the new technology. It is not simply a collection of findings from the previous studies, but an analysis of the activity system and interactions with external media relevant to the system design. What is important to the design of new human-centred technology is not so much the cognitive states and process of the actors, but how the actors interact with each other and with their external tools and media.

The task model is a crucial link between the stages of work analysis and system design. It provides system designers with a description of the task structure and human activity. It identifies difficulties and breakdowns in performance of everyday activity that might be overcome by introducing new technology. It provides a shared source of reference for everyone involved in the project to uncover misconceptions or differences in viewpoint, not only between the system designers and its direct users, but also between the different types of people affected by the introduction of the technology, such as teachers and students, or managers and employees. Lastly, when the new technology is implemented and deployed it creates new socio-technical systems, providing new ways of working, thinking and interacting. A comparison of the previous task model with an account of the new activities, cognitions and interactions can give an understanding of the effects and benefits of introducing the new technology.

For the MR Tutor, the main actors are trainers and trainee radiologists at different levels of expertise, but other perspectives include patients, general practitioners, surgeons and others who contribute towards the patient record or who read and act on the radiology report. These people interact in a complex activity system, within and outside the hospital, with the case conference being a major focus of interaction. Their resources and media include MRI scanners, light boxes, patient records and informal notes.

The task model for the MR Tutor is further complicated by the general requirement to move from the current ad hoc and often inconsistent styles of reporting towards a more structured approach, so there is a tension between describing current practices and analyzing the conduct and performance of the new approaches to structured reporting. The IDL and procedures for structured image description are the result of a long process of validation, using statistical techniques to validate the accuracy of the terminology and user trials to assess its ease of use. We are currently undertaking a series of knowledge engineering studies to explore radiologists' understanding of disease categories and uncertainty in diagnosis.

The software, task, and knowledge engineering aspects of the task model are described in (Sharples, du Boulay, Jeffery, Teather, & Teather, 1996; Sharples et al., 1995; Sharples et al., 1997; Teather, Sharples, Jeffery, Teather, & du Boulay, 1994). In summary, the task model provides an analytic account of:

– The main actors (radiology trainers and trainees) and their activity systems (including training sessions, reporting sessions, case conferences and the process of patient management);

– How the actors employ tools and resources (such as lightboxes and case notes) to mediate their interaction and to externalise cognition;

– How the actors represent knowledge to themselves and others (including their language, their conceptions of MR images and their understanding of the distribution of cases across diseases, uncertainty in diagnosis and diagnostic probabilities);

– The methods and techniques that the actors employ, including differences in approach and strategy (for the MR Tutor, this includes the methods of structured image description and diagnosis with the IDL);

– The contexts in which the work occurs (both the formal settings for training and other locations where trainees work and interact);

– The actors' conceptions of their work, including sources of difficulty and breakdown in activity and their attitudes towards the introduction of new technology.

5.5 Design concept

The design concept is a translation of the Task Model into a coherent design for new technology. Whereas the Task Model is an analytic account of existing practices (or trials of new approaches), the design concept provides a system image and style of interaction. The 'desktop' is the most familiar system image for computer systems, but others include the 'electronic whiteboard' and the 'pocket organiser'.

Producing a design concept is a creative process that involves understanding the representations that the main actors currently employ to externalize and mediate their understanding, and also developing new methods of representing the structure, process and content of the task. For the MR Tutor, the three main types of representation are the MR Images, the image descriptions and diagnoses, and the distribution of cases across diseases. These form the main components of the interface.

