UNIVERSITY OF SOUTHAMPTON

From Classroom Tutor to Hypermedia Advisor: A Case Study in Medical Education

by

Michael William Kemp

A thesis submitted for the degree of

Doctor of Philosophy

In the

Faculty of Engineering and Applied Science

Department of Electronics and Computer Science

June 2000

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
DEPARTMENT OF ELECTONICS AND COMPUTER SCIENCE
Doctor of Philosophy

From Classroom Tutor to Hypertext Advisor:
A Case Study in Medical Education

by Michael William Kemp

History is abundant with predictions about the influence of new technologies on educational practice. Reality however has seen traditional attitudes resolute in resisting such change. Proponents of extreme changes to education subscribe to an education ethos of no teachers and schools: a society that denounces the traditional epistemological principals grounded in the belief that intellectual progress consists of moving from the concrete to the abstract, and embraces a more situated way of learning.

With the continuing rise of staff-student ratios, a new impetus is fuelling the imaginations of policy makers and practitioners who are looking towards modern technologies to ensure the ongoing pedagogic quality of education at all levels. In today's educational technology arena, hypermedia is being hailed as the latest technology to revolutionise educational practices.

Informed by modern educational theory and the latest innovations in technology, this thesis contributes to the scientific domain of educational technology by presenting a case study of replacing undergraduate medical students' traditional pathology laboratory practicals with a hypermedia learning environment. Central to this instantiation is functionality that facilitates knowledge misconception resolve through student-tutor and peer dialogue within an open hypermedia learning environment.

Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to thank Wendy Hall and William Roche for their supervision, their constant encouragement and support, and for believing in me. Thanks also to Wendy for securing part sponsorship for my research.

I would also like the thank the members of the Multimedia Research Group for their support and for providing a conducive working environment. Thanks also to Steve Rake for many informative discussions on intelligent tutoring systems. An excellent researcher's sounding board, Steve is inspiring, critical, supportive and to the point.

A word of thanks to university pathology staff and the hundreds of medical students who participated in the evaluations presented in this volume. I would also like to acknowledge and thank Lesley Millard from the Medical Education Unit for part sponsoring this research, and Dick Bacon and the SToMP consortium for offering the use of their MCQ Engine.

Also worthy of a mention are those who have directly or otherwise helped shape my career. Ian Perkins, Mark Bolton and Gail Glencross were so influential in encouraging me to make the transition from industry to academia. They provided the push while Eric Cooke of the Multimedia Research Group provided the pull. And once again I would like to thank Wendy for being so influential in facilitating my move from undergraduate to postgraduate education.

Special thanks to my family for their support and for believing in me. Without them this work would not have been possible.

A very special thanks goes to James and Amanda. Thanks to James for understanding why I have been a student for so many years, during which time I often neglected his youth. And Amanda, for giving me the reasons for this work, and for whom my love and appreciation could never be explained by any documented eloquence. Without you, there would be no point.

Chapter 1: Introduction

1.1 Failing Education Change – Past Predictions

History is abundant with predictions about the influence of new technologies on educational practices. Reality however has seen traditional attitudes resolute in resisting change. Cuban's overview of the use of educational technology since 1920 (1986) offers examples of such predictions. As early as 1913 Thomas Edison predicted that the motion picture would have a major impact on education:

> 'Books will soon be obsolete in the schools... It is possible to touch every branch of human knowledge with the motion picture.'

Darrow (1932) was another early education visionary who believed that the technological innovation of the radio would revolutionise education:

> 'The central and dominant aim of education by radio is to bring the world to the classroom, to make universally available the services of the finest teachers, the inspiration of the greatest leaders'.

Over a decade later Levenson (1945) exemplified continuing predictions about the role of the radio in education:

> 'The time may come when a portable radio receiver will be as common in the classroom as is the blackboard. Radio instruction will be integrated into school life as an accepted educational medium'.

One of history's more prominent advocates of change in education was Andrew Carnegie. Driven by the belief that education was life's key, Carnegie was convinced in the power of access to information. Proclaiming that 'each petty school calls aloud that it has the truth, the whole truth, and nothing but the truth, but no school can embrace the whole' (Carnegie, 1895), Carnegie thought that libraries would positively replace schools as the arena for education. Consequently one of his lifelong interests was the establishment of free public libraries to make available to everyone a means of self-education. At the onset of his library building program in 1881 only a few public libraries existed worldwide. As testament to his vision, by the end of the program in 1917 his corporation had financed the building of 2509 libraries throughout the English speaking world.

It is argued that inaccurate predictions about the revolutionary use of technology in education has had a negative impact on the attitudes of education practitioners. Papert (1993) believes that 'the failure of past reformers to bring about dramatically better learning has armed those within the educational establishment with the argument that future proposals will prove no more capable of bringing about radically improved learning'. Referring to some of the more seminal thinkers of educational theory as Dewey, Freire, Piaget, and Vygotsky; he was acknowledging radical reformers armed with little more than a 'strong philosophical sense'. Mayes (1995) identifies with these failures from a technological perspective warning that 'Learning technology has been through several previous cycles involving high expectation of a revolution in educational practice, followed by complete disappointment'. Laurillard (1993) comments, 'There is a folk wisdom in academic circles that educational technologies come and go, and all the expensive machines end up gathering dust in cupboards'.

It is the computer, Papert (1993) proclaims, that 'with much more persuasive power than the philosophy of even so radical a thinker as Dewey' offers new hope to those who advocate change in educational practices in favour of epistemological values that subscribe to our natural learning abilities. Such optimism is reiterated by Mayes (1995) with the added caution that 'the expectation will again be unfulfilled unless the technology meets real learner needs', with such needs being realised by empirically proven theories of learning.

Whether or not technology in the form of the computer can make a fundamental difference to educational practices is yet to be seen. Today, educational practitioners still exhibit an air of pessimism towards the use of computers.

1.2 The Need for Change

Proponents of extreme change in education subscribe to an ethos of no teachers and schools. A society that denounces the traditional epistemological principals grounded in the belief that intellectual progress consists of moving from the concrete to the abstract, and embraces a more concrete way of learning. Key to such radical calls for change is the belief in students' natural learning abilities, which Papert (1993) suggests goes against the grain of school methods, and according to Illich (1970) is best nurtured in libraries.

Calls for change are not just heard from the education community. The hypermedia community has its own radical thinkers. In a talk given at MediaActive '94, held at Liverpool John Moores University, Ted Nelson (1994) denounced schools as wasteful and harmful, again suggesting reliance on students' natural enthusiasm and interests to guide individual learning. Nelson's cries echoed Perelman (1992), who also called for a move away from teaching in schools.

With the continuing rise of staff-student ratios, a new impetus is fuelling the imaginations of policy makers and practitioners who are looking towards modern technologies to ensure the ongoing pedagogic quality of education at all levels. Laurillard (1993) acknowledges the positive effect psychologists have had in schools, especially at primary level, but raises concerns about the way universities continue to subscribe to the traditional classical practice of 'imparting knowledge' with a reliance on lectures and textbooks. With the computer being hailed as the latest medium to revolutionise education practices, this may be about to change. The World Wide Web (WWW) (Berners-Lee et al. 1994) now offers a medium through which computer based learning material, or courseware (Mayes 1995), can be made available to a global audience (Hawkridge 1996). While radical advocates suggest 'campusless universities' where education takes place entirely within virtual environments, computers are increasingly being employed to augment or replace facets of existing teaching environments.

1.3 Improving Medical Education

Medicine is one domain of higher education that has gone further in recognising the need for change in academic learning. Section 5 of the Medical Act 1983 charged the Education Committee of the General Medical Council, by statute, with responsibility for 'promoting high standards of medical education and co-ordinating all stages of medical education'. Ten years later, in pursuance of these responsibilities, the General Medical Council released its recommendations on undergraduate medical education (Tomorrow's Doctors 1993). Fourteen principal recommendations were made, the first of which addressed the traditional didactic approach of imparting knowledge:

> 'The **burden of factual information** imposed on students in undergraduate medical curricula should be substantially reduced'.

The second recommendation highlighted the committee's recognition of natural learning principals through learning by exploring (Schank 1994, Schank and Cleary 1995):

> '**Learning** through curiosity, the exploration of knowledge, and the critical evaluation of evidence should be promoted and should ensure a capacity for self-education; the undergraduate course should be seen as the first stage in the continuum of medical education that extends throughout professional life'.

Having addressed educational theory, a further recommendation served to highlight the committee's commitment to the use of educational technology:

> '**Learning systems** should be informed by modern educational theory and should draw on the wide range of technological resources available; medical schools should be prepared to share these resources to their mutual advantage'.

The constituent research of this thesis presents an innovative way of employing courseware technology to satisfy these recommendations.

1.4 The Thesis Position

1.4.1 Research Motivation

The research presented in this thesis constitutes work carried out as part of the Southampton Computer Assisted Learning Pathology Education Laboratory (SCALPEL) project. Initiated by Professor Roche from the Department of Pathology, and funded by the University of Southampton's Medical Education Unit and the Department of Electronics and Computer Science, the project was charged with researching the impact of replacing traditional pathology laboratory teaching with pedagogic courseware. Consequently, this research was primarily driven by the continuing need for empirical data on the employment of courseware in the education arena.

The initial aims of the SCALPEL project were as follows:

- Identify the requirements for implementing courseware to replace existing pathology laboratory case-based problem solving exercises.
- Engineer and implement courseware.
- Design and conduct an empirical study into the effectiveness of employing courseware within the domain of pathology laboratory teaching.

Analysis of the quantitative and qualitative data collected during the first two years of the project highlighted a need to address student misconceptions of the domain material with students' expressing a continuing need for demonstrators to assist with misconceptions developed whilst working through the case-based practicals. Consequently the research focus shifted to augmenting SCALPEL with functionality that would replicate the role of the traditional laboratory demonstrator through offering students the ability to engage in dialogue about the pathology material with demonstrators and their peers. The aims of this stage of the project were as follows:

- Investigate innovative technologies within the domain of computer-aided support of domain knowledge.
- Engineer and implement new functionality into existing SCALPEL environment.

- Design and conduct empirical study of new functionality.

1.4.2 Contribution of Thesis

This thesis contributes to the domain of computers in education by:

- Adding to the wealth of empirical data relevant to courseware evaluation and patterns of user interaction within courseware environments.
- Offering an innovative computer-mediated communication model to realise students' continued need to engage in dialogue about the domain material with demonstrators and peers. This is achieved by building on existing models of FAQ functionality by dynamically filtering context sensitive FAQ's.

From a holistic viewpoint, this thesis communicates a case study (Fenton and Pfleeger 1997, Kitchenham and Pickard 1998) which replaces a traditional case-based teaching environment with courseware informed by modern educational theory and underpinned by a novel technological model. Lessons learnt from conducting this research are of value to educational practitioners at all levels and lend themselves to form the basis for future research within the domain of computers in education.

1.5 Outline of Thesis

It is important to realise that authoring courseware requires an understanding and belief in an educational model of pedagogic value. In order to subscribe to an educational model an understanding is required of what has gone before, what is considered pedagogically correct today, and the educational models of tomorrow. A review of educational theory and practice is subsequently presented in chapter 2. Chapter 3 follows this theme with a focus on educational technology.

The remaining chapters have been structured to add a chronological perspective to this research. Chapter 4 discusses the authoring of SCALPEL. An insight is offered into: the traditional pathology laboratory environment; the technology underpinning the SCALPEL courseware; and the lessons learnt from integrating new technology into undergraduate teaching. not on a pilot basis but as part of real teaching.

An understanding of the impact of introducing this major change was necessary to ensure that the pedagogic value of pathology teaching would not be compromised. Chapter 5 discusses the research methodology underpinning the empirical evaluation and comparison of the traditional laboratory and SCALPEL teaching environments.

Chapters 6. 7. 8 and 9 provide details of the experiments carried out to assess the impact of SCALPEL. Analysing the results of these experiments prompted a change of research focus. Attention turned to investigating novel ways of embracing the role of the demonstrator, as is typically found in the traditional laboratory environment. i.e. namely, functionality that attempts to resolve student misconceptions and encourage dialogue between students and demonstrators. The design and implementation of this functionality is discussed from a technological and educational perspective in chapter 10.

Informed by the same research methodology as detailed in chapter 5. chapter 11 presents the empirical evaluation of the dialogue functionality introduced in chapter 10 to satisfy the requirements of misconception resolve through discourse within SCALPEL.

Finally, chapter 12 offers a synopsis of the constituent chapters of this thesis. The major contributions are once again addressed along with: a vision of what benefits innovative technologies might bring to education in the future; and a discussion on how future research can augment the work presented in this thesis.

Chapter 2: Theoretical Influences on Education

2.1 Introduction

It is important to realise that authoring courseware requires an understanding and belief in an educational model that provides pedagogic value. This fundamental point is well considered by Laurillard (1993) who believes that 'any attempt to clarify the pedagogical principles associated with the design of educational media has to begin with some consideration of an underlying educational model'. Referring to what is now commonly labelled as 'learning technology', Hawkridge (1996) suggests that the 'psychological underpinning for the next educational technology is not very firm'. Hawkridge raises concerns about a field in which practitioners from a diversity of specialist areas approach the task of integrating technology without any notion about teaching and learning other than those influenced from personal experiences.

Given that this thesis is about introducing learning technology into an undergraduate learning environment this chapter offers a perspective on the theoretical influences of contemporary moves to introduce computers into education. Educational technology models are broadly categorised as being influenced by either behaviourism or constructivism, although the latter is generally considered to have had more of an impact in schools than in higher education (Hawkridge 1996). Each of these psychology fields has a particular theory of knowing that has underpinned, albeit with only a tenuous empirical and/or logical link (Laurillard 1993), the integration of technology into the education arena. To this end, it is important to have an insight into the central tenets of these theories. Following this, in chronological order, is an introduction to the three educational technology models that have attracted the most attention from contemporary education: Instructional Design, Intelligent Tutoring Systems (ITS), and Constructivist Learning Environments.

2.2 Theories of Knowing

von Glaserfield (1995) attributes his motivation to work on a constructivist theory of knowing to a decline in education: 'Education may never have been considered good enough, but whatever its methods and effectiveness were, it seems to have suffered a decline during the last 20 or 30 years'. He believes that this decline has been driven by a misconception in education circles that if you give teachers more money, they will teach the right answers. Commenting on what he describes as a 'fatal attitude', von Glaserfield advocates that: 'Money does not change the philosophy of education, and a philosophy of education that believes in teaching right answers is not worth having'. Acknowledging that as a constructivist his views cannot be taken as objective, von Glaserfield suggests that this misconception came about as a result of 50 years in this century being influenced by the 'virtually undisputed domination of a mindless behaviourism'. All learning, von Glaserfield proclaimed, was 'reduced to a model that had been derived from experiments with captive pigeons and rats'. The fundamental principle that underpinned this model was what Thorndike called The Law of Effect (von Glaserfield 1995): behaviour is stamped-in when followed by certain consequences (Skinner 1953). Thorndike (1931) observed that animals, including humans, tend to repeat actions that through experience lead to satisfactory results. von Glaserfield comments that the behaviourists turned this theory into a learning theory based on the power of reinforcement: responses that are reinforced (or rewarded) will be repeated.

The behaviourists being referred to were informed by the work of the leading psychologist B. F. Skinner. In her contribution to over 100 monographs on 'Thinkers on Education', Smith (1997) writes of B. F. Skinner as being 'the most important American psychologist of the twentieth century, and arguably the most important world psychologist since, or including, Freud'. The behaviourist learning theory, according to von Glaserfield (1995), has had a negative impact within education as it has deflected attention away from the reasons that prompt students to respond or act in a particular way in favour of performance. The consequence of the behaviourist learning theory is exemplified by students who adequately give the right answers to standard questions, but when asked to solve simple problems that do not precisely represent the preceding course of instruction, exhibit no understanding of the concepts. The absence of problem solving skills in students is attributed to the widespread behaviourist born misconception that competence in intelligent behaviour is a consequence of drilling performance. A proponent of constructivism, von Glaserfield (1995) promotes the conceptual understanding of abstract building blocks and their myriad of relationships as the ingredients that give students 'a chance of success when faced with novel problems'. Awareness for the need for conceptual development has brought to the fore questions about how conceptual development should be approached and how it could be fostered. Constructivism addresses these questions by offering a new perspective on the theory of knowledge, its structure and acquisition - a new epistemology.

The foundation of the contemporary constructivist theory of knowing was laid by Vygotsky (1978; for monograph see, Ivic 1997), and Piaget's (1963; for monograph see, Munari 1997) 60 years of work as a developmental cognitive psychologist (for comparative works on Piaget and Vygotsky see for example: Brockmeier 1996; Brown et al. 1996; Martí 1996). The tenets of their theories are grounded in the belief that knowledge cannot simply be transmitted or conveyed wholesale to another person. Of the many forms of constructivism, it is this first principle that unifies them all (Ernest 1995). von Glaserfield (1995) reiterates this by pointing out the constructivist belief that concepts have to be conceived, and that a conceptual repertoire cannot evolve by simply transferring knowledge from teachers to students.

2.3 Educational Technology Models

Behaviourist principles influenced early implementations of educational technology, the traditional field of Computer Assisted Instruction (CAI), and remained a prominent force through the traditions of Intelligent Tutoring Systems (ITS), and Computer Assisted Learning (CAL) of the eighties and early nineties. The cognitive science field of constructivism has become the dominant modern approach, although methods of the now discredited behaviourism continue to be recognised in many of the modern learning environments. This section introduces the three chief educational technology models that have been employed in contemporary education: Instructional Design, Intelligent Tutoring Systems (ITS), and Constructivist Learning Environments.

2.3.1 Instructional Design

Skinner (1960) designed 'teaching machines' based on his behaviourist theories. Concerned not with what students think, but on teaching students to behave correctly to a given stimulus, these mechanical machines presented a list of short-answer questions to be answered sequentially. If a question was answered incorrectly the student was immediately informed of the correct answer but no attempt was made to identify the reasoning behind the misconception. Students were not required to apply the information in a realistic setting, they simply continued answering questions until they produced the correct responses.

The availability of computers in the late sixties and seventies saw the rigid paths of the mechanical machines being replaced with branching systems. However, influenced by 'instructional design' (Gagné 1965, Gagné and Briggs 1979), a formal approach to computer based teaching remained prominent. From the designers' perspective, this 'highly disciplined' (Boyle 1997) procedure involves working backwards from the learning goals, identifying every sub-task the student has to do and every piece of information the student needs to acquire. Fundamentally, the task of the designer is to build an appreciation of the appropriate instructional events in order to encapsulate prescriptive guidance: target behaviour is systematically decomposed into constituent skills that are subsequently used to derive a pathway through the material starting with simpler skills before moving on to the complex.

Underpinned by the premise that learning consists of behaviour modification through conditioning, this approach has earned the label 'drill and practice' (Nix 1990, Schank 1994) and influenced the traditional CAI (Computer Aided Instruction) arena. Jelovsek and Adebonojo (1993) offer a summary of empirical studies of the use of CAI in medical education. 49 randomised clinical trials of medical education courseware published in English between January 1966 and June 1992 were surveyed. Trials were included in the survey if they studied 'the effectiveness of CAI, whether they measured knowledge retention (primarily short-term), changes in behaviour, the time needed for CAI as compared to other methods of instruction, or the study of participants' attitudes toward CAI'. A study was considered randomised if interventions were alternated or assignment was based on the subject's identification number or geographic site. Analysing the results, the authors found that 61% of the 49 studies highlighted improved performance after the use of CAI. Notably, they found only one study that showed CAI to produce 'significantly lower scores' when compared with traditional methods. Concluding, Jelovsek and Adebonojo suggest that 'students learn as well with CAI as with traditional methods, and in some cases they learn better', and caution that each application needs to be 'tailored to its specific educational objectives'.

Laurillard (1993) criticises instructional design as having 'only a tenuous link to any empirical base' and dismisses it as 'not a progressive force'. Papert (1993) denounces CAI, for which he gives Suppes the accolade of being the 'intellectual father', as 'using the computer to program the students'. The approach of Suppes (1980) was to allow the computer to generate an optimal sequence of presentations based on the past history of the individual learner. Papert (1993) points out that this approach is grounded in the behaviourist theory of learning, which he believes is 'not "learning" in the sense of something a learner does but "instruction" in the sense of something the instructor does to the learner'.

However, the influence of instructional design on CAL (Computer Assisted Learning) environments cannot be dismissed. Schank (1994) suggests that most contemporary educational technology applications exhibit a 'page-turning architecture' that is largely modelled on outdated learning theories (e.g. instructional design) influenced by Skinner's discredited behaviourism. Schank criticises applications that present students with screens of text with buttons that allow sequential navigation between pages as being passive on-line books with a gloss that looks impressive but offers little educational value. Laurillard (1993) suggests that the successful influence of instructional design is probably attributable to a prescriptive list of instructional events to be carried out by the teacher that 'have an intuitive logical appeal' and 'seem unobjectionable'.

Carroll (1990) believes that the problems with instructional design technologies stem from a bias towards 'monolithic analysis, diagnosis, and solutions'. As a consequence Carroll advocates the design and implementation of individualistic educational practices empirically driven by observations of 'what people want to do, how they want to do it, and the problems they are observed to experience'. This 'value-free' ethos subscribed to by Carroll is as a result of what he calls the 'ephemeral status of particular orientations to educational philosophy', a philosophy which he believes to be cyclical not evolutionary. Carroll's view is exemplified by the back-to-basics movement of the late 80's calling for the return of rote memorisation of fundamental facts.

Carroll (1990) found what he calls 'fundamental flaws in the standard systems approach to instruction' advocated by Gagné and Briggs (1979). By observing users' patterns of interaction with self-instruction training material of the kind typically integrated into desktop applications (e.g. word processors) Carroll observed that users were too involved with making sense of the situation to be concerned with the prescriptive nature of the instruction. Empirically driven by these observations, Carroll proposed an alternative 'minimalist' user centred instructional approach where the student not the system determines the model and the methods of instruction, a model that minimises prescriptive instruction and supports the spontaneous nature of students' natural learning strategies. The minimalist approach is concerned with obstructing as little as possible the learner's proactive efforts to find meaning in the learning activity by engaging students with realistic tasks from the very start whilst minimising didactic passive activities and helping to make errors and error recovering pedagogically constructive. In Carroll's words: 'The minimalist model tries to capitalize on what learners do spontaneously instead of being embarrassed by it' (Carroll 1990).

2.3.2 Intelligent Tutoring Systems

CAI, and its behaviourist traditions, influenced an area of research into the use of computers in education known today as ITS (Intelligent Tutoring Systems). Wenger (1987) comments on the history of ITS research. Much work in this area of Artificial Intelligence (AI) was originally conducted under the name 'Intelligent Computer-Aided Instruction' (ICAI), a label that evolved out of 'Computer-Aided Instructed' (CAI), just one of the names commonly given to the use of computers in education.

Wenger (1987) describes the purpose of ITS research as 'not to provide a software-engineering framework within which experts can compose instructional interactions by representing their decisions in the form of programs', but to 'capture the very knowledge that allows experts to compose an instructional interaction in the first place'. Using explicitly represented knowledge, ITS programs are responsible for composing dynamic instructional interactions according to models of the domain, models of the students, and models of the communication or tutoring processes. It is the emphasis on actually giving systems access to a representation of knowledge that they attempt to communicate that prompts Wenger (1987) to call these systems 'knowledge communication systems'. He defines knowledge communication as 'the ability to cause and/or support the acquisition of one's knowledge by someone else, via a restricted set of communication operations' (Wenger 1987). Recognising the behaviourist underpinnings of earlier work in this field, Wenger promotes a radical methodological shift away from the traditional ITS focus on the programming of decisions, towards a cognitive orientation: a computational understanding of processes of knowledge communication.

Moyse and Elsom-Cook (1992) argue that Wenger's definition of knowledge communication implies that 'some single, pre-existent mass of knowledge has to be carefully poured into a politely passive student'. They promote recent moves in the ITS community away from the idea that 'the tutor takes a directive role controlling the interaction', a model that implies that the tutor has the correct knowledge about the domain and holds a position of authority over the student. Moyse and Elsom-Cook (1992) highlight the need to recognise that many domains do not have a single correct representation, and that 'interpretation of the domain, or "viewpoint", must be jointly constructed between teacher and learner'. They attribute the 'earliest and most explicit' recognition for the need for multiple viewpoints in tutoring systems to the work of Stevens et al. (1979) who found that students' accounts of the process of heavy coastal rainfall after tutoring contained a number of bugs as a result of the importation and incorrect use of concepts external to the intended domain. The ITS community has now widely embraced the need for what Self (1992) calls 'computational viewpoints' and a negotiation-based approach to tutoring. Moyse and Elsom-Cook (1992) suggest that this is testament to a move away from an approach that they quote Self (1989) as proclaiming to be 'education as transmission', and subscribe to the belief that students are active participants in the tutoring process rather than passive recipients.

Wenger (1987) attributes Carbonell with launching the ITS paradigm by introducing SCHOLAR (Carbonell 1970), a tutoring system that facilitates mixed-initiative dialogue about the geography of South America. Some systems monitor students very closely, adapting to responses but never relinquishing control. In mixed-initiative dialogues, control is shared by the system and student as questions and answers are exchanged. Systems that subscribe to this approach must be able to respond to a student, but may also ask questions to help the student understand what they are attempting to do. Wenger (1987) comments that as a result of representing knowledge with semantic nets, a technique that uses nodes to represent knowledge, and links to represent their relationships; SCHOLAR had little use as a tutoring system. However SCHOLAR is regarded as the background to much of the research on ITS systems and inspired the initial conception of MYCIN (Shortliffe 1976), a consultation system that concentrates on offering the appropriate treatment of patients who have infectious diseases. MYCIN is described by many as one of the earliest and best-known examples of an expert system (Charniak and McDermott 1985, Shadbolt and O'Hara 1997, Wenger 1987). Work on expert systems was seen to offer an ideal basis on which to build intelligent tutoring systems (Wenger 1987). Turban and Aronson (1998) define an expert system (ES) as 'a system that uses human knowledge captured in a computer to solve problems that ordinarily require human expertise'. Because the computers of the 1970s were unable to support a program of MYCIN's size on a cost-effective machine with acceptable response times, MYCIN was never used clinically. However, evaluations of MYCIN showed that the advice offered compared favourably to that of experts on infectious disease (Shortliffe 1990).

MYCIN was not a tutoring system, but was used as the basis for GUIDON (Clancey 1979), viewed by some as one of the most provocative examples of an ITS system (Hoffer and Barnett 1990). Alpay (1992) believes that GUIDON represents a significant milestone in the design of ITS systems for teaching medical diagnosis. Clancey (1979) developed GUIDON by adapting MYCIN's explanation functionality to produce an active tutor. GUIDON works by communicating successive cases through mixed-initiative dialogue in order to convey MYCIN's knowledge to students in a realistic problem-solving context. Once a case has been selected and described by GUIDON, the student adopts a diagnostic strategy by asking questions to gather important data and propose hypothesis. During the tutoring process GUIDON continually compares the student's behaviour with that of an expert. The system intervenes when the student asks for help or when performance is determined by GUIDON to be sub optimal. As with MYCIN, and most other AI based educational applications, GUIDON had limited success outside of the research setting.

The use of AI, and in particular the sub classifications of expert and intelligent tutoring systems in education, proved useful in specific instances only. Where successful, the systems required large computers and consumed many man years of research and development (Shadbolt and O'Hara 1997). Laurel (1997) reiterates this point believing that AI is a discipline 'whose star is currently in eclipse' with many holding the view that it doesn't work, and even if it did, an instantiation would require more processing power than it's worth. Shadbolt and O'Hara (1997) add caution to the success of expert systems in particular by pointing out that the majority of systems only reached the prototype stage, and attribute this to AI not being 'a mature and exploitable technology'. Negroponte (1997) believes that 'AI suffered a turn for the worse around 1975 when computing resources started to achieve the kind of power that might be needed to solve intuitive problems and to exhibit intelligent behaviour'. As a consequence science opted for achievable and marketable applications like commercial expert systems, leaving untouched what Negroponte refers to as 'the more profound and basic questions of intelligence and learning'. However, Negroponte and others believe that AI is once again moving onto the centre stage of scientific research under the guise of software agents - entities that function continuously and autonomously in a particular environment, often with other agents or processes (Bradshaw 1997, Shoham 1997).

2.3.3 Constructivist Learning Environments

Papert is attributed as being the seminal proponent of learning environments grounded in the constructivist view of learning (McKnight et al. 1991). Papert (1990) builds on Piaget's constructivism by offering a constructionist attitude to teaching: 'The word with the "V" (constructivism) expresses the theory that knowledge is built by the learner, not supplied by the teacher. The word with the "N" (constructionism) expresses the further idea that this happens especially felicitously when the learner is engaged in the construction of something external or at least sharable'. By something external or at least sharable, Papert (1993) was referring to products that can be 'shown, discussed, examined, probed, and admired' - a sand castle, a cake, a Lego house, a computer program, a book, a poem.

Papert (1980) instantiated his constructionist theories through a still popular LOGO language specifically aimed at children. The idea of the LOGO environment is for children to program the computer, rather than to be drilled by it. Instead of presenting the student with a set of predetermined unambiguous facts, the learner is encouraged to construct something within the computer environment. The principle of the LOGO environment is to encourage the formalisation of creative, self-expressive problem solving strategies - processes that can be used to solve problems in other LOGO and LOGO-like environments, and that can be generalised to other domains. Building on LOGO, Papert (1980) introduced another power idea called "microworlds", computer environments in which certain questions are relevant and others are not. Given that a microworld is authored with its own set of assumptions and constraints, Papert (1980) believes that children get to know what it is to explore the properties of a given microworld undisturbed by extraneous questions. Many concepts seem too abstract and alien for children to grasp, a microworld helps children make sense of tasks in terms of everyday familiar experience, but supports transcending the particular into the world of the abstract more effectively than the abstract descriptions in a textbook.

Papert provides little quantitative empirical evidence to support his constructionist ideas believing that: 'The method of controlled experimentation that evaluates an idea by implementing it, taking care to keep everything else the same, and measuring the result, may be an appropriate way to evaluate the effects of a small modification. However it can tell us nothing about ideas that might lead to deep change' (Papert 1980). He suggests that deep change evolves through a close harmony with social evolution which is steered less by the outcome of tests and measurements than by intuitive understanding. Instead, Papert employs a technique of communicating a series of learning stories or anecdotes to reinforce his theories of education (see for example, Papert 1980).

Ackermann (1995) sums up Papert's contribution to Piaget's constructivist theory as specifying 'some of the conditions under which such constructive learning is more likely to happen'. Hawkridge (1996) points out that many existing computer assisted learning environments claim to be founded on the construct[iv/ion]ism principles, but cognitive psychology has tended to be more focused on schools than on higher education.

2.4 Summary

This chapter offered a perspective on the theoretical influences on contemporary learning environments. Behaviourist and constructivist theories of learning were introduced before the three primary educational technology models that have demanded the most attention within education circles. The next chapter introduces hypermedia, a technology that continues to have an overwhelming impact in education.

Chapter 3: Hypermedia Learning Environments

3.1 Introduction

The previous chapter offered a perspective on the learning theories that have been influential on contemporary learning environments. Hypermedia is now seen as the dominant force in educational technology. This chapter introduces the origins of hypermedia and discusses its influence on the use of computers in education. Before focusing on hypermedia, section 3.2 discusses how multimedia in general affords learning. The focus on hypermedia begins in section 3.4 with an introduction to hypertext and how technological advances with differing media formats has brought about the concept of hypermedia. This section concludes by introducing examples of hypermedia systems. Section 3.5 discusses the use of hypermedia in education and offers examples of hypermedia learning systems. Finally, section 3.6 offers a critical reflection on the use of hypermedia in education

3.2 Multimedia in Education

Woolf and Hall (1995) point out that a multimedia learning environment is only useful if the learner remains active and motivated. Students need to be challenged to reason about the material being presented, which is achieved through offering an experience that is 'authentic and relevant to the learner's life'. Authenticity of learning settings or 'situated learning' is a concept that is discussed in great depth by Laurillard (1993). Acknowledging that the aim of university teaching is to make learning possible (Ramsden 1992), Laurillard believes that 'teaching must not simply impart decontextualised knowledge, but must emulate the success of everyday learning by contextualising, situating knowledge in real-world activity'. Laurillard (1991, 1993) makes a distinction between the natural environments which favour learning of 'percepts' in everyday life, and artificial environments which are designed to communicate 'precepts', or givens, in education. In learning preceptual knowledge, students are learning someone else's descriptions of the world, which gives teaching a 'second-order character'. Making learning possible for Laurillard is a rhetorical activity. It is mediated learning in that it is informed by the analogy of situated learning in everyday life, but is adapted to the learning of descriptions of the world devised by others.

Although many education psychologists believe social interaction should be a key consideration of any educational model, courseware authors often omit to recognise this. Mayes (1997) suggests that the pedagogic value provided by lectures is social, although it doesn't necessarily follow that arranging learners into physical assemblies facilitates cooperative peer learning. For cooperative learning to occur, Mayes (1995) believes that there needs to be 'genuine task-structure which encourages relevant group interaction', and proposes an educational model that views the basic unit of learning as a continuous cycle of conceptualisation and re-conceptualisation (Mayes 1995, McKendree and Mayes 1997). The model, which is shown in Figure 1, suggests three stages in the interactive process of understanding a knowledge domain.

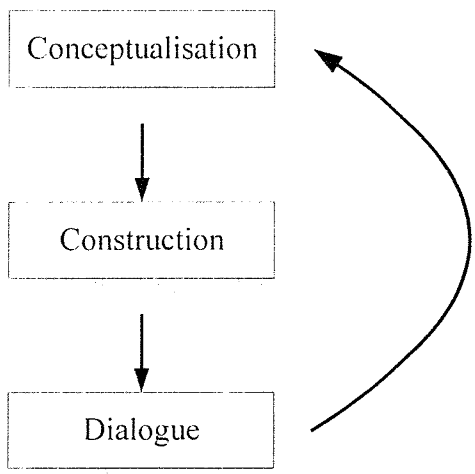


Figure 1 Conceptualisation Cycle

Conceptualisation refers to the student's current understanding of the knowledge domain. This stage of learning occurs through dialogue with other peoples', and reflecting on pre-existing conceptualisations. The construction stage allows the student to test their understanding of the knowledge in performing a task. If the test satisfies the student's understanding of the knowledge then the cycle is complete. Dialogue with teachers and peers promotes reflection, creating new conceptions or prompting a student to question the validity of pre-existing conceptions, hence satisfying the iterative process. Laurillard (1993) acknowledges the fundamental concept of dialogue between teachers and peers affirming that all teaching models should provide dialogue that reveals the conceptions of both teacher and student.

Recognising the pedagogic limitations of passive software where users have no active part in learning the domain knowledge, Schank (1994) believes that "multimedia systems should be designed for compatibility with people's powerful natural learning mechanisms". Schank and Cleary (1995) describe the process of learning in the real world as adopting goals, generating questions, and finding answers (see Figure 2), and stresses the importance of educational establishments to be configured to support, not 'short-circuit', this natural learning waterfall.

The fundamental concept being communicated is that learning begins with a goal that is adopted willingly. Courseware should allow students to control the environment in which they learn, allowing them to adapt the knowledge domain to their individual learning requirements as their interests and goals guide them. The issue with natural learning is that it requires 'an inordinate amount of individual attention from teachers, parents friends and others' (Schank 1994), a resource which in most cases is too expensive to provide. Schank and Cleary (1995) believe that multimedia environments can make one-on-one tuition a real possibility. The problem lies with identifying the characteristics by which we learn naturally and instantiating systems that tap into these characteristics. Informed by theoretical and empirical perspectives, Schank and Cleary (1995) offer five teaching models that they feel are particularly amenable not just to multimedia systems, or to computer-based learning systems in general, but education in general. They suggest that each model is based on an underlying teaching method that is not dependent on computers - 'they may be furthered through computers, but by no means do they depend on them' (Schank and Cleary (1995). Figure 2 shows which part of the natural learning waterfall is targeted by each of these models.