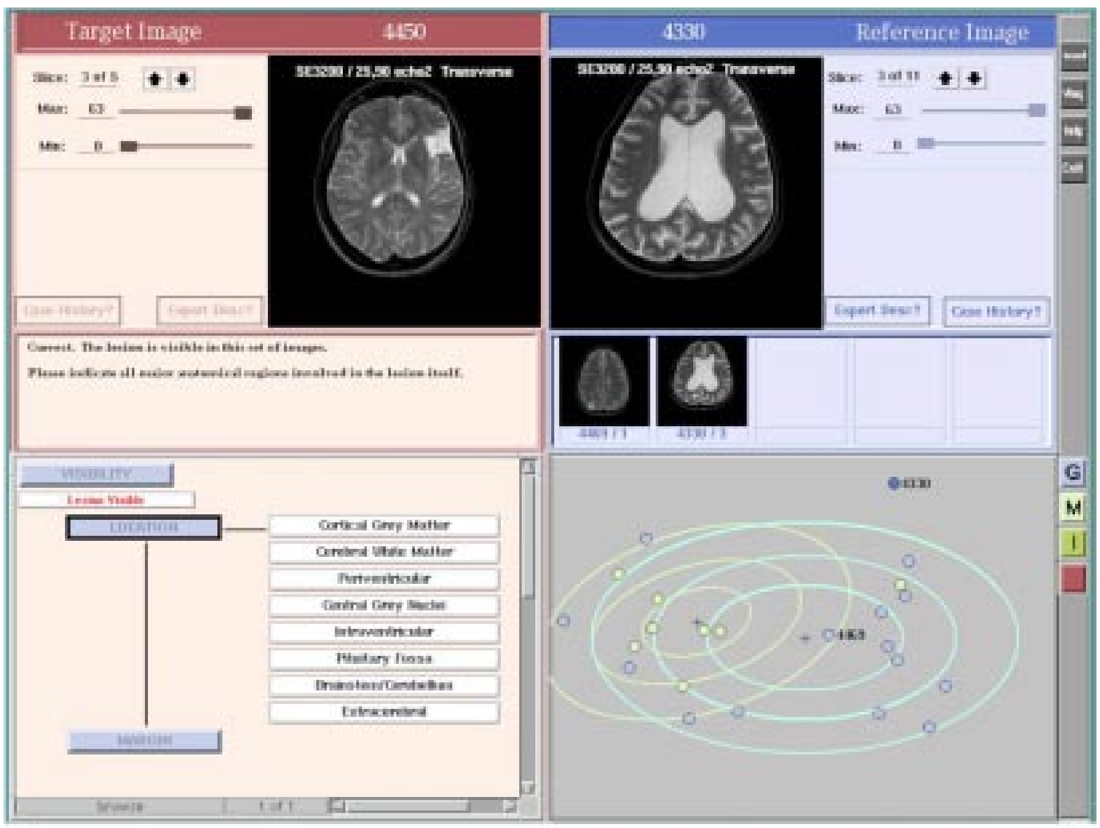

Figure 2. The interface to the initial prototype MR Tutor

Figure 2 shows the interface to one version of the prototype MR Tutor. The right side of the screen allows the trainee to view and interact with the image archive through an interactive 'overview plot' which provides a direct representation of typicality, similarity and disease membership of the cases. It is computed automatically from the structured descriptions of the cases and enables a trainee to view the distribution of cases within and across diseases. The plot can be viewed as a 'radar screen' with each point representing a case from the archive. The small crosses represent the centres of disease categories and the ellipses show scaled 'contours of typicality' for the categories. The closer a case point lies to the centre of a disease category, the more typical it is of that disease (in terms of the abnormal image features it shares with other members of that disease category). The closer any two cases are to each other, the more similar are their abnormal features. Thus, a trainee can see directly whether any case is typical or atypical of a disease, which are the most similar cases, and whether an outlier (atypical) case belongs unambiguously to one disease, or if it lies in the space between two or more disease regions and thus may be difficult to diagnose.

The left half of the screen is the interface to the tutoring component, based on the terminology of the Image Description Language. The trainee describes the target case by selecting terms from pop-up menus of feature descriptors and the system provides a tutorial response to the selection.

Across the top of the screen are the images associated with the case to be tutored and reference cases from the archive. We are currently designing an additional 'lightbox' view that will mimic and extend a conventional light box for viewing film images, so that the images associated with each case can either be placed in a stack or side by side, for viewing and comparison.

5.6 Design Space

The design concept provides a coherent metaphor and interface design for the system, but does not determine the methods of interaction, nor how the metaphor might be adapted to different types of user (such as novice, intermediate and expert trainees). The design space is a detailed exploration of options for the design of the interface and interaction, by means of a visual map of design issues, alternatives and decisions.