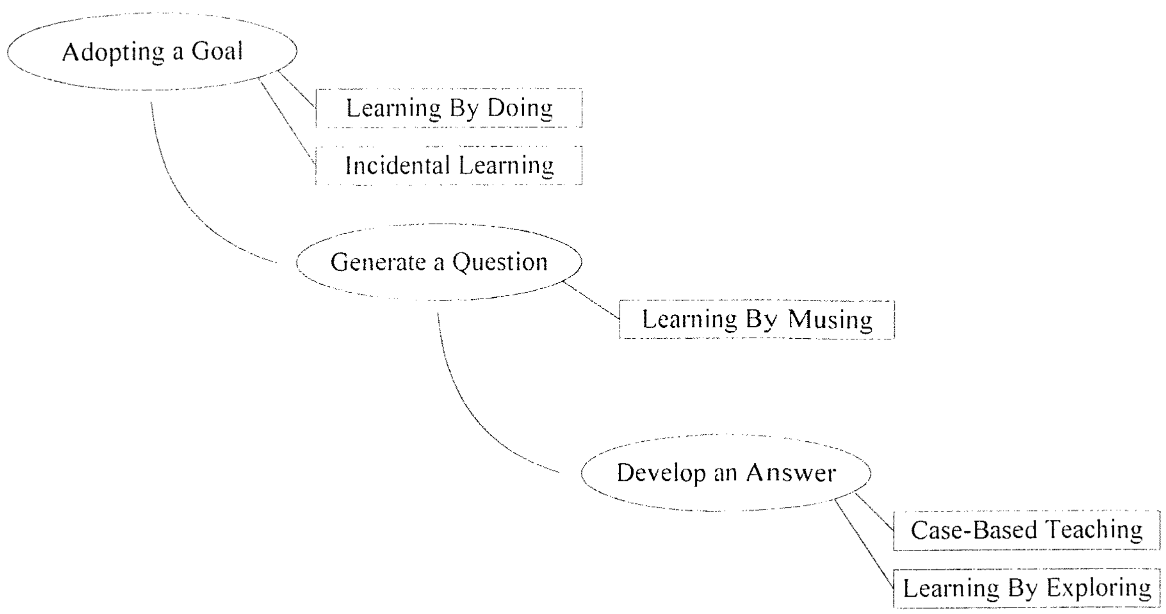


Figure 2 Teaching models and the natural learning process (Schank and Cleary 1995).

These models provide an agreeable framework for discussing how multimedia environments afford learning. Multimedia technology can support a variety of learning settings and best practice often constitutes hybrid adaptations of these architectures. These models are not mutually exclusive in instantiations of learning environments. Any segregation across these models has been made for descriptive purposes only.

3.2.1 Simulation-based learning-by-doing

Simulations encourage students to learn skills by immersing the user in an active learning environment (Boyle 1997). O'Shea and Self (1983) comment on the benefits of simulations as being to 'provide a student with a safe, inexpensive view of certain phenomena, such as nuclear reactions, or space travel'. Laurillard (1993) warns of the confusion in terminology when 'simulation' is used to refer to a program that runs a model without any inputs. A simulation of this constitution, where the program generates its own inputs to the model and the user simply watches, is relegated to non-interactive animations or demonstrations. Woolf and Hall (1995) broadly categorise simulations as being either scenario-based or knowledge-based. Scenario-based simulations employ multimedia technologies (see section 3.4) to engage the user in realistic situations, but offer static representations of the problem domain and therefore cannot respond to student questions or explain the domain beyond predetermined settings. Knowledge-based simulations are technically more complex as they model the situation and user to make assumptions about the user's state of knowledge and learning needs. Knowledge-based simulations can therefore dynamically adapt and offer appropriate advice by comparing user actions to that of an expert.

Smith et al. (1997) offer a different perspective on simulations. They criticise most simulations of not permitting users to modify their fundamental behaviours and assumptions. Whereas knowledge-based simulations dynamically adapt to user interactions, Smith et al. propose a symbolic simulation tool kit 'KidSim' that allows learners to actually construct and modify existing behaviours and assumptions of a microworld. A symbolic (as opposed to numeric) simulation is defined as a 'computer-controlled microworld made up of individual objects (agents) which move around a game board interacting with one another'. Conceptually, software agents in a KidSim microworld are analogous to those in Papert's (1980, 1990, 1993) Logo microworlds (see also, Boyle 1997; Laurillard 1993; Nix 1990; Wenger 1987). The difference is that in Logo microworlds learners program objects using Logo, in contrast users of KidSim construct objects using a technique called 'graphical rewrite rules'. Rewrite or 'if-then' rules are informed by production systems of artificial intelligence (see, for example, Anderson 1990, 1993; Charniak and McDermott 1985; Clancey 1997) and work by visually transforming an object or small scene in the simulation if a rule is matched. KidSim introduces learners with objects or 'characters' of a microworld and encourages interaction as with ordinary video games. A learner can then decide to change the behaviour of an existing character or elect to define new ones. The act of changing and building new characters is pedagogically important because it forces learners to think: they have to decide what to change, how to change it, and how to fix it when their changes go wrong.

Whalley (1998) takes the idea of simulations further by harnessing the technological advances of the Internet. He believes that simulations can provide highly motivating hands-on experiences for learners, especially when they can communicate and collaborate with each other. To this end Whalley's focus is on bringing together the ideas concerning virtual science simulations and the potential for collaborative dialogue and resource sharing offered by computer mediated communication (CMC). He developed an exemplar system (Virtual Spring Collaborative Microworld) where users located on the Internet benefit from a rich understanding of the simulation domain through their shared dialogues. Stratfold (1998) proposes using online simulations for improving learning dialogues. A proponent of learning dialogues on the Internet, he envisions the results of online simulations feeding into a discussion about individual differences or trends across a group. Such activities could help stimulate discussions and take away the pressure from tutors to lead and drive debates.

3.2.2 Incidental learning

Incidental learning is best suited when a dull or boring task is to be learnt, often involving the memorisation of facts about the world. The aim is to create interesting and fun ways of imparting dull information. The important premise that underpins incidental learning is that when students are participating in something that is fun, they can be learning a great deal in passing without even noticing it. An example domain put forward is learning geography (Schank 1994, Schank and Cleary 1995), a task which can be an extremely dull exercise in rote memorisation. A program called Road Trip is offered as an example of how incidental learning facilitates the learning of state capitals by simulating a car trip around the United States. Once a destination is reached, students can view video clips about interesting activities and events in that location, e.g. sports highlights, movie clips, music videos, amusement parks and historical footage. Students are motivated to achieve a goal that they have real interest in, e.g. see clips of their favourite basketball team, while the uninteresting information is learnt by being intrinsically related to the goal, in that the uninteresting information is used to achieve the goal.

3.2.3 Learning by reflection

Learning by reflection is founded on the idea that its not always the case that a student needs to be told something as much as the need to know how to ask the right questions. We are often our own best teachers if there is someone around to listen to our ideas. When a student has a half-formed idea, Schank and Cleary (1995) see the role of the teacher as asking questions to help the student see the shortcomings in thinking. By acting as sounding boards, teachers can allow students to 'speculate, wonder, imagine, and be creative'. Although multimedia systems cannot provide the empathy that humans can, instantiations can make good sounding boards because they exhibit infinite patients while prompting students to reflect on aspects of a domain by asking probing questions.

3.2.4 Case-based teaching

This model exploits the human capacity to learn isolated facts from stories. Schank and Cleary (1995) believe that 'case-based reasoning is the predominant way in which people think about their worlds'. In deciding what action to take in any given situation, we rely on similar past experiences as a guide. When students learn-by-doing they experience knowledge failures, times when they realise that new information is required to progress. Case-based teaching can help student progress by providing situated knowledge at the relevant time. Harnessing students' natural case-based reasoning qualities is dependent on two principle ideas: experts have a repository of cases, and good teachers are good storytellers.

Boyle (1997) views case-bases as a more radical approach to structuring content within an interactive multimedia learning environment. Because content is introduced as part of an unfolding story relevant to an authentic task, Boyle calls this approach the 'thematic' approach to structuring content. Bransford et al. (1990) offer a model of anchored instruction, and draw parallels between this and case-based approaches to learning. They emphasise the importance of creating an anchor or focus that generates interest and allows the student to deal with a general goal, e.g. organise a trip to a South American jungle, before introducing concepts that are relevant to their anchored perceptions. The central principle is to 'enable students to notice critical features of problem situations and to experience the changes in their perceptions and understanding of the anchor as they view the situation from new points of view'. In the context of case-based learning, students are presented with cases of problems to be solved, and as new concepts are introduced, they see the effects of this information on the problem they confront.

Masterton (1998) uses a case-based approach to deliver previous years discussion threads that may be relevant to a current electronic conference about a course. A retrieval algorithm, or Case Matching Agent (Zdrahal and Domingue 1998), attempts to present students with previous cases of dialogue by matching accumulated keywords from the current discussion with cases from the case-base.

3.2.5 Learning by exploring

Learning by exploring fulfils the need to satisfy students' natural curiosity once they become involved in a subject. Schank and Cleary (1995) believe that learning environments based on this architecture must allow students to follow their own paths through the domain whilst providing multiple expert viewpoints on the knowledge domain. The potential for learner centred exploration is considered to be the primary factor in influencing the shift away from the CAL tradition of prescriptive instruction.

3.3 Authoring Multimedia Learning Environments

Boyle (1997) offers an overview of the main components involved in the development of a multimedia learning environment. Presented schematically in Figure 3, the basic requirements for development are domain knowledge, and design knowledge and skills. Multimedia design knowledge and skills is further broken down into two principle disciplines: conceptual design, and presentation design. Conceptual design is concerned with creating the learning architecture of the system and involves, in consultation with domain experts, the structuring of the domain knowledge. Presentation design involves authoring the system from a technological perspective, with a strong focus on human-computer interaction (HCI) issues and the aesthetics of the interface. The central layer of Boyle's schematic addresses the development process. The favoured software engineering paradigm employed is the spiral model (Pressman 1994). This approach adopts an iterative stance whereby initial concepts are instantiated from the knowledge that feeds into the project (represented by the boxes in Figure 3) and evaluated to give formative feedback for iterative refinement (see section 6.2).

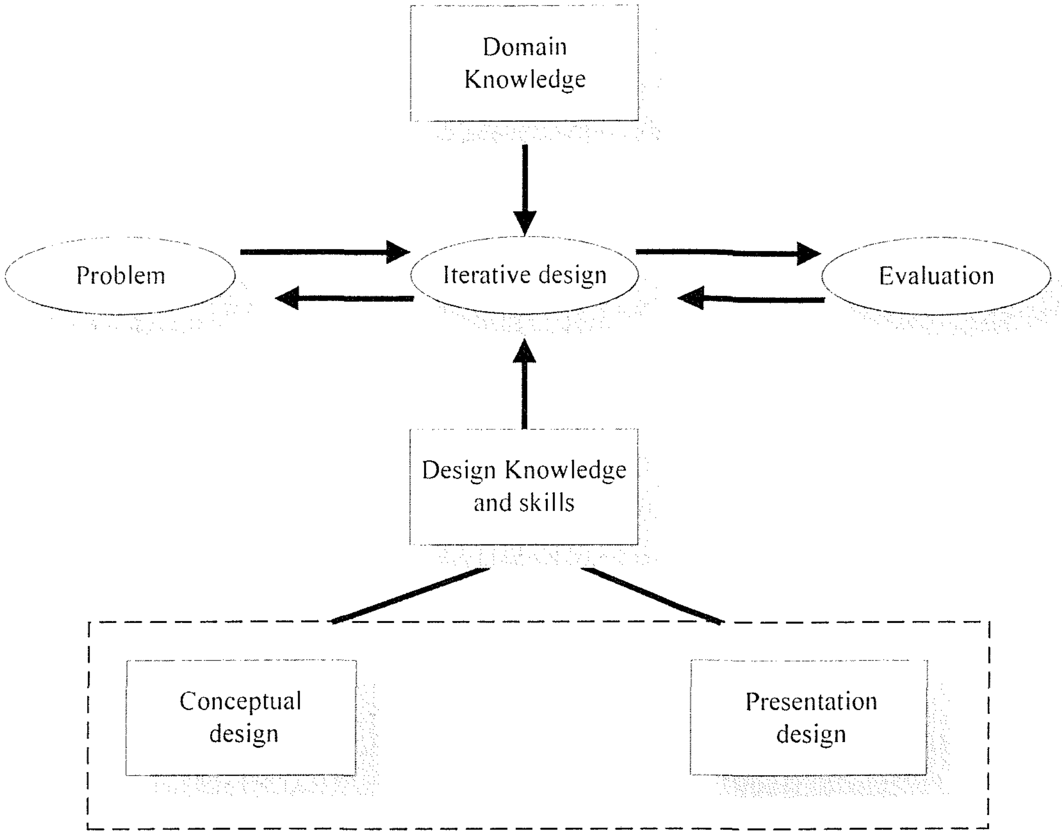


Figure 3. A schematic overview of the principle factors in the design and development of Multimedia Learning Environments (Boyle 1997).

3.4 Hypertext and Hypermedia

In presenting a historical perspective of hypertext, Conklin (1987) suggests that the concept is 'not so much a new idea as an evolving conception of the possible applications of the computer'. Documents that reference other parts of the document or other documents, e.g. dictionaries and encyclopaedias, can be viewed as manual hypertext systems (Conklin 1987), as can text that use footnotes (McKnight et al. 1991). What sets hypertext apart from conceptually inter-linked paper documents is that navigation between nodes in the information space of hypertext is automated by computer support (Conklin 1987, McKnight et al. 1991).

The original hypertext concept was conceived by Vanevar Bush (1945) who, anticipating the information explosion, conceptualised 'memex', a machine for browsing and annotating an extensive corpus of text and graphics. Unfortunately computer technology was not advanced enough to allow a visionary like Bush to realise a working memex. Conklin (1987) comments that it was just less than two decades later in 1963 when Douglas Engelbert (1963, 1984) first proposed a working hypertext system modelled on the ideas of Bush. The system was later implemented in 1968 as NLS (oN Line Systems) and has evolved over the years to be commonly known as 'Augment'. Inspired by Bush and Engelbert, the naming of this now influential field is attributed to Ted Nelson (1988), who coined the term 'hypertext' whilst working on his Xanadu project: a vision of creating a 'docuverse' where the entire literature of the world would be linked.

Conklin (1987) offers the following definition for the concept of hypertext: 'windows on the screen are associated with objects in a database, and links are provided between these objects, both graphically (as labelled tokens) and in the database (as pointers)' (see Figure 4). The objects in the database and their associated links are essentially a network of nodes (Smith and Weiss 1991) that can be thought of as a kind of hyperdocument.

Where the objects in the database are text only, the inferred hyperdocument is called hypertext. In recent years the term 'multimedia' has evolved to describe the storage and display of audio-visual material as well as text and graphics in computer systems (Woolf and Hall 1995). Advances in multimedia technologies have brought about applications that combine a variety of information sources, such as voice, graphics, animation, images, audio, and video into a wide range of applications (Furht 1994). Hutchings et al. (1994) differentiates hypermedia from multimedia. Although multimedia is considered to be a superset of hypermedia and other interactive technologies, instantiations of multimedia do not always exhibit linking functionality that allows them to be sufficiently classified as hypermedia. To this end, the more general term 'hypermedia' is used when the objects in a database infer a hyperdocument of differing media formats (Boyle 1997, Conklin 1987, McKnight et al. 1991, Yankelovich et al. 1991). Throughout this thesis the terms hypertext and hypermedia refer to computer systems of linked information.

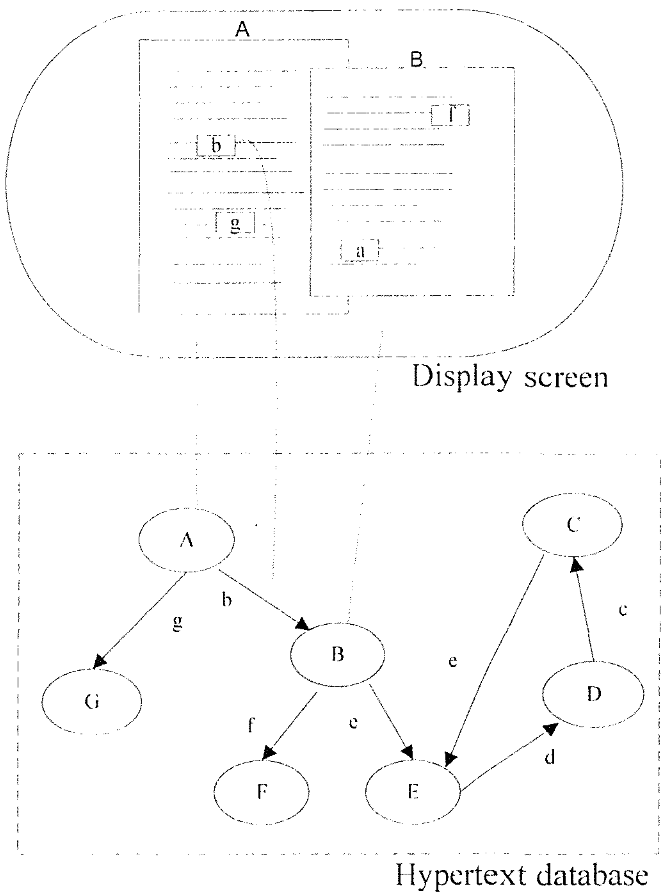


Figure 4 The concept of hypertext: the correspondence between windows and links in the display, and nodes and links in the database (Conklin 1987).

3.4.1 Navigation and Orientation

It is generally considered that navigation was the single greatest difficulty for users of early hypermedia systems and continues to be a prominent issue. Frequent reference is made to getting 'lost in hyperspace' (e.g., Conklin 1987; Duchastel 1991; Edwards and Hardman 1989; Hutchings et al. 1992a; Mayes et al. 1990b; McAleese 1989; McKnight et al. 1990; Nielson and Lyngaek 1990), a shortcoming that Conklin (1987) describes as a consequence of disorientation, of not being sure where you are in relation to other parts of the network, or not knowing how to find other information that you know, or think, exists. A second fundamental problem with using hypermedia is what Conklin (1987) calls 'cognitive overhead'. Here, the requirement to keep track of links imposes additional cognitive load that would otherwise be assigned to thinking about the domain material.

The navigation and orientation problems inherent in hypermedia has influenced research into providing navigation tools. Allinson and Hammond (1989) comment that a hypermedia system can 'provide great flexibility over the sequencing of material, but without provision of overlaying navigation tools it can do little to assist the varying learning abilities and strategies of individual learners'. Empirical investigations suggest that different navigation techniques are suitable in different circumstances (Simpson and McKnight 1990). Wright and Lickorish (1990) call for authors of hypermedia to consider both the inherent structure of the domain material and the tasks that users will be trying to accomplish when deciding upon navigation tools.

Hypermedia navigation tools range from static graphical representations of a hyperdocument through to those that employ knowledge-based intelligence (Boyle and Snell 1990, Kibby and Mayes 1989, Nicolson 1990, Woolf and Hall 1995). Hutchings et al. (1992a) categorise tools that have been developed in response to problems with hypermedia navigation and orientation as being direct access devices, history devices, or tours.

3.4.1.1 Direct Access Devices

Direct access devices are either textual or graphical in nature and allow a user to jump to any node in a hyperdocument with one action. Techniques used to navigate the traditional linear structure of information, e.g. index and contents, were carried over to hypermedia (Hall et al. 1990) and constitute the textual facet of this approach. Whilst investigating appropriate hypermedia navigation techniques, Simpson and McKnight (1990) found that navigation conventions employed in the paper medium were beneficial to a hypermedia environment. Benest (1990) offers a managed medium that is visually and navigationally based on the book metaphor. Adopting what is fundamentally a linear text approach, hypermedia link functionality is used to replicate the activity of using indexes, references and footprints/bookmarks in a book. Benest argues that the book metaphor deals with the problem of disorientation by maintaining a sense of position and openness. Traditional textual navigation techniques aside, the chief research focus centred on graphical representation techniques, the most popular of which was visual representation of a domain's cognitive or structural map (Allinson and Hammond 1989, Hutchings et al. 1993, Jonassen 1990, Kidd et al. 1992, Mukherjea et al. 1994, Zizi and Beaudouin-Lafon 1994). The problem with maps is that as the hyperdocument increases in size and complexity, the geography of a map becomes too difficult to be of any benefit (Hutchings et al. 1992a, Mayes et al. 1990b).

3.4.1.2 History Devices

The idea behind history devices is to provide a record of viewed nodes so the user can return to any previously visited, or recognise a node that has already been accessed. Monk (1990) offers the concept of a 'personal browser'. Here users can add nodes that are known to contain useful information to a personalised history, therefore facilitating the navigation of known locations.

3.4.1.3 Guided Tours

Tours work by temporarily removing the problem of navigation by providing a linear, predetermined route through a non-linear information space (Hill et al. 1993). Reasoning that the novice requires more structured guidance through the hyperdocument than the more experienced user, a number of authors subscribe to the use of guided tours in hypermedia systems (Allinson and Hammond 1989, Baird and Percival 1989, Furuta et al. 1997, Hauck 1996). The advantage lies in the ability for users to branch off at any time to explore the hyperdocument, but with the option of returning to the tour if disorientation becomes a problem.

3.4.2 Examples of Hypermedia Systems

A number of hypermedia authoring systems are available that offer functionality to help both authors and users of hypermedia applications. An introduction to five of these systems is offered below, for a more comprehensive synopsis see Mendes and Hall (1999, Mendes 1999). Intermedia is included in this review given its influence within the field of hypermedia. The other systems reviewed are included because they were considered as potential implementation environments for SCALPEL (see section 4.3.1).

3.4.2.1 Intermedia

Until the end of the project in 1992 Intermedia (Yankelovich et al. 1988) had been used to author educational applications, and influenced the design of open hypermedia systems (e.g. Microcosm, see section 3.4.2.3) with its ability to link documents created with different source applications

Intermedia consists of five integrated tools that facilitate the authoring and display of hyperdocuments. InterText is Intermedia's proprietary word processor; InterVal is used to create timelines; InterDraw allows the creation and display of graphics; InterPix allows scanned images to be created and displayed as bitmaps; and InterSpect which is used to view three-dimensional structures.

Each tool has integrated hypermedia functionality that allows links to be created and followed while material is being created or edited. In addition, the system supports bi-directional links. These are made explicit by small maker icons being visible at the source and destination of the link. To help with the navigation of the hyperdocument, Intermedia provides two types of tools: link explainers and maps. Link explainers provide link descriptions that indicate the nature of the information at the destination of the link while maps are employed to display the current document along with all other documents that are linked to and from it. As user navigates the hyperdocument, the map dynamically updates to reflect the change of position. To help with the conceptual structure of large and complex hyperdocuments, Intermedia offers the facility to assign descriptive properties to links and documents within the system. Referred to as typed links, if a particular document offers links to several other documents users can elect to view a list of explanations about each of the links. Explanations about the links and documents are not embedded within the documents but superimposed on them. Intermedia web structures maintain the explanation information and allow multiple users to follow and create links independently of each other, therefore allowing individual paths through the hyperdocument without any imposition to other users.

3.4.2.2 HYDESIGN

HYDESIGN (Marmann and Schlageter 1992) is the prototype of an extensible hypermedia system which supports the modelling of complex hyperdocuments. Marmann and Schlageter (1992) believe that a restriction of many existing hypermedia systems is that they do not support the design of higher level hypermedia objects beyond the basic node-link paradigm. In addition they believe that further restrictions hinder the modularization of the overall design and reuse of complex hypermedia resources. In HYDESIGN, different links and nodes can be categorised as being more abstract, or of a higher level, which allows for the definition of higher level operations and the subsequent modularization of complex hyperdocuments which in turn facilitate reuse.

The HYDESIGN model comprises two components, the HYDESIGN-engine, and the graphical user interface HYDESIGN-GUI. The HYDESIGN-engine is the data management component of the system and deals with the management of the differing node and link types, namely: Atomic Nodes, Reference Nodes, SBL-nodes, Links and Aggregate Links.

Atomic Nodes can be of any media type and can share their contents with Reference Nodes, which in turn offer the ability to share the same information in different semantic situations through different link contexts. SBL-nodes, considered to be of a higher level, represent particular composite nodes and offer the capability of defining (nested) local environments with particular behaviour. This is achieved through the use of three properties or attributes: Structure, Behaviour, and Locality. Structure represents the way nodes and links of a particular type are connected. Behaviour refers to the way a particular SBL-node and link responds to SBL-node oriented operations while Locality represents the sub-network of the hyperdocument that the SBL-node is associated with and is considered desirable when designing large complex hyperdocuments.

As a composite node, SBL-nodes can consist of Atomic nodes, further SBL-nodes and links. This model allows nested documents with conceptually differing levels of abstraction. Each SBL-node is classified as being on of the following types: Standard Hypermedia Graph (SHG), Directed Rooted Graph (DRG), Hierarchy, Sequence, or Set. A

SHG node can be made up of nodes of different types but may only have standard referential links. Hierarchy and Sequence nodes represents nodes organised as their classification implies while a Set represents nodes that have no organisation and do not contain links. Apart from the standard hypertext referential link, HYDESIGN introduces the concept of Aggregate links which are used with SBL-nodes classified as directed rooted graphs.

Furthermore, Marmann and Schlageter (1992) suggest that HYDESIGN supports navigation as well as query based access in an integrated approach. As a whole, they believe that HYDESIGN provides better support for the hypermedia design process by providing powerful structuring facilities.

3.4.2.3 Microcosm

Microcosm (Davis et al. 1992a, 1994; Fountain et al. 1990; Li et al. 1992; Hall et al. 1996) is an open hypermedia system that allows the integration of disparate pieces of multimedia information into a collection of documents. One of the most powerful features of Microcosm is the ability to link documents together, irrespective of format, without imposing any mark-up on the document (Davis et al. 1992a). This is a fundamental characteristic of an open hypermedia system. Microcosm maintains all information about links separately from the nodes. Consequently, all documents registered with Microcosm retain their original format, allowing them to be continually viewed by their host applications. This approach also allows several link bases, representing different views of the hyperdocument, to be modelled by the system. In the context of education, this functionality allows hypermedia environments to be authored to represent differing paths through the knowledge domain according individual needs, e.g. accommodate learners at different levels on the novice to expert continuum. In addition to the rich functionality provided by Microcosm, the architecture allows the integration of third party applications (Davis et al. 1994), thus allowing authors to add educational value to learning environments by developing simulations.

Microcosm offers authors four types of links: button, specific, local, and generic, all of which connect to a specific point in a destination document. The difference between the links lies with how Microcosm identifies their source anchors. Buttons are links from a visible anchor in a source document. Specific links are similar to buttons except that the anchor is invisible to the user. The anchor of a local link is invisible and is active on every occurrence of the specified anchor local to a given document. Generic links are also invisible and are activated when the anchor is selected from anywhere in the hyperdocument. Microcosm offers functionality called 'Show Links' to identify hidden (specific, local and generic) links within the hyperdocument. This process is particularly useful when searching for links that the author of the application may have included. Show Links works by checking each significant word in a user selection to see if a link is available. All relevant invisible links are presented to the user in a results window. Another powerful tool provided by Microcosm is 'Compute Links'. This functionality allows users to select some text and quickly find other documents that contain a high occurrence of the significant words in the selection. Links to relevant documents are constructed and presented to users in a results window. Third party applications that can communicate with Microcosm, via the Microcosm API, are referred to as 'Microcosm Aware' and have the ability to take advantage of the powerful linking functionality offered by Microcosm.

3.4.2.4 The Web

Originally developed to be a pool of knowledge that would allow collaborators in remote sites to share ideas, the World Wide Web (W3), or the Web, (Berners-Lee et al. 1994) has become the de facto hypermedia system. The Web is an open community of hypermedia-enabled document severs and readers on the Internet, a collection of networks that share digital information via a common set of networking and software protocols. Through a common authoring language, HyperText Markup Language (HTML), information is viewed via a browser. Browser programs run on client computers which talk to Web servers over to Internet to access and retrieve electronic documents.

Although links can point to any entity that can be displayed by a browser, the linking strategy employed is unidirectional. The link information itself is embedded as part of the document and cannot be changed without revising and republishing the document. A change to the address of a document requires every source document that refers to it to be changed to maintain link integrity. This makes it difficult for authors to move or delete a document.

The portability of the Web across different platforms and its ease of use have fuelled an explosion of Web users worldwide, which makes it the ideal medium for information dissemination and retrieval.

3.4.2.5 Webcosm

Webcosm evolved from work carried out on the Distributed Link Service (DLS) (Carr et al. 1994, 1995, 1998, Hill et al. 1995) which in turn was informed by experiences in providing open hypermedia services with Microcosm (see section 3.4.2.3). The motivation behind both the DLS and Webcosm is to provide similar services that are offered by Microcosm to the World Wide Web (see section 3.4.2.4). By separating the links from documents, multiple link bases can be authored to communicate different paths through the hyperdocument. This model is the first step in realising an infrastructure that offers generic and computed link functionality over the Web. Webcosm therefore offers the potential to combine the portability and large user base advantages of the Web with the rich linking capabilities of Microcosm.

3.5 Hypertext and Hypermedia in Education

The advent of hypertext and hypermedia systems in the late eighties had an overwhelming impact on the use of computers in education. Mayes et al. (1990a) comments: 'most of us working with exploratory learning environments in the CAL tradition were unaware until we read the proceedings of Hypertext '87 that our work was closely related to an idea incorporated in such disparate applications as information management, specialist writing and design environments'. McAleese (1989) suggests that one of the aims of hypermedia in education is to enhance existing learning strategies, and proposed one such strategy to be learning by exploration, a strategy that subscribes to an ethos that control over learning should be the responsibility of the learner and not the pedagogue. However, the benefits of hypermedia in education have not always been so direct as influencing core learning strategies. Hypermedia technology has also had an indirect role to play, exemplified by its impact on providing help to students through on-line help functionality (Colbourn and Cockerton-Turner 1990).

The potential of hypertext is grounded in its ability to structure information in radically new ways. Non-linearity of node and link structures has been favourably compared with the fixed linear structures of both paper-based and standard electronic texts (see, for example, Beeman et al. 1987; Cooke and Williams 1989; Duncan 1989; Spiro and Jehng 1990; Spiro et al. 1995; Rouet 1992; Stanton and Stammers 1990). In education, linear structures are typically seen as constraining both students and authors of conventional CAL environments. Duncan (1989) comments on the ability of hypertext systems to represent knowledge in different ways according to varying perspectives of users of the system. In contrast to linear text, the dynamic and interactive nature of hypertext can be harnessed to allow learners to explore the knowledge domains in ways not previously determined by the system. Beeman et al. (1987) argue that hypertext not only offers different information presentation techniques but can also encourage a particular (and desirable) style of thinking. They see education as encouraging a pluralistic, non-lineal critical style of thinking that encourages students to see the world in inter-related relativistic terms and not as isolated chunks of information, and view hypertext as a solution to the paradoxical attempts of education to realise non-lineal thinking through 'lineal communication, presentation and instruction'.

Researching hypermedia and its use in exploratory learning environments, Mayes et al. (1990a, 1990b) found the learner's freedom to explore significant in favouring hypermedia over learning environments that depend on building representations both of domain to be learned, and of the state of knowledge of the learner. The ability to access relevant information rapidly by direct selection is viewed as an essential characteristic of both hypermedia and exploratory learning environments. McAleese (1989) believes that 'freedom to browse, navigate and take part in a journey or voyage of discovery at will is the most distinguishing feature of hypertext', and that this feature can be used to exploit learning by exploration, one of the most powerful learning strategies. McAleese draws parallels between hypermedia and learning by exploration: 'Exploring in this way is occasioned by known concepts triggering off new ideas or by the learner attempting to make a link between two previously known ideas. In hypermedia this has its parallel with a node triggering another node and a learner or designer making a link or association between existing nodes'.

Hypermedia has always supported learning by exploration, from the information retrieval and dissemination systems underpinned by CD-ROM based encyclopaedias and reference books, to the current phenomena of the Internet. In education, the arrival of multimedia has brought about a resource based learning (Boyle 1997, Davis et al. 1993, Hall et al. 1995c) approach of integrating a wealth of information. However it is generally accepted that completely free navigation of hypermedia applications is insufficient for learning (Hammond and Allinson 1989, Laurillard 1993, Mayes et al. 1990a, McAleese 1989) and that functionality is required to facilitate and support the learning process. Believing the information available via the Internet to be inadequate, Henderson (1995) argues for the development and implementation of specific structured learning environments. Laurillard (1993) attacks resource based learning as not addressing students' needs. Referring to the authoring of courseware and providing tools to students to aid their navigation through it, Laurillard comments: "beneath the rhetoric of 'giving students control over their learning' is a dereliction of duty. We never supposed students could do that with a 'real library'; why should they be able to do it with an electronic one"?

3.5.1 Examples of Hypermedia Learning Applications

Examples of the use of hypermedia are presented below. Each system offers an example of the different navigation techniques and/or offers empirical evidence on the use of hypermedia in education.

3.5.1.1 Cell Motility

Hall et al. (1990) report on the development and evaluation of a hypertext interface to an interactive videodisc application used to teach cell biology to undergraduate students. Hypertext was seen to offer the appropriate technology to link text, graphics and video in a network topology that would encourage students to determine their own path through the material. The system was realised using the commercially available HyperCard, bundled by Apple with every Macintosh. Nodes in the hyperdocument contained text, graphics, or animations designed to present concepts in a dynamic way.

To evaluate the effectiveness of the system it was implemented as a complement to a first year Cell Biology course. Because of insufficient resources, 79 students were asked to use the system on a voluntary basis. Students who used the system were asked to fill out a questionnaire designed to elicit attitudes towards the system. At the end of the study 25 questionnaires had been returned and 15 students observed and interviewed. Aside from problems with software bugs none of the students interviewed criticised the quality of the teaching material or the systems overall value as a teaching resource. This enthusiasm is backed up by the questionnaire data which infers that the majority of students found the program to be an effective learning resource.

Subsequent to this first evaluation, Hutchings et al. (1992b, 1993, 1994) reported on the evaluation of the Cell Motility application using three successive cohorts of undergraduate students. During these evaluations students were required to use the system at least twice as a compulsory but non-assessed part of their course. A 15 minute introduction was given before each student was given between 35 - 45 minutes to investigate a specific subject. At the end of each session students were asked to complete a questionnaire. Analysis of the questionnaire data collected over the three years indicated a generally agreeable attitude to the system. Encouraged by the results of the Cell Motility program Hutchings et al. (1992a, 1992b) developed a set of authoring tools to allow users with no knowledge of HyperCard to create similar applications. The set of tools, called StackMaker, removes the need for domain experts to learn how to create text fields, highlight words or create links within the HyperCard environment. Instead text can be entered unformatted and linked simply by clicking one or two buttons provided by StackMaker.

3.5.1.2 Glasgow Online

Baird and Percival (1989) developed Glasgow Online, a hypermedia system (in this instance text and images only) on the life and times of the City of Glasgow. The aim of Glasgow online was not to 'emulate the many books, pamphlets, guides etc. on Glasgow' but to 'integrate, repackage and restructure that information to provide users with an electronic conspectus on the city and its culture' (Baird and Percival 1989). Baird and Percival refer to Glasgow Online as a community of information targeted at a heterogeneous population that includes tourists and businesspersons, as well as researchers, teachers, and students. The primary navigation tool provided by the system is the guided tour metaphor (see section 3.4.1.3).

Commenting on the importance of evaluating a system Baird and Percival believe that 'it is only by examining how readers use the system, that we, as authors, can assess whether the presentation of data and establishment of links is sufficient to meet user needs' (1989). As a consequence a comparative evaluation was undertaken to determine the effectiveness of Glasgow Online on two target population groups: 22 non computer literate women attending a 10 week course at the University of Strathclyde and 28 computer literate mathematicians from the Second European Conference of Mathematicians in Industry (ECMI). The evaluation exercise was conducted using a questionnaire consisting of twenty-three questions (based on a seven-point scale) gathered into five categories: first impressions of the system; clarity and organisation of information on screen; terminology and system information; learning from the system; speed and ease of use.

Starting with first impressions, Baird and Percival (1989) report that both groups thought that the system could be improved by offering better information retrieval facilities or faster link executions. In the second category on clarity and organisation of information, as would be expected, the computer literate ECMI group found the system clearer to use than the group of non computer literate women. For the category of terminology and system information, both groups expressed negative attitudes towards the feedback offered by the system. In the learning from the system category, both groups considered the system easy to learn. Finally, the ECMI group was less satisfied about system speed than the group of non computer literate women, an observation attributed to having no fixed preconceptions on what to expect. Baird and Percival (1989) comment that the results are of interest from a formative perspective in that they inform research into improving Glasgow Online.