There are well developed methods of generating and displaying design spaces, including IBIS (Issue Based Information System) (Conklin & Yakemovic, 1991; Rittel & Kunz, 1970) and QOC (Questions, Options, Criteria) (MacLean, Young, Bellotti, & Moran, 1991). The basic approach is to pose each design issue as a question (such as "how should the tutorial responses be displayed to the trainee?") and the designers attempt to generate many alternative options in answer (for example, "text caption", "pop up response box", "computer generated speech") with arguments for and against each option (e.g. "for text caption: always available on screen"; "against text caption: takes up screen space; trainee often does not read the text"). Each argument should be justified in terms of principles in human-computer interaction, theories of human cognition and social interaction, or results of empirical studies. Finally, one option for each issue is chosen for implementation (or in a few important cases where there is no justified way to choose, two or more options are implemented and compared through user testing).

The design space is not only a way of recording design options and decisions, it is also a 'conceptual playground', showing the structure of the design space, prompting designers to add more issues and options to under-explored areas of the design, and providing a means of altering unsuccessful prototypes by reviewing the design options. IBIS and QOC provide a notation to display the design space on paper or screen as a labeled graph.

Design issues for the MR Tutor include how to provide tutorial feedback to the trainees, how to present the distribution of cases across diseases, how to provide a familiar yet powerful 'lightbox' mode, how to represent and provide guidance on sources of uncertainty in description and diagnosis, and how to provide levels of assistance that match the abilities of users. We are exploring these issues through a combination of discussion with expert neuroradiologists, experimental studies of radiologists' representations of uncertainty, and testing of prototype interfaces with trainee and expert radiologists (for example, to determine whether, by assessments of the similarity to existing cases, they are able to place cases on the overview plot in positions that match those produced by the statistical plot).

5.7 System Specification, Implementation, Deployment and Testing

The remainder of the process of socio-cognitive engineering follows a standard method for interactive system design (Newman & Lamming, 1995) with system specification leading to a series of prototype systems that are then deployed for use in the workplace. We have employed the OVID method for design of object-oriented systems developed by IBM (Roberts, Berry, Isensee, & Mullaly, 1998). This specifies the system in terms of Objects, Views, Tasks and Interactions, using the formal notation of the Unified Modeling Language (UML). The Object Models are formal representations of users' concepts; Views are those aspects of the system that the users need to see and interact with; Tasks are representations of how the users interact with objects through the views; and Interactions provide a map of the system states and actions that change state. The OVID diagrams can be used to generate object-oriented code by employing a UML CASE (Computer-Assisted Software Engineering) tool such as Rational Rose.

The design of an interactive system may involve many cycles of specification, implementation, testing and redesign to produce a system that meets the requirements and is robust, usable and appropriate to the different actors' needs and contexts of work. Testing is carried out at all stages of the design: to ensure that the task model matches the observed activities and the perceptions of the different actors; to create a design concept that is clear, consistent, elegant and matches the requirements and task model; to choose appropriate options from the design space; to check that the system specification is complete and appropriate; to ensure that the prototype implementations meet the general requirements and conform to the system specification; and to asses the usability, usefulness, and cost effectiveness of the deployed system. The results of tests during the first stage of design can be carried forward to inform decisions about how the system should be implemented and deployed. Results from later evaluations are fed back to the design of new versions.

There are also larger cycles of implementation, evaluation, reflection and redesign. When a complex piece of human-centred technology is deployed it will affect, sometimes in profound ways, how people work and interact. Often these new activities and systems are unforeseen and arise from people exploiting the affordances and opportunities of the new technology in ways that meet their personal needs or desires. A good example is the Simple Messaging System (SMS) of mobile phones. This was added, largely as an afterthought, by the designers of early mobile phones to enable users to send email or pager messages. As mobile phones became first an everyday device and then a fashion accessory, so use of the SMS grew, particularly by teenagers. Mobile phone companies are now developing and marketing new phones with enhanced SMS facilities such as larger screens and message icons to appeal to this new activity and market. Thus, the deployment and use of a system will lead to a re-examination of the task model and to new concepts for design.

The MR Tutor is still at the implementation stage, but testing of early prototypes suggests that teaching and supporting radiologists in a structured approach to describing images could have a deep effect on the quality of their reporting. It provides them with a shared language, of precisely defined terms for describing and exchanging radiological findings, supported by an archive of reference cases for comparison and shared discussion. This presents two fundamental questions: how will learning the methods of structured image description affect the cognitive processes and workplace practices of trainee radiologists; and will these new processes and practices improve the quality of their reporting compared to radiologists taught by existing methods? These questions cannot simply be answered by performing a comparative assessment of the performance of trainees who have used the MR Tutor compared to those trained by traditional means. Learning a new, computer-supported, approach to radiological reporting is a long-term process of conceptual and organizational change, as the new technology is trialled, debated, refined and deployed.

Even if it is adopted, there will be a long period during which the old and new training methods run together with much overlap (and potential confusion) between the terminology and methods. During this period the technology and the methods will be reviewed and refined, with no clear point at which two definitive approaches or technologies can be compared. It is generally the case with socio-cognitive engineering that there is no single outcome, but rather a continuing process of analysis, design, implementation, deployment, further analysis and refinement.

6 Conclusions

The methodology of socio-cognitive engineering has itself undergone a continuing process of design, analysis and refinement. The early projects resulted in prototype systems (Cyclops, The Writer's Assistant, CORECT) that were innovative and suited to the working methods and cognitive processes of the potential users. They also provided insights into human cognition and activity, for example in the form of an account of the process of creative writing (Sharples, 1999) and reports to client companies on their workplace practices.

In recent projects we have extended the methodology to provide clearer guidelines for analysis and design, incorporating industry standard methods and notations such as OVID and UML. We are teaching it to students at undergraduate and postgraduate level, and it has been employed by a successful undergraduate Group Design Project to develop personal mobile technology for lifelong learning (Sharples, 2000).

For the MR Tutor project, we have reached the stage of a working prototype system, implemented in Java. Initial formative evaluations of the system have indicated that is it usable and meets the general requirements. More generally, the studies have identified and clarified issues of importance to the design of computer-based training in radiology, including:

- the need to place computer-based training within a process of patient management. The teaching sequence and the presentation of individual cases should ideally take account of the content and quality of the case notes, and the recipients and purposes for which the radiological report will be used.
- providing a system that can be used within the limited and unpredictable time available for self-study, indicating the need for software implemented on a portable computer with the facility to rapidly save and resume a session;
- the need for a system that is authoritative, and perceived to be so by trainees and consultants. In general, radiologists have less confidence in the quality of computer-based material than information provided in book form.

This paper has proposed a methodology for socio-cognitive engineering, as a contribution towards a science of interactive system design. The methodology has been tested and refined through a series of projects over twenty years, and is being taught to undergraduate and postgraduate students as a practical method for the design of human-centred technology.

Acknowledgements

The workplace studies were funded by grant L127251035 from the Cognitive Engineering Initiative of the UK Economic and Social Research Council. Current work on developing the MR Tutor is funded through grants from the Engineering and Physical Sciences Research Council, GR/L 53588 and GR/L94598. The authors wish to thank Leicester Royal Infirmary and the Institute of Neurology, London for their cooperation with the studies. The methodology of socio-cognitive engineering has been developed through the contributions of many people, including Lyn Pemberton, James Goodlet, Lydia Plowman and Ian Rogers.

References

Azevedo, R., & Lajoie, S. P. (1995). *Learning styles underlying radiological expertise* (Project Safari Report 005). Montreal: McGill University.

Beyer, H., & Holtzblatt, K. (1998). *Contextual Design: Defining Customer-Centred Systems*. San Francisco, California: Morgan Kaufmann Publishers Inc.

Catterall, B. J., Taylor, B. C., & Galer, M. D. (1991). The HUFIT planning, analysis and specification toolset: Human factors as a normal part of the IT product design processing. In J. Karat (Ed.), *Taking Software Design Seriously*. London: Academic Press.

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

Conklin, J. E., & Yakemovic, K. C. B. (1991). A process-oriented approach to design rationale. *Human-Computer Interaction, 6*(3&4), 357-391.

du Boulay, G. H., Teather, B. A., Teather, D., Jeffery, N. P., Higgott, M. A., & Plummer, D. (1994). Discriminating multiple sclerosis from other diseases of similar presentation – can a formal description approach help? *Rivista di Neuroradiologia*(7), 37 – 45.

du Boulay, G. H., Teather, B. A., Teather, D., Santosh, C., & Best, J. (1994). Structured reporting of MRI of the head in HIV. *Neuroradiology, 37*(S), 144.

Eraut, M. (1994). *Developing Professional Knowledge and Competence*. London: The Falmer Press.