3.5.1.3 Hitch-Hiker's Guide

Allinson and Hammond (1989) developed Hitch-Hiker's Guide, a learning support environment (LSE) for non-formal domains. They define a LSE as a 'set of tools to support the exploration of knowledge domains'. Psychology is a discipline put forward as an example of a non-formal knowledge domain given that it contains controversy, conflicting explanations of the same experimental evidence, historical perspectives and personal opinions. This is in contrast to most formal domains within mathematics or the physical sciences were there is often an underlying structure that holds together the raw data of the domain. Allinson and Hammond suggest that programmed learning environments are not appropriate for non-formal domains as they restrict students from developing their own critical abilities in coping with conflicting information. Instead they subscribe to the use of hypertext as it provides flexibility over the sequencing of material. They consider this important in realising learning environments to support non-formal domains: domains in which there is no single route to a unique answer but where deep understanding relies on the ability to 'organise and structure the material by isolating the main points and forming the necessary connections between the related items' (Allinson and Hammond 1989).

Hitch-Hiker's Guide was authored as a large network of display frames. To promote active rather than passive learning a series of multiple-choice quizzes, interactive demonstrations and selected experiments were made available to the student.

Evaluations conducted during the development of the system highlighted problems with students not always reaching their goals. To deal with this navigation and orientation problem (see section 3.4.1) Allinson and Hammond made use of indexes, maps (see section 3.4.1.1) and, by using the metaphor of a travel holiday, a guided tour (see section 3.4.1.3). An evaluation was conducted to assess the use of these navigation tools. Data were collected through a time-stamped performance log, a short questionnaire administered to each student at end of each session, and an extended questionnaire to a selected group of students. Questionnaire data were analysed from 59 undergraduate students using a Cognitive Psychology module.

The results show that the system is easy to use and that the navigational tools were used by the students to achieve their aims. In addition, it was observed that increased familiarity with the material results in a shift in navigation patterns from system control (i.e. tours), through shared control (i.e. maps) to user control (i.e. links). Allinson and Hammond (1989) conclude that there are navigation and subsequent learning performance problems when hypertext is used alone, and that these problems can be overcome by providing appropriate navigation tools. They comment on the success of Hitch-Hiker's Guide not only as a LSE but also as a means of investigating the success or otherwise of various navigation tools.

3.5.1.4 Software Teaching of Modular Physics (SToMP)

SToMP (Bacon 1994, Sinclair and Bacon 1997) is a Microcosm (see section 3.4.2.3) based hypermedia learning environment that offers self-paced student-centred learning modules to students of higher education physics. Text, sound, images and interactive simulations are used to delivery primary topics which are augmented by case studies, descriptions of applications, revision material, and reference materials including chapters of textbooks, a glossary, tables and lists of fundamental formulae and biographical synopses. The initial aim of the SToMP consortium was to provide a distance learning application. However, used in more than 140 establishments world-wide, in practice the use of SToMP ranges from a supporting resource to a complete replacement of selected lectures (Sinclair and Bacon 1997).

A formal evaluation of SToMP was carried out in four UK and one Australian universities and is reported in Calverley and Watkins (1995). The techniques employed to gather evaluation data were questionnaires, observations, structured interviews, and discussions with students. This mixed methods approach was chosen to satisfy a pre-defined agenda as well as to invite discovery of other aspects that may otherwise be overlooked. Questionnaires administered pre and post exposure emphasised attitudes towards computer-based learning and its comparison with traditional methods. In addition sections were authored to assess attitudes towards the SToMP material itself. Throughout the evaluation observations were made of students working through the SToMP material while structured interviews and informal discussions were conducted with individual and groups of students.

The evaluation highlighted four principle findings. First, the examination results indicate that the performance of students who use SToMP is no worse than those taught under the traditional method of teaching. Second, training is essential for both staff and students if the educational benefits of SToMP are to be realised. Third, SToMP should be implemented as a significant part of several modules in a degree course and not just as an addition to one or two lectures. Finally, staff teaching time is similar to the traditional methods. Calverley and Watkins (1995) conclude that the response to SToMP being used as an obligatory resource in support of traditional lectures is positive but highlight that attitudes towards its use as a replacement to lectures have not been so encouraging.

3.5.1.5 StrathTutor

Mayes et al. (1988; see also, Kibby and Mayes 1989; Mayes et al. 1990a, 1990b; Mayes 1993) developed StrathTutor, a hypertext system designed for tutoring applications. The system was conceived as a learning-by-browsing system that offers a number of ways to navigate through the domain material. As well as conventional hypertext links, StrathTutor realises the concept of dynamic hypertext, a method of generating relevant links at run time. StrathTutor employs a routing algorithm (Kibby and Mayes 1989) which matches frames of information based on author-provided attributes. During execution the student can choose to navigate by accepting a related frame offered by the system or can ignore the prompt and elect a different path though the material. A key feature of StrathTutor is the ability for students to interrogate the system by selecting a combination of attributes from which the system will compute a guided tour of the all the frames that have been coded with that particular subset of attributes. The tutoring element of StrathTutor encourages browsing in conceptual space rather than the spatially oriented network of conventional text. Mayes et al. (1990b) believe that under certain conditions disorientation in conceptual space facilitates deep learning. Since the point of discovery learning is that the student is continually trying to map the information being discovered onto their own developing framework of understanding, making wrong turns is part of that learning process.

Mayes et al. (1990a, 1990b) carried out evaluations of StrathTutor using methods of protocol analysis (Ericsson and Simon 1984). The technique involved transcribing and annotating the dialogue resulting from pairs of subjects making collaborative decisions about how to proceed through the material. Analysis was performed on a number of interactions between same-sex pairs of students who used StrathTutor to learn about glaciation. The results reveal that active exploration of the hyperdocument is not the same as active exploration at the conceptual level. Mayes et al. (1990a) report that in some cases the students were too absorbed in the task of learning the material to be concerned with learning the navigation functionality offered by the system. They found that students either learn to navigate or they learn the material, but not both, and offer this as evidence in support of 'cognitive overhead' influencing to effectiveness of a learning environment (Mayes et al. 1990b) in as much as students allocate most of their cognitive resources to learning the domain material and neglect to learn functionality that would actually make this easier. It was also observed that students tended to regress to an earlier stage in their use of the system as their attention became more focused on the material. However, Mayes et al. (1990a) remained confident that an increasingly flexible use of the system would pay dividends.

3.6 Critical Reflections on Hypermedia in Education

Enthusiasm for hypermedia has not been shared by all. McKnight et al. (1991) suggest that since 'many people have seen hypertext as a replacement for printed information, it is not unnatural to suppose that it might have an impact in the field of education', but err on the side of caution by proclaiming that inline with any new technology that is brought into education 'more claims have been made on hypertext's behalf than the experimental evidence has been able to support'. This stance is maintained as a consequence of their own empirical evidence that leads them to conclude that 'for some texts and some tasks hypertext is not the universal panacea which some have claimed it to be' (McKnight et al. 1990). As an information storage and retrieval system Laurillard (1993) believes hypermedia to be a 'very well-designed medium', but in the domain of education, 'enabling the student to develop their academic understanding, it has little to offer'.

Other criticisms regarding the euphoria that surrounds hypermedia in education are based on the misconception that hypermedia learning environments are easy to author. This view is shared by Laurillard (1993) who believes that the 'ease of authoring the links between nodes has created great excitement in the world of educational media because it appears to solve the perpetual problem of how to create sufficient good quality courseware very easily'. McKnight et al. (1989) argue that the task of the document author gets trivialised amongst the 'general euphoric tone of much writing about hypertext', and suggest that authoring requires considerable effort on the part of the author if user requirements are to be sufficiently supported.

3.7 Summary

This chapter discussed the origins of hypermedia and its influence on education. The problems with navigation and orientation of hypermedia were introduced along with the differing approaches of employing tools to overcome these problems. The impact of hypermedia on education practises was discussed together with the potential of hypermedia to afford situated, dialogue rich learning environments that take advantage of students' natural learning capabilities. Finally, a critical reflection on the use of hypermedia in education highlighted concerns about the lack of empirical evidence to support such claims, and the misconception that hypermedia learning environments are easy to author. The next chapter discusses the implementation of SCALPEL, a hypermedia environment design to replace traditional pathology laboratory teaching.

Chapter 4: Pathology Practicals and Authoring SCALPEL

4.1 Introduction

The previous chapter discussed the influence of hypermedia in education. This chapter discusses the design and implementation of SCALPEL, a hypermedia learning environment charged with replacing traditional pathology laboratory teaching. The chapter focuses on the domain and design knowledge, and is structured around the principle disciplines of design knowledge: conceptual and presentation design (see section 3.3). Beginning with conceptual design, the traditional pathology laboratory practical environment is discussed before offering, through a synopsis of participant observations, a qualitative empirical perspective on student learning in this environment. The domain knowledge is introduced and reasons are given for keeping the domain material identical in the traditional laboratory and SCALPEL environments before a discussion is offered on the education principles that underpin the learning architecture of SCALPEL. The chapter then addresses the presentation design of SCALPEL. Choosing an implementation environment is discussed before highlighting the consideration of hardware specifications and aesthetic issues. Next a discussion is offered on the use of a MCQ engine. Finally each of the media types utilised by SCALPEL are presented with a focus on their pedagogic value and presentation quality from a technical viewpoint.

4.2 Conceptual Design of SCALPEL

This section addresses the conceptual design of the SCALPEL courseware environment. Conceptual design is concerned with creating the learning architecture of the system and involves, in consultation with domain experts, the structuring of the domain knowledge.

4.2.1 Pathology Laboratory Practicals

Recognising the need to improve student attendance and appreciation of pathology practicals at the University of Southampton, McCullagh and Roche (1992) introduced case-based teaching (see section 3.2.4). Fictitious case histories and a series of investigations and questions were designed to demonstrate the importance of basic science in a clinical context and to introduce new concepts and diseases to the students. The practicals were set in a traditional laboratory environment, and demonstrators with physiological and clinical expertise were available to discuss the pathology material with students. The format of pathology practicals required students to work through case-based pathology case histories presented in a pathology practical notebook (see Appendix 5). Typically, a pathology practical presents students with a case history which is used in conjunction with supplementary material including: anatomy and other images, data charts and tables, and histology slides to be viewed under a microscope. Questions relative to the context of the case history are periodically presented to students. All answers are recorded in the pathology notebooks. Unless demonstrator assistance is sought, no feedback is offered to the students with regards to the validity of their answers. In this environment students have a measure of control over their learning; they are able to determine the speed at which they work through the domain material, and to what extent they wish to engage in peer group and demonstrator discussions.

As a single component of the students' pathology learning, pathology practicals were designed to complement the two other principle components: lectures and tutorial discussions. Through a learning-by-doing experience (see section 3.2.1) a practical builds on the pathology material delivered in the preceding lecture, while post practical group tutorials address student misconceptions by encouraging dialogue and reflection about the concepts under consideration.

4.2.2 Observations of Learning Patterns in the Traditional Laboratory

Participant observation (see section 5.4.3) was carried out in the traditional laboratory with a view of formulating a situated, holistic and well-rounded account (Bryman 1998) of a typical laboratory practical. Figures 5, 6, 7 and 8 are presented to complement the following synopsis of observations during a typical laboratory-based practical.

The overall atmosphere in the laboratory was very informal. Most students had arrived at the laboratory and began working through the practicals before the official start time, and the arrival of demonstrators. Students took full advantage of the large work surfaces by laying out text books and other supplementary pathology material, e.g. anatomical images etc.

There were no stipulations about the size of peer groups, prompting students to work individually or in groups of their own design. Student groups varied in size from two to as many as eight. In addition to the in-group discussions, some students were observed temporarily moving between groups to discuss the practicals. Generally, demonstrators were consulted only after concerted efforts by group members to resolve misconceptions through dialogue, although an occasion was observed where a group of two called for demonstrator assistance having made no attempt to resolve a problem using the pathology material available to them.

When demonstrators were consulted, questions were asked on behalf of the group with most students participating in the ensuing open discussion. Demonstrators tended to deal with students' enquiries by asking additional questions to those presented in the practical book. Another approach was to tell anecdotes about the domain material. Situating the pathology material in this way engaged the interest of students and was observed to be effective in resolving problems.

The noise level in the lab was consistent with a cohort of approximately 80 students concurrently engaging in conversation. Students took advantage of social interaction made possible through the open layout of the laboratory and, utilising the large work surfaces, the ability to layout the course material for the group to see.

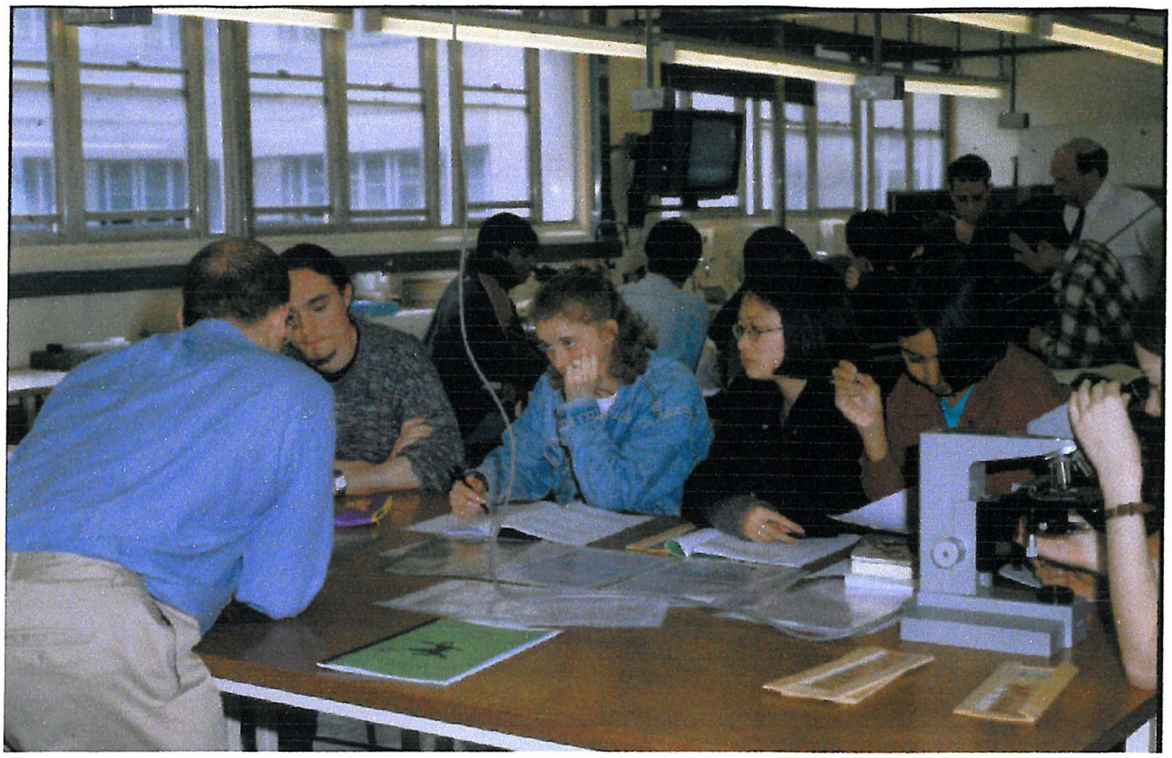


Figure 5 Demonstrators discussing pathology material with students.

One observation considered disadvantageous was that the environment made it possible for some students to take a backseat. Students were observed relying on other group members to work through the practical and answer the questions. Furthermore, other students with extravert and domineering personalities were observed taking control of the group, sometimes dismissing the input of others. In these situations, the extent at which all students benefit from the laboratory experience is brought into question.

Where the use of microscopes was required, it was observed that a single student in the group would manage the microscope, and share the explanation with other students to answer the relevant question in the practical book. When necessary, demonstrators discussed histology slides with the entire cohort by taking advantage of the monitors situated around the lab.

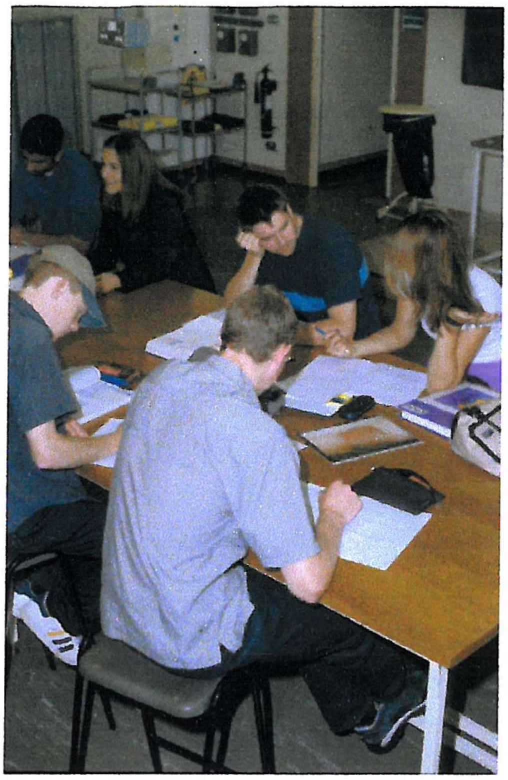


Figure 6 One to one peer dialogue.

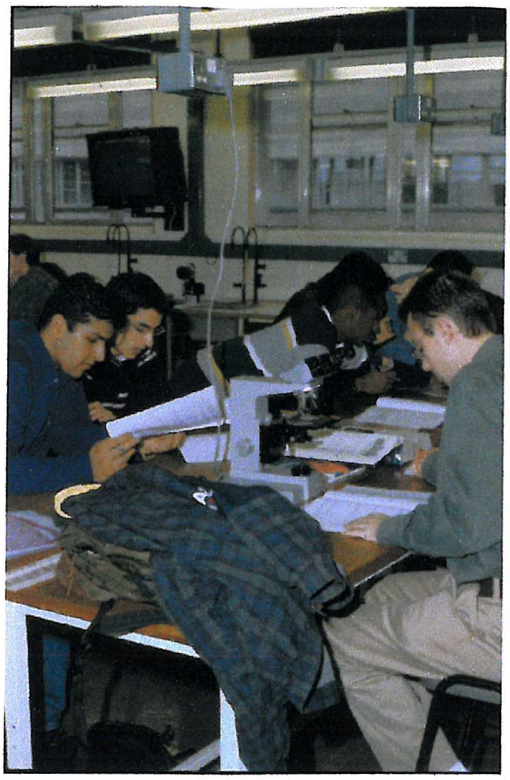


Figure 8 Group discussions.