Getty, D. J., Pickett, R. M., D'Orsi, C. J., & Swets, J. A. (1988). Enhanced Interpretation of Diagnostic Images. *Investigative Radiology, 23*(4), 240-252.

Greenbaum, J., & Kyng, M. (Eds.). (1991). *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Lesgold, A., Robinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: diagnosing X-ray pictures. In M. Chi & R. Glaser & M. J. Farr (Eds.), *The Nature of Expertise* (pp. 311-342). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

MacLean, A., Young, R. M., Bellotti, V. M. E., & Moran, T. P. (1991). Questions, options, and criteria: elements of design space analysis. *Human-Computer Interaction, 6*(3&4), 201-250.

McConnell, D., & Sharples, M. (1983). Distance Teaching by CYCLOPS: an Educational Evaluation of the Open University's Telewriting System. *British Journal of Educational Technology, 14*(2), 109-126.

Newman, W. M., & Lamming, M. G. (1995). *Interactive System Design*.: Addison Wesley.

Norman, D. A. (1986). Cognitive Engineering. In D. A. Norman & S. W. Draper (Eds.), *User Centred System Design* (pp. 31-61). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Plowman, L., Rogers, Y., & Ramage, M. (1995). What are workplace studies for?, *Proceedings of European Conference on Computer-Supported Cooperative Work* (pp. 309-324): Kluwer.

Rittel, H., & Kunz, W. (1970). *Issues as elements of information systems* (Working paper 131): Institut für Grundlagen der Planung I.A., University of Stuttgart.

Roberts, D., Berry, D., Isensee, S., & Mullaly, J. (1998). *Designing for the User with OVID: Bridging User Interface Design and Software Engineering*. Indianapolis, IN: Macmillan Technical Publishing.

Rogers, E. (1996). A study of visual reasoning in medical diagnosis. In G. W. Cottrell (Ed.), *Proceedings of Eighteenth Annual Conference of the Cognitive Science Society* (pp. 213-218): Mahwah, NJ: Lawrence Erlbaum.

Rogers, Y., & Bellotti, V. (1997). Grounding blue-sky research: How can ethnography help? *Interactions, May - June 1997,* 58-63.

Royal College of Radiologists. (1995). *Structured Training in Clinical Radiology*. London: Education Board of the Faculty of Clinical Radiology, Royal College of Radiologists.

Sachs, P. (1995). Transforming work: collaboration, learning and design. *Communications of the ACM, 38*(9), 36-44.

Sharples, M. (1999). *How We Write: Writing as Creative Design*. London: Routledge.

Sharples, M. (2000). The Design of Personal Mobile Technologies for Lifelong Learning. *Computers and Education, 34*(177-193).

Sharples, M., du Boulay, B., Jeffery, N., Teather, D., & Teather, B. (1996). Interactive Display of Typicality and Similarity Using Multiple Correspondence, *Proceedings of HCI 96 Conference on Human-Computer Interaction* (pp. 162 - 167). London.

Sharples, M., du Boulay, B., Teather, D., Teather, B. A., Jeffery, N., & du Boulay, G. H. (1995). The MR Tutor: Computer-based Training and Professional Practice, *Proceedings of World Conference on Artificial Intelligence and Education (AI-ED '95)* (pp. 429 – 436). Washington DC.

Sharples, M., Goodlet, J., & Pemberton, L. (1992). Developing a Writer's Assistant. In J. Hartley (Ed.), *Technology and Writing: Readings in the Psychology of Written Communication* (pp. 209–220). London: Jessica Kingsley.

Sharples, M., Jeffery, N., Teather, D., Teather, B., & du Boulay, G. (1997). A Socio-Cognitive Engineering Approach to the Development of a Knowledge-based Training System for Neuroradiology, *Proceedings of World Conference on Artificial Intelligence in Education (AI-ED '97)* (pp. 402 – 409).

Teather, B. A., Sharples, M., Jeffery, N., Teather, D., & du Boulay, B. (1994). Statistical Modelling and Structured Image Description for Intelligent Tutoring in MR Imaging of the Head. *Rivista di Neuroradiologia, 7*, 29–35.

Towle, A. (1991). *Critical Thinking: The Future of Undergraduate Medical Education* (A Study by the King's Fund Centre in Collaboration with St Bartholomew's Hospital Medical Centre). London: King's Fund Centre.