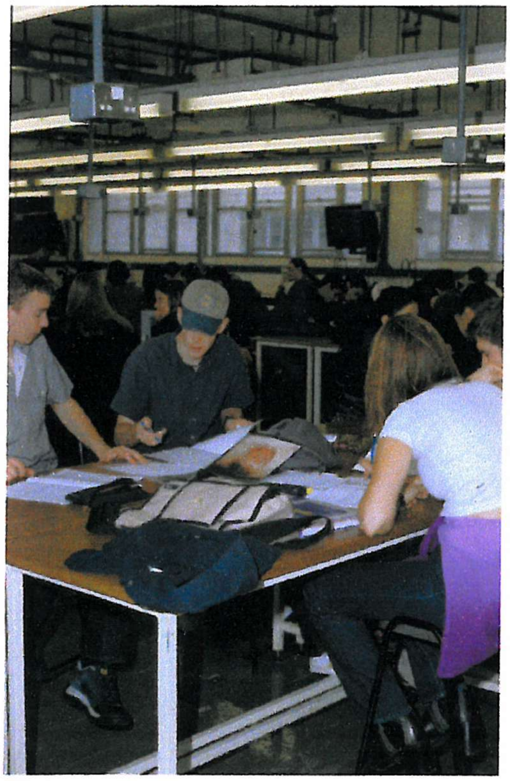


Figure 7 Small groups were not uncommon.

Pictures of students working in the traditional laboratory. Peer groups varied in size. Microscopes (Figure 8) were integral to the practicals although it was often the case that a single student would perform the histology analyses and share her observations with the rest of the group. Monitors around the laboratory (Figures 7 and 8) were used to deliver video material. Demonstrators were available to discuss misconceptions with the student, while the large work areas were conducive to using text books.

4.2.3 Educational Theory

Presented schematically in Figure 9, parallels can been drawn between the pathology practicals and the construction stage of Mayes' conceptualisation cycle (Mayes 1995, McKendree and Mayes 1997) (see section 3.2).

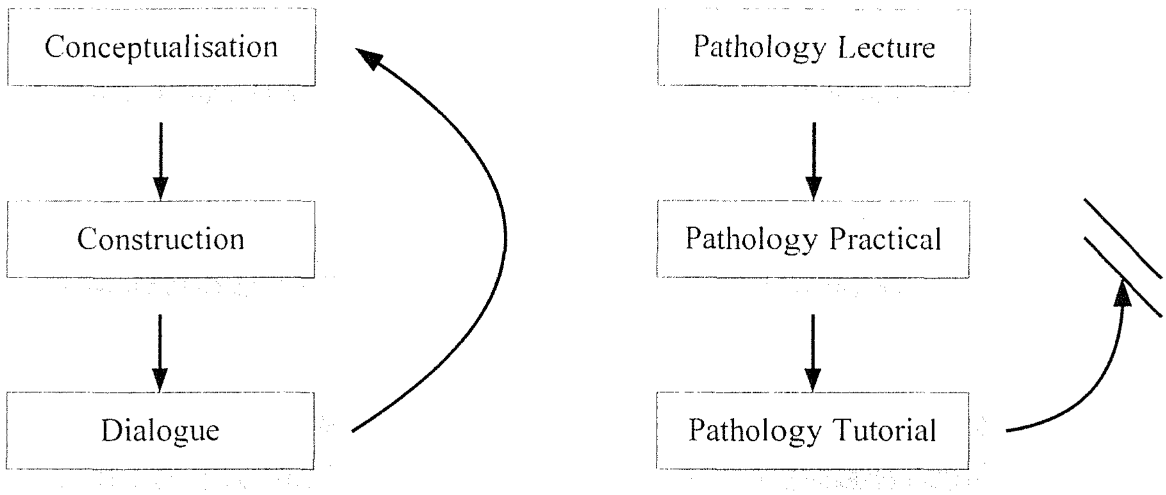


Figure 9 Analogy between Mayes' Conceptualisation Cycle and the Pathology Lecture, Practical and Tutorial.

The pathology lecture delivered prior to a practical is analogous to the conceptualisation stage, and post practical group tutorials where students discuss and reflect on the domain material, the dialogue stage. Breakdown in the analogy is evident when considering the cyclic nature of Mayes' re-conceptualisation. The pathology instantiation does not have a mechanism for reintroducing students with misconceptions back into the cycle: additional lectures and the ability for students to revisit the practicals in the laboratory are not available. If misconceptions are not captured in the tutorial, they remain for all but the most resourceful of students.

The breakdown in the re-conceptualisation cycle, as a consequence of delivering pathology practicals in the traditional pathology, is addressed by moving away from the traditional laboratory and integrating courseware. This is achieved by offering students the opportunity to return to the practicals to address any misconceptions. The ability for students to visit and revisit the pathology practicals without any timetable constraints was considered a key advantage to implementing a courseware environment.

Hoffer and Barnett (1990) believe that an important aspect of courseware development that is overlooked is the integration of computer-based materials with the curriculum, with most applications being used as supplementary material: placed in libraries and used by students on their own initiative. Hoffer and Barnett suggest that the most effective use of courseware is to integrate them into the standard curriculum. In the context of medical education, they promote the use of computer-based educational systems for laboratory exercises. Integrating SCALPEL into the curriculum was considered important to the success of replacing the traditional laboratory environment with courseware.

4.2.4 Domain Experts

Authoring courseware requires domain experts to provide content. This has been highlighted as a problem in past courseware projects where 'finding specialists with sufficient time and inclination to become involved in the authoring of applications' has proven to be one of the 'greatest hurdles faced' (Hutchings et al. 1994). Given that the head of the pathology department motivated this initiative, authoring SCALPEL was not hindered by this problem.

4.2.5 Conceptualising Hypermedia Based Pathology Practicals

When deciding upon the best approach for implementing hypermedia based pathology practicals, a prime consideration was that the domain material should be as identical as possible to the content of the practicals delivered in the traditional laboratory. This decision was taken on the basis that case-based reasoning, the format of the traditional practicals, is well informed by education theory (Schank 1994, Schank and Cleary 1995), and has empirically been shown to be successful in the context of pathology practicals at

Southampton (McCullagh and Roche 1992). In addition, keeping the domain material identical had implications from an evaluation perspective. It was reasoned that confidence could be given that statistical differences from the evaluation of the two teaching environments would not be biased by differences in content, thus allowing any such findings to be attributed to differences in the enabling environment and not the domain content.

Considering the potential of undermining the empirical comparison between the laboratory and SCALPEL environments highlighted what Conklin (1987) refers to as the cognitive task scheduling problem. Would the students be too occupied with learning the functionality required to navigate SCALPEL, to concentrate on the domain material? The cognitive overhead required by users to navigate hypertext systems, and any new system in general, deflects information processing capacity away from the domain material and to the meta-level details of working the software application. With the potential of the task scheduling problem biasing those students assigned to SCALPEL during the cross-over trial (chapter 7; see also, Kemp et al. 1997, 1998), and acknowledging the differing levels of IT literacy amongst students, links were only authored to provide access to the same supplementary material offered in the traditional laboratory.

Mendes (1999) defines structural knowledge as 'the knowledge of how concepts within a domain are interrelated' and nodes as 'the chunks of information (documents) that comprise a hypermedia application'. The format of the laboratory-based practical material (see section 4.2.1) transferred well to a hypermedia-based composition. In the laboratory explicit text in the notebooks is used to refer students to supplementary material, e.g. 'see slide 1', or the more indirect 'slide 1 contains sections of…'. This structural knowledge is synonymous to links within a hypertext environment. The nodes are predefined since each practical case-base constitutes a primary case-base node and supplementary resource nodes, and the structural knowledge is predefined through explicit text in the case-base representing the anchors for the contextual links to the supplementary material. Bernstein (1998) calls a hypertext structured like the pathology practicals a cycle: a user starts at a node (case-base) and follows a path (supplementary material) before returning to a previously visited node (case-base) to departing along a new path.

In keeping with the format of the material as presented in the traditional laboratory, the main visual components of each SCALPEL practical were conceptually determined to be: a main practical case window, and an MCQ engine. The main practical window would communicate case histories, refer (link) students to supplementary material, and pose questions to students in much the same way as the course notebook in the traditional laboratory. The MCQ engine would offer feedback on the validity of the answers chosen for each of the questions posed.

4.3 Presentation Design of SCALPEL

This section offers an insight into the physical design of the hypermedia based SCALPEL courseware environment. Issues pertaining to the choice of a hypermedia system and considerations of authoring hypermedia are presented relative to the realisation of the conceptual design of SCALPEL presented in the previous section.

4.3.1 Hypermedia Authoring Models, Methodologies and Systems

The use of hypermedia in education was discussed in chapter 3. Today the World Wide Web (Berners-Lee et al. 1994) is the predominant hypermedia medium in widespread use, however, a number of hypermedia authoring systems are available that offer functionality to help both authors and users of hypermedia applications, some of which have be discussed in chapter 3. Before authoring a hypermedia learning environment, it is important to have an understanding of the different hypermedia environments and their respective features in order to be able to make an informed decision about best practice. Mendes and Hall (1999, Mendes 1999) offer a comprehensive synopsis of hypermedia authoring models, methodologies and systems. They define models and methods as facilitating the orderly arrangements of ideas; methodologies as representing bodies of methods used in particular activities or processes; and systems as instantiations of applications.

Mendes and Hall (1997a, 1997b, 1999, Mendes 1999) carried out a qualitative survey aimed at analysing the processes involved in the development of hypermedia applications for education. From a volunteer sample of thirteen authors who collectively authored twenty hypermedia applications for education, Mendes and Hall highlighted the requirements of potential users of hypermedia learning environments. These requirements include a model's, methodology's or system's ability to support: a desirable authoring approach, e.g. top-down; the conceptual design, e.g. structuring the knowledge domain without considering link strategies; and physical capabilities, e.g. presentation of cognitive maps, contextual and typed links, and the ability to support different resources, e.g. text, images, video etc. A further requirement highlighted as being important by the qualitative evaluation was the support of co-operative authoring (Melly 1995) for the differing phases of design and implementation, e.g. interface design, structural knowledge, links, and authoring the contents of nodes.

Considering these requirements, and more general features from the literature, namely the mandatory support of ill-structured domains (Spiro and Jehng 1990, Spiro et al. 1995), Mendes and Hall (1999, Mendes 1999) performed a feature analysis (Kitchenham 1996a, 1996b; Kitchenham and Jones 1997a, 1997b, 1997c, 1997d) of a number of hypermedia models, methodologies and systems with a view of identifying those which best fit the requirements of potential users. Their findings found that none of the models and methodologies evaluated supported the structuring of ill-structured domains. Of the hypermedia systems under evaluation only HYDESIGN (see section 3.4.2.2), Microcosm (see section 3.4.2.3) and Webcosm (see section 3.4.2.5) offered support of contextual links.

Consequently, informed by the research of Mendes and Hall, the choice of a hypermedia system for the implementation of SCALPEL was narrowed to include the Web, HYDESIGN, Microcosm and Webcosm.

The Web was eliminated for two principle reasons. First, without development of additional functionality (Anderson 1997), the Web was considered inadequate for the purposes of delivering a hypermedia learning environment. Second, the performance of the Web needed to be considered, even on the broadband network infrastructures of Universities. The development of SCALPEL began in 1996, at this time Webcosm was still in its infancy and as such had little in the way of empirical support from the literature, as was the case with HYDESIGN. However, the final decision to choose Microcosm to implement SCALPEL was not only made on the basis of the functionality offered and the extensive literature to support the use of Microcosm in education (Bacon 1994, Calverley and Watkins 1995, Davis et al. 1992b, Davis et al. 1993, Hall and Colson 1991, Hall et al. 1995a, 1995b, 1995c, Piper et al. 1995, Sinclair and Bacon 1997, Woolf and Hall 1995), but also because the University of Southampton offers support to developers and users of Microcosm based educational courseware. A key consideration was Microcosm's open architecture and therefore its ability to integrate third party applications. This proved most useful when consideration SCALPEL's requirement for MCQ functionality (see section 4.2.5). The SToMP consortium (see section 3.5.1.4), researchers of an existing Microcosm-based learning environment, had implemented such functionality and made it available to this research.

4.3.2 Hardware Considerations

The media types employed to realise the corpus of pathology material within SCALPEL include text, graphics, animation and video. Consequently, consideration was given to the technological limitations that govern the quality of media sources presented by hypermedia applications (Furht 1994). A compromise between quality, storage space and retrieval time is often required to achieve acceptable results: higher quality requires more storage space and leads to longer retrieval and display times.

When deciding upon values for the variables that influence the presentation quality for each of the media types embraced by SCALPEL, the specification of the delivery computers had to be taken into consideration. To this end, SCALPEL was developed to work efficiently on two clusters of low-end workstations:

Cluster 1 - 13 x Elonex PC433, 8Mbytes RAM, 8 bit colour, Windows 3.11

Cluster 2 - 27 x Elonex PC466, 8Mbytes RAM, 8 bit colour, Windows 3.11

4.3.3 Aesthetic Considerations

Stark (1990) offers empirical evidence in favour of using 'pop-up' as opposed to 'replacement' windows in closed loops: hypermedia structures that allow a user to follow a link to a node whose only exit is a return to the source node. Whilst investigating patterns of undergraduate students interaction with hypermedia systems, Hutchings et al. (1994) found that the use of hypermedia navigation tools were influenced by the nature of the interface. When tools were in constant view, participants made significantly more use of them than when they had to be requested or found under a heap of overlapping windows. Consequently, windows within the SCALPEL environment were arranged to keep overlapping to a minimum (see Figure 10).

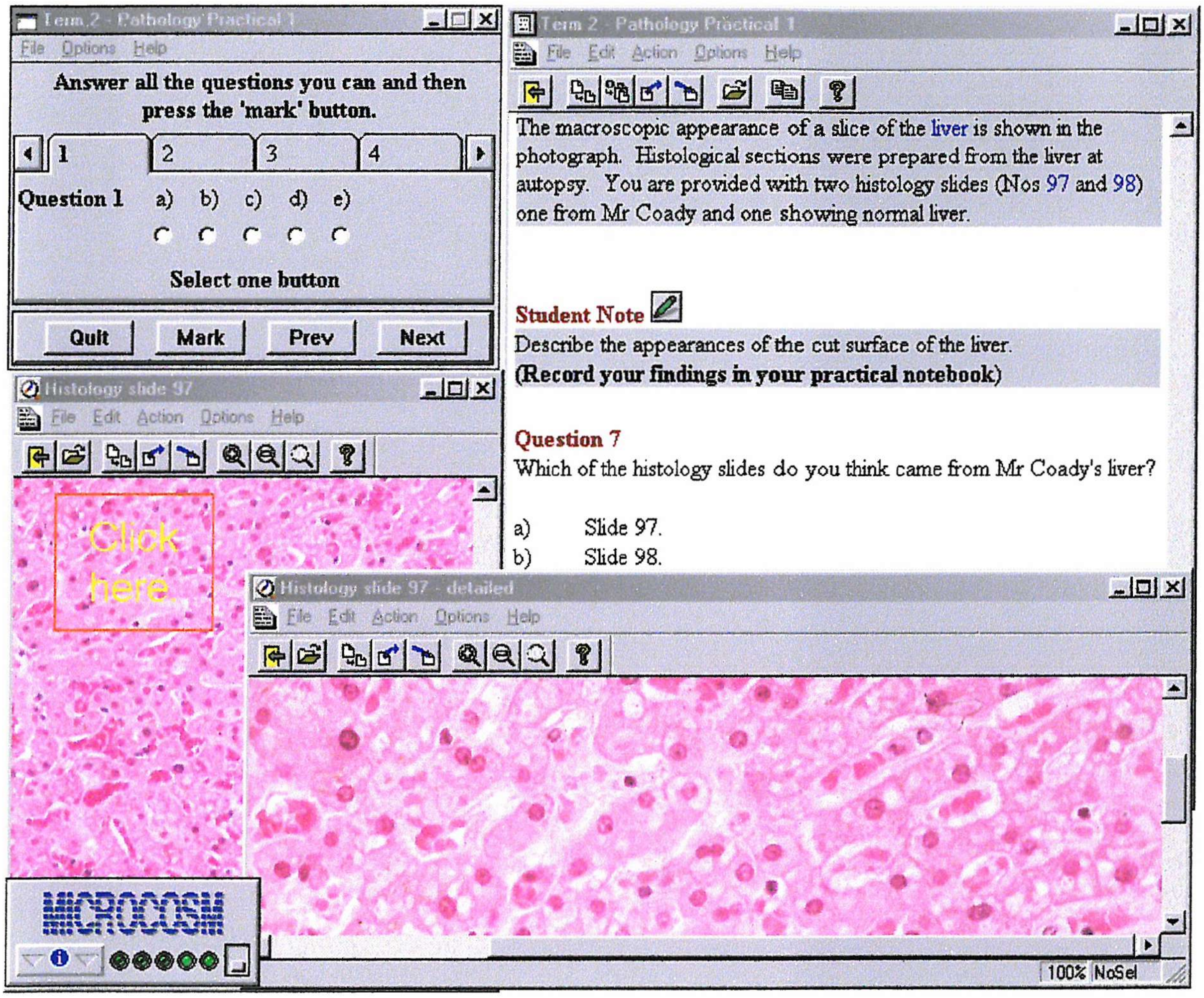


Figure 10 The SCALPEL interface showing the MCQ engine (top left), the main practical text, two histology images, and the microcosm status function.

Preece et al. (1994) comment on the importance of offering visual feedback on user actions: 'If feedback is not current, correct or closely expressed a user may think or do the wrong thing. Moreover, if there is no feedback the user will be left wondering what is happening and may feel that he is not in control'. Process visualisation (Baird and Percival 1989) is the concept of giving visual feedback to users about what is going on in the system. If any time at all is to elapse during a process, the user should be made aware that something is happening and that an error has not occurred.

The best example of process visualisation is when the mouse pointer in the Microsoft Windows environment turns to an hourglass to reflect that a process or operation is active. In the context of SCALPEL, and hypermedia in general, delays are not uncommon between a user activating a link, and the destination node being displayed. Microcosm's status function was utilised to offer visual feedback to students on the status of the computer's side of an interaction. The status function works by presenting a line of five green lights (see Figure 10) that turn on and off in sequence when the system is idle. When the system is acting on an event, e.g. an activated link, all five lights turn to red, informing the user that their request is being dealt with.

4.3.4 The Multiple Choice Question (MCQ) Engine

Central to the SCALPEL courseware is an MCQ engine designed by the SToMP consortium (see section 3.5.1.4) to provide students with feedback when answering the case-based questions (see Figure 11). In the context of SCALPEL, the MCQ engine works in conjunction with the main practical window. Whilst working through the case-base, if a question is encountered, the student is required to answer the question by moving to the corresponding question number in the MCQ engine, select an answer (e.g. a, b, or c, etc.), and then press 'Mark' button. The MCQ engine then informs the student if the question was answered correctly. Employing an MCQ approach to evaluation in courseware requires an understanding of the pedagogic value of feedback. Laurillard (1993) believes that 'action without feedback is completely unproductive for a learner'. Arguments for subscribing to the use of feedback are underpinned by the importance of feedback in everyday life. Laurillard (1993) highlights this believing that 'as we learn about the world through acting on it, there is continual feedback of some kind, and if we can make the right connection between action and feedback, then we can adjust the action accordingly and this constitutes an aspect of learning'. To this end, as opposed to simply offering "yes" or "no" answers in response to a student's MCQ answer for a given question, the MCQ engine is authored by domain experts to give informative feedback regardless of whether or not a question is answered correctly (see Figure 11).

There are no constraints on the number of times a question can be attempted, and the students performance is not assessed (although the SToMP MCQ engine is capable of assessing students).

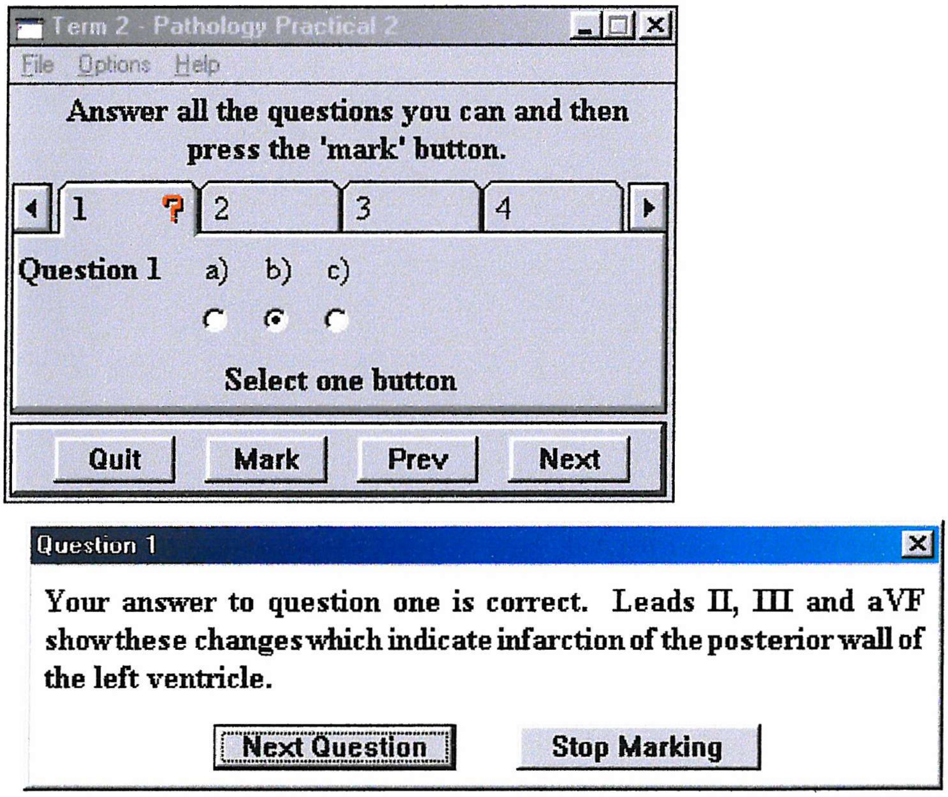


Figure 11 Example of SToMP's MCQ Engine in SCALPEL offering feedback.

As a Microcosm aware application (see section 3.4.2.3), the MCQ engine offers additional functionality that allows relevant supplementary domain knowledge to be accessed by means of Microcosm's show links feature (see section 3.4.2.3). This requires keywords to be explicitly identified by domain experts and mapped to supplementary material registered with Microcosm at design time. For any given question, if a domain expert has mapped supplementary material, students can look at a list of all relevant Microcosm links by pressing the 'Explain' button on the MCQ engine. If a domain expert has not explicitly mapped a question to supplementary material, the 'Explain' button is not presented to the students. The MCQ engine is configured at runtime by reading a configuration script (see Appendix 6). A drawback of the MCQ engine is that configuration scripts have to be authored manually. Although the complexity of doing so is reduced by following previous examples, an interface could easily be developed to automate this task.

4.3.5 Static Domain Material

4.3.5.1 Text

The structure of the original pathology practicals was carried over into SCALPEL giving a consistent look and feel for all pathology practicals. All text documents were prepared using Microsoft Word and imported into SCALPEL as RTF (Rich Text Format) documents.

4.3.5.2 Images

Pathology imagery made available to students in the traditional laboratory was digitally scanned for use within SCALPEL. The following image sources were used:

- Scanned 35mm slides of images used in the laboratory.
- Scanned 35mm Histology slides, captured with a digital camera coupled to a microscope.
- Scanned 35mm slides of pathology cultures prepared by laboratory staff.

Each image was scanned to a specification of 24 bit colour 150dpi and subsequently reduced to 8 bit colours for use within SCALPEL. The reduction in colour resolution was made to improve image quality and retrieval times after considering the specification of the delivery computers.

4.3.6 Dynamic Domain Material

Dynamic media such as animation, video, and voice contribute 'a dramatic, attention grabbing dimension to multimedia' (Boyle 1997). Sutcliffe et al. (1997) comment: 'Dynamic media attract attention much more effectively than static media such as text and still images'. Within the context of this research, video and animation were integrated with SCALPEL to add impact to the presentation of the learning environment.

4.3.6.1 Video

Blattner (1994) believes that education becomes 'more effective and engaging with multimedia interfaces that incorporate video and audio'. Through interactive laser videodisc hardware, video proved to be an effective resource in early computer based learning environments (see, for example, Hall et al. 1990; Hall et al. 1995b; Hoffer and Barnett 1990; Hutchings et al. 1992a, 1992b, 1993; Shortliffe 1990; Speedie and McKay 1990; Woolf and Hall 1995). From a technical perspective, having resolved the difficulties of displaying full-motion video on computers, namely as a result of international standards for still and moving picture compression (JPEG and MPEG respectively) and performance improvements in affordable hardware (Furht 1994), the prominent research focus shifted to content and context-based retrieval of video (see, for example, Flickner et al. 1997; Ho-Shing and Chan 1997; Lewis et al. 1996; Pentland 1997; Zhang et al. 1997).

Recognising that video can contribute to the impact of a hypermedia learning environment, video sequences identified by domain experts as being relevant to the pathology practicals were captured in digital form and authored into SCALPEL.

Quality of the digital end product was influenced by two primary considerations: the quality of the original analogue recording, and the specifications of the delivery computers. To achieve good quality, analogue video segments were recorded by professional audio-visual members of staff. Taking into consideration the specification of the delivery computers, the analogue video sequences were sampled to produce 8 bit colour sequences at 15 frames per second. The end result was the realisation of full motion effect Microsoft AVI video files to be displayed by the Microcosm Graphics Viewer.

As a timed based medium, once video starts it takes control until it has finished. Boyle (1997) comments on the importance of giving users control over the video, allowing them to interrupt sequences and navigate through the frames. The Microcosm Graphics Viewer's Media Bar (see Figure 12) offers control over video segments by allowing users to start, stop and step forward and backwards through the frames of a video sequence. This functionality allows students using SCALPEL to view the video segments in consultation with the case-base text and other supplementary resources.

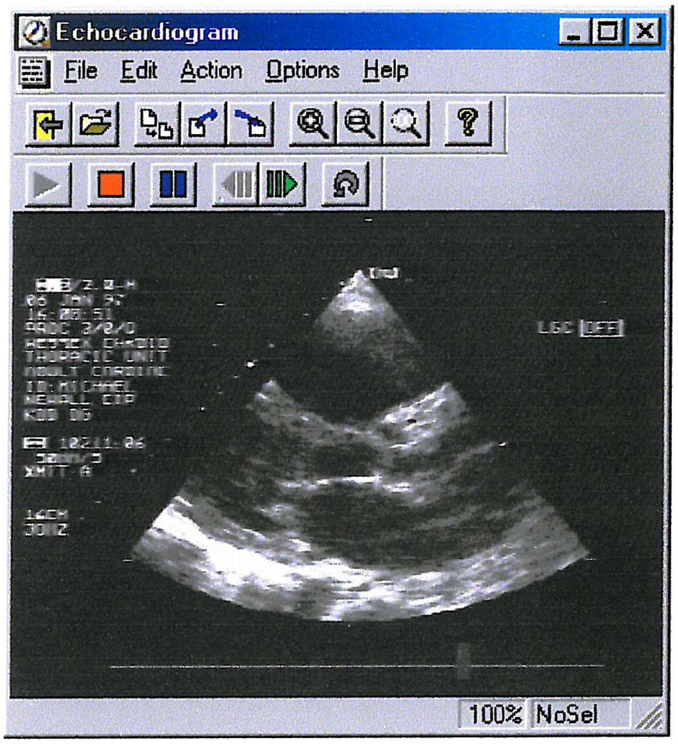


Figure 12 Snapshot of an echocardiogram video segment as displayed through the Microcosm Graphics Viewer

4.3.6.2 Simulations and Animations

The educational benefits of employing simulations or animations in learning environments were discussed in section 3.2.1. Advantage was taken of Microcosms' open architecture (Davis et al. 1994) to add impact to the presentation of an electrocardiograph (ECG) trace within SCALPEL. In the traditional laboratory, students were required to parse an A4 representation of the ECG trace. By employing animation concepts, students using SCALPEL were able to observe a moving ECG trace as if situated in a real clinical setting (see Figure 13).

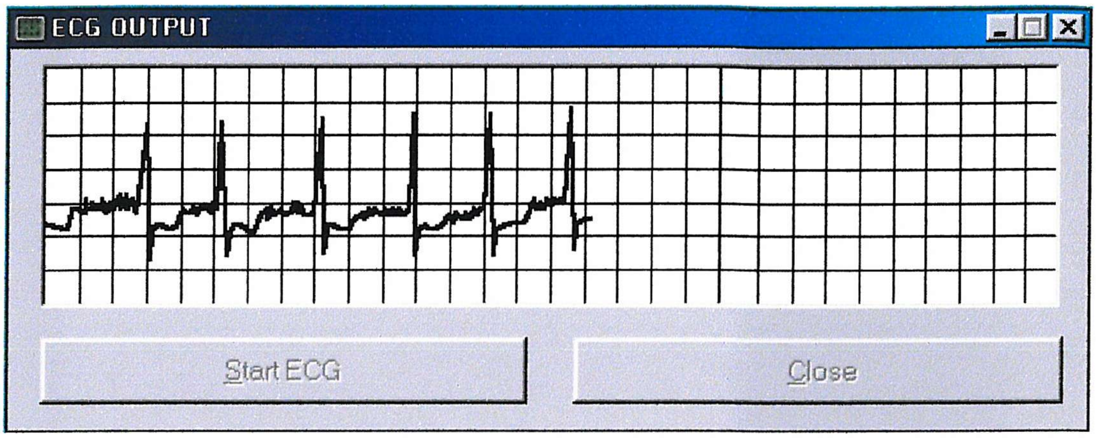


Figure 13 Snapshot of an ECG animation

4.4 Resources and Authoring Effort

As discussed in section 3.6, authoring a hypermedia application is considered to be a non-trivial task. The authoring effort required to realise SCALPEL was reduced for the following reasons: domain expert availability; predefined hyperdocument structure and content; and the open architecture of Microcosm.

A significant consideration is the availability of domain experts to provide the domain material for the hyperdocument. As discussed in section 4.2.4 this was not a problem for authoring SCALPEL. The structural knowledge and nodes of the hyperdocument had already been implicitly defined by virtue of the way the pathology material was presented to students in the traditional laboratory (see section 4.2.5), and because of the open architecture of Microcosm (see section 3.4.2.3) there was no need to develop a new MCQ engine as the implementation used by SToMP (see section 3.5.1.4) could be used.

It is difficult to be precise when attempting to quantify the authoring effort of any software application. However, as a result of the factors discussed in this section, it is estimated that a single pathology practical took approximately 48 person hours to author, with the majority of time as a consequence of the requirement for a domain expert to revise the case questions for use in the MCQ engine (i.e. provide a list of alternative answers, and relevant feedback).

4.5 Summary

This chapter introduced the key issues addressed during the implementation of SCALPEL; discussing considerations pertaining to the domain material, domain experts, choice of implementation environment, and media types. The rest of this thesis discusses the evaluation and refinement of this learning environment. This begins in the next chapter with a discussion on the research methodology and techniques that underpin the subsequent evaluations.

Chapter 5: Research Methodology and Techniques

5.1 Introduction

Having authored SCALPEL to replace traditional pathology practical teaching, an understanding was required of its effectiveness as a learning environment. This chapter introduces the research methodologies and techniques that underpin the empirical studies designed to offer such an understanding. These empirical studies are presented in subsequent chapters of this thesis. Confidence given to the inferences of empirical studies can only be determined on the basis of the information presented (Moser and Kalton 1979). Altman (1991) believes that

> 'A consequence of the fundamental role of study design is that the most important part of a research paper is the methods section.'

In recognition, this chapter presents an account of the different research designs and techniques underpinning the evaluations presented in this thesis. Information is provided here to make it possible and valid to draw more general conclusions from the inferences presented in subsequent chapters (Oppenheim 1998).

The empirical evaluations presented in this thesis followed closely the research process (see Figure 14) proposed by Bryman and Cramer (1997). The sections of this chapter consider each stage in turn and address issues that influenced the evaluations and their subsequent conclusions. Section 5.2 discusses how theoretical considerations and third-party objectives influenced the decision to employ two categories of research agenda. Section 5.3 introduces issues that were taken into consideration during the formulation of the hypotheses while 5.4 looks at the research techniques used to operationalize the constituent concepts. Section 5.5 looks at the population of and section 5.6 presents the different categories of research design. Section 5.7 discusses some of the issues informing the analysis of the recorded data, and finally section 5.8 discusses the benefits of replication.

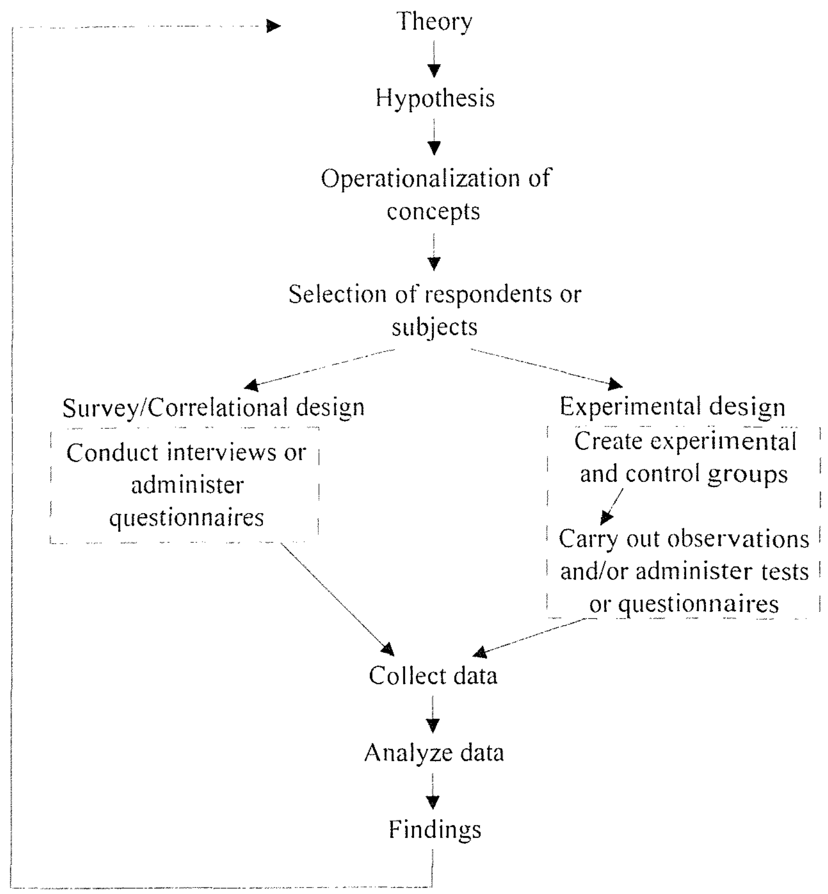


Figure 14 The research process (Bryman and Cramer 1997)

5.2 The Theoretical Domain

Argyrous (1997) defines pre-specified research as being informed by a clearly defined theoretical framework or by directives determined by other people or bodies that stipulate the key concepts to be investigated. Conversely, curiosity driven research is described as an investigation 'purely on the basis of a hunch: a loosely conceived feeling that something useful or important might be revealed by looking at a particular variable'. Analysis of the research objectives introduced in chapter 1, and consideration of the literature from the education (chapter 2) and computer science (chapter 3) domains, highlighted the need to embrace both pre-specified and curiosity driven research during the different phases of this work.

5.3 Identifying the Hypotheses

Research hypotheses were identified during every stage of the research to satisfy the analytical aspects of the research agendas: investigations that deal with associations or cause and effect relationships (see section 5.6). For each research hypothesis, the constituent abstract concepts were conceptualized to introduce working definitions that specified their qualities relative to this research. Relationships were forged between these concepts to introduce the sub hypotheses. Argyrous (1997) believes that hypotheses should be well-defined statements or claims about the characteristics of a population, with each exhibiting the following qualities:

- It must be clearly capable of being accepted and rejected.
- It must be amenable to quantification.
- It must state the population(s) for which an inference is made.

Consideration was given to ensure that the hypotheses derived for evaluation within this study satisfied these conditions. The operational statement (alternative hypothesis H_1) and null hypothesis (H_0) of each sub hypothesis was put forward for direct enquiry (Siegel 1956).

Hypotheses were not generated to satisfy the descriptive (see section 5.6) investigations that make up the latter stages of this research. Here statements about the research objectives were considered sufficient in describing the issues under investigation.

5.4 Operationalization of Concepts

Once the constituent concepts of the hypotheses had been identified they were operationalized to specify the rules, procedures, and operations (Argyrous 1997, Bryman and Cramer 1997) used to observe and measure them in the population. This process defines the research techniques to be employed for data generation and collection. Oppenheim (1998) defines research techniques as the methods 'concerned with measurement, quantification and instrument building and with making sure that our instruments are appropriate, valid and reliable'. The measurement of concepts in this research was principally achieved through self and group-administered questionnaires (Oppenheim 1998), although observation and informal interviewing techniques were employed at various phases.

5.4.1 Questionnaires

Self-administered questionnaires are usually presented to respondents by an interviewer or by someone in an official position. Once the purpose of the enquiring is explained, respondents are left alone to complete the questionnaires, which are picked up later. The group-administered questionnaire is given to groups of respondents assembled together (Oppenheim 1998) and is usually controlled by two or more persons. This method of administration is usually employed when an added degree of control is required, e.g. to ensure that all respondents have the same amount of time. This research employed both these questionnaire administration techniques to ensure a high response rate, accurate sampling and a minimum of interviewer bias, while allowing for appropriate administrative assistance and a degree of personal contact.

5.4.1.1 Questionnaire Planning

Oppenheim (1998) suggests that one of the biggest problems with questionnaire planning is finding ways to gain respondents' co-operation and of motivating them to respond to questions, and recommends the consideration of numerous factors that increase response rates when designing questionnaires. The following factors were thought relevant to this research and taken into consideration:

Anonymity - respondents were given anonymity during all evaluation phases. For the first evaluation, where data from more than one measurement instrument needed to be matched for individual respondents, code numbers were used to promote anonymity.

Appearance - no clear conclusions have emerged from numerous experiments on the general layout, type face, colour and quality of paper of questionnaires (Oppenheim 1998).

As a consequence, a conservative appearance was adopted to achieve the best results for each of the questionnaires.

Confidentiality - steps were taken to ensure that no information would be published that could identify respondents. The statement, 'Your answers will be treated confidentially', was used on each of the questionnaires to allay any apprehensions about the confidentiality of the information provided.

Incentives - a future cash incentive was used for the final phase of the research in an attempt to improve the poor response rate experienced in the second evaluation phase (see chapter 11).

Length - investigations into the length and time required to complete a questionnaire suggest that these variables depend on the topic and its degree of interest. Long and complex instruments are often completed successfully if the topic is of interest to the respondents or they believe that their responses will have a direct influence. Given that the issues being addressed by the questionnaires in this research were directly relevant to the quality of education for each respondent, it was hoped that the length of the questionnaires would not be a contributing factor to the response rate.

It has often been shown that respondents give some kind of answer to most questions even if they are ill informed and know it, and that they are likely to express an opinion on matters they have given little thought to or which they do not understand (Moser and Kalton 1979). With this in mind, all questions were authored with careful consideration about the knowledge the respondents were likely to possess.

5.4.1.2 Measuring Factual Data

Aside from the free-response (open) classification questions that were asked to determined sample demographics, given the problems associated with coding qualitative data (section 5.4.3 below), pre-coded (closed) questions were used to collect factual data (Moser and Kalton 1979, Oppenheim 1998). Here, the adjective 'factual' refers to the type of data being sought and does not refer to the accuracy with which responses are given: so a measure described as factual does not imply that the responses are necessarily accurate. For each pre-coded measure a limited but well established list of mutually exclusive alternatives was provided. A problem with pre-coded questions is that respondents who have never considered the concept being addressed may seize upon a lead given by the list of alternative answers (Moser and Kalton 1979). However, recognising the possibility that the list would not be exhaustive, and acknowledging the possibility of responses being forced into categories where they do not belong, opportunity was given for respondents to record answers not covered by the pre-specified alternatives.

5.4.1.3 Measuring Opinions and Attitudes

Pre-coded techniques were used to collect non-factual data. To avoid the misclassification problem associated with pre-coded questions (see section 5.4.1.2), space can be provided to allow respondents to make finer attitudinal distinctions (Moser and Kalton 1979), the disadvantage however is that this technique inherits the difficulties associated with coding qualitative data (section 5.4.3 below). For this reason a decision was made to offer a finite list of pre-coded alternatives with no room for qualification for each non-factual measure. Free-response questions were used throughout the questionnaires to give respondents the opportunity to express unencumbered what they considered important.

The problems associated with authoring measures of non-factual questions are well documented (Argyrous 1997, Bryman and Cramer 1997, Moser and Kalton 1979, Oppenheim 1998). In the case of factual questions, as long as it is made clear what information is sought, it is assumed that the respondent either does or does not know the answer. Attitudes and opinions are more complex and multi-faceted than issues of fact and are more susceptible to changes in wording, emphasis, context, events uppermost in the respondents' mind, and other biases. To this end, although some attitudinal concepts can be measured with a single indicator, it is often necessary to develop several indicators to measure the many facets of an attitude. This technique, referred to as a multiple-item measure, entails asking a respondent their position on a scale for a number of indicators that measure facets of the same concept. A score is determined by taking the mean of the aggregated indicators, and represents a respondent's position in relation to the totality of the complex attitude. This technique has the advantage of taking missing data into consideration by looking at the overall score for a multiple-item measure. For example, if a respondent strongly agrees (coded 5) with each of 4 indicators of a multi-item measure, the overall score would be 20 (5 * 4); however, if a respondent strongly agrees with all items but for some reason omits to answer one of the indicators, the overall score would be 15 (5 * 3), and would not reflect the true attitude. Taking the mean score of the non-missing values standardises the index so the overall score would be 5 for the first (20 / 4 = 5) and second (15 / 3 = 5) respondent. In this research, for any given respondent, if 10 per cent or more of the scores for a multiple-item measure were missing, the overall score for that measure was recorded as missing. Multiple-item measures have the added advantage of minimising the risk of a respondent being incorrectly classified, as would be the case if a single indicator is used to measure a concept and is misunderstood. If more than one indicator is used, as with multiple-item measures, a misunderstood statement is diluted by those that are interpreted correctly.

To measure the level of agreement or disagreement of attitudes, this research employed Likert (Moser and Kalton 1979, Bryman and Cramer 1997, Oppenheim 1998) scales consisting of five categories: 'strongly agree', 'agree', 'neutral', 'disagree', and 'strongly disagree'. A score of 1, 2, 3, 4 and 5 was assigned to each category respectively. Recognising the potential bias of people who agree or disagree to everything (Bryman and Cramer 1997, Moser and Kalton 1979), the wording of some indicators was reversed to change the direction of the measure. Consequently, a score for a negatively worded item was reversed to make the total score of a multiple-item measure meaningful.

5.4.2 Examinations

Multiple choice question (MCQ) (Laurillard 1993, Schank and Cleary 1995) and short answer test techniques were employed during the first evaluation phase. All examinations were administered under strict examination conditions. Negative marking principles were applied to MCQ scripts with one mark given for a correct answer and one subtracted for a wrong answer. For the short answer tests one point was awarded for a correct answer and negative marking was not applied.

For both methods of assessment, the marks were uncorrected, i.e. closed marking did not apply and the marks were not normalised in any way. In addition the examinations were blinded (Altman 1991, Wiederhold and Perreault 1990) at marking, giving no way of knowing if a particular script was from a student in the experimental or control group.

5.4.3 Observations

Observation techniques are used when there is a requirement to record information directly, rather than through reports of others (Moser and Kalton 1979). This method of enquiry is advantageous over information gathering by questioning when there is a concern that respondents are unable to offer exact information. In the context of this research, there was a concern that respondents could provide false information as a consequence of misunderstanding a question, or through thinking that a particular answer is expected. In addition observation techniques were used with a view of formulating a situated, holistic and well-rounded account (Barry 1995, Bryman 1998, Squires 1997) of the concepts under investigation. The technique employed during this research is categorised (Bryman 1998, Moser and Kalton 1979) as participant observation since observations were made by interpreting events of a closed group whilst engaging in conversations with the subjects. Difficulties with quantifying data collected through observation techniques begin with the bias introduced by how the observer perceives the information and continues through to the interpretation or coding of data for quantitative analyses (Bryman 1998, Mason 1994, Moser and Kalton 1979, Oppenheim 1998, Ritchie and Spencer 1994). Acknowledging this, no attempt was made to code and quantify the observations: synopsis of observations and dialogue are presented only. Observation techniques were not employed instead of quantitative techniques, they were used to augment them.

Where photography was used during this research, the ethical issue of respondents' fundamental right to privacy and to participate in the research was recognised at all times (Oppenheim 1998). To this end, respondents were notified when photography was to be used and were given prior opportunity to decline participation.

5.4.4 Focus Groups

Exploratory group interviewing techniques (Oppenheim 1998) - also known as informal

(Moser and Kalton 1979) or unstructured (Bryman 1998) - were used to inform the execution of a focus group. In contrast to standardised (formal or structured) interviewing where the questions, their sequence and wording are worked out beforehand for the specific purpose of gathering facts and statistics, the informal group discussions were encouraged to be free-flowing with little directive influence from the interviewer. The job of the interviewer was to collect ideas and not data. This was considered to be the best approach to promote a lively discussion among the respondents in the hope that this would encourage new ideas in each other (Grant 1997). Problems with coding data collected from unstructured interviews are akin to those experienced with other qualitative data collection techniques (see section 5.4.3). Consequently, as with the data collected from participant observations, no attempt was made to code and quantify the discussions of the informal focus group.

5.4.5 Pilot Work

Pilot work was carried out to refine and revise each of the questionnaires used for this evaluation. Consideration was given to the ideal of piloting on a sample representative of the respondents in the main enquiry (Oppenheim 1998). However, problems with finding a motivated sample together with time and resource constraints made this an unrealistic proposition. Each questionnaire was therefore piloted on colleagues who subsequently advised on problems with the general layout, notes on completion and wording of indicators.

5.4.6 Reliability and Validity of Measures

Reliability refers to the consistency of a measure, the probability of obtaining the same results if the measure were to be duplicated under the same conditions. Validity is concerned with whether a measure really measures what it is supposed to (Bryman and Cramer 1997, Kitchenham and Pickard 1998, Moser and Kalton 1979, Oppenheim 1998). Reliability and Validity are interconnected to the extent that a measure can be highly reliable yet invalid; but if a measure is valid then it must also be reliable (Oppenheim 1998).

5.4.6.1 Reliability

Test-retest (Bryman and Cramer 1997, Moser and Kalton 1979) is one method often employed to test the reliability of a measure. The procedure involves administering a measure to a sample, then re-administering the measure to the same sample after a period of time. A strong correspondence between the responses indicates an externally reliable measure. However, when testing external reliability in this way consideration needs to be given to the potential of bias. If the test and retest are administered too close in time subjects may recall earlier answers, leading to an artificial consistency. Alternatively results may be biased as a result of respondents having too long to think about the subject domain; they may make less effort to give accurate answers; or intervening events may change their attitudes on the subject. The problem is to choose an interval long enough to deal with memory effect whilst short enough to minimise the risk of events changing opinions and attitudes on the concepts under investigation. Given the problems associated with measuring reliability in this way, techniques that use internal checks are often preferred (Oppenheim 1998). For factual data reliability is often realised by introducing variations of the same question: in spite of using different techniques a respondent is expected to be consistent. Inconsistencies with related items would therefore highlight problems with wording or other sources of error. Since attitudinal questions are more sensitive to changes in question wording it is not appropriate to ask the same question in a different form. Instead multiple-item measures are employed to cancel out any bias introduced by ill-formed questions. The internal reliability of a multiple-item measure (see section 5.4.1.3) assesses the extent to which the constituent items are internally consistent. Estimation of internal reliability in this research was performed using Cronbach's alpha coefficient. This statistical test calculates the average of all possible split-half reliability coefficients for a set of items. Generating a value between 0 and 1, internal reliability is stronger the nearer the coefficient is to 1. Reliability is generally accepted with a coefficient of 0.8 or more (Bryman and Cramer 1997).

5.4.6.2 Validity

Determining the validity of a measure is best achieved if a criterion can be obtained. Oppenheim (1998) defines a criterion as 'an independent measure of the same variable' to which the results of a test or questionnaire can be compared. This approach is referred to as concurrent validity (Moser and Kalton 1979) and is concerned with assessing how well one measure converges or harmonises with another (Bryman and Cramer 1997). The problem with establishing this validity is that external sources of information are often unavailable. For attitudinal measures the use of samples with known attitude characteristics is proposed, but Oppenheim (1998) again acknowledges the difficulty in obtaining accurate samples. Construct validity is an approach based on theoretical considerations (Bryman and Cramer 1997, Moser and Kalton 1979): relationships are assumed between measures and observations are made to see if the association is as expected. Caution is advised here (Bryman and Cramer 1997, Moser and Kalton 1979) given the possibility of an erroneous postulation about the relationship, as the analysis of the association is as much a test of the theory as of the measure's validity. Bryman and Cramer (1997) maintain that a measure, at the very least, should be shown to have face validity: the constituent items of a measure reflect the content of the concept in question.

5.5 The Population of Interest

Generalising an inference from a sample to a population relies on the assumption that the studies' participants are representative of the population under study (Altman 1991, Argyrous 1997, Bryman and Cramer 1997, Moser and Kalton 1979). For each of the evaluations undertaken for this research, the population of interest was the undergraduate medical student body at the University of Southampton. The representative samples used were first year medical student populations. Assuming that the selection criteria of medical students is consistent from one year to the next gave confidence that the operation of random variation or selection bias was adequately controlled. Confidence can therefore be given when generalising the inferences made about the sample to the wider population of all undergraduate medical students. However, caution is advised when generalising the inferences to a wider population.

5.6 Research Design Categories

Altman (1991) offers a taxonomy of the categories of most research designs (see Figure 15).

Crudely classified as experimental and observational, this first classification is concerned with the nature of the study, while the various combinations of sub-classifications describe the way in which the data are collected. In observational research the variables of interest are not manipulated. Conversely, experimental research is concerned with deliberately manipulating events in order to observe the effects of an intervention. The sub-classifications prospective and retrospective reflect when data are collected. Prospective studies are those that are concerned with collecting data forwards in time from the start of the study while retrospective studies are concerned with the analysis of data from past events. The further sub-classifications longitudinal and cross-sectional refer to the number of observations. Longitudinal studies investigate changes over time, usually in relation to an intervention, and require observations to be taken on more than one occasion. Cross-sectional studies are those in which subjects are observed only once.

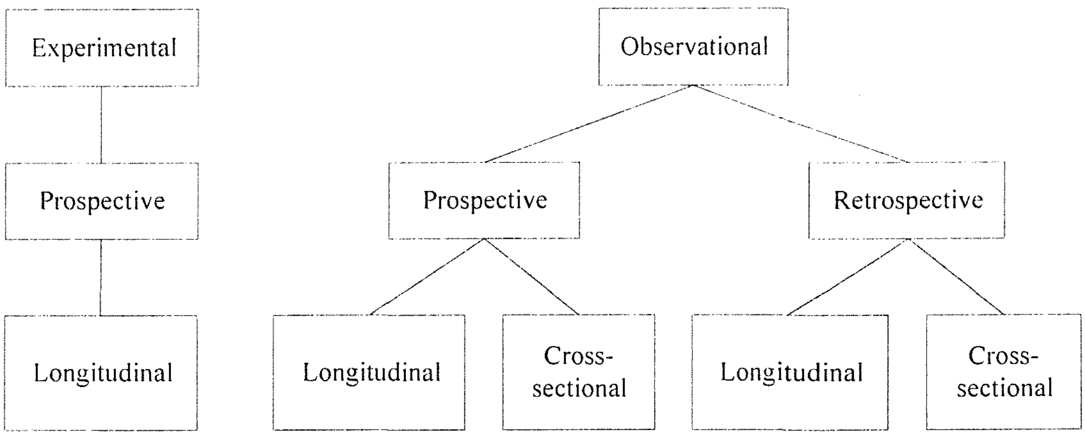


Figure 15 Types of research design (Altman 1991)

Oppenheim (1998) goes further in categorising research design as being either 'Descriptive' or 'Analytical'. Descriptive designs serve the purpose to count, to determine how often events occur, and are not designed to explain anything or to show causal relationships; while analytical designs deal with associations or cause and effect relationships. Relative to Altman's taxonomy, experimental designs are analytical, while observational (also referred to as survey/correlational, Figure 14) can be either analytical or descriptive.

A hybrid research design employing experimental and various combinations of observational strategies was required to satisfy both the pre-specified and curiosity driven facets of this research.

5.7 Analysing Data

This section discusses the factors taken into consideration during the statistical analyses phases.

5.7.1 Parametric and Non-Parametric Tests

Bryman and Cramer 1997 highlight the on-going debate on when it is appropriate to apply parametric or non-parametric statistical tests with some analysts arguing that the use of parametric tests should only be permitted when the data satisfies the following three conditions: the underlying scale of measurement is of equal interval or ratio scaling; the distribution of the population scores is normal; and the variance of the variables are equal. When choosing a statistical test, consideration was given to the level of measurement of the constituent variables. Where these variables were measured at different levels, the choice of a statistical test depended on the lowest level of measurement (Argyrous 1997). Further consideration is given to the circumstances requiring the use of different statistical tests in the sections of this thesis that address statistical analyses.

5.7.2 Chosen Confidence Level

When generalising the descriptive statistics from the sample to the population of interest, random variation was considered at a confidence level of 95 per cent. The level of significance or alpha level (α) for the inferential statistics was therefore chosen to be:

$$(\alpha) = 0.05$$

Consequently, it was accepted that the probability of the operation of random variation generating a biased result is 5 times in every 100. This confidence level was chosen as the appropriate compromise which optimises the balance between making a type I or type II error (Argyrous 1997, Bryman and Cramer 1997, Siegel 1956).

5.7.3 Establishing Causality

Testing hypotheses to establish causality (Bryman 1998, Bryman and Cramer 1997) was the primary method of statistical evaluation employed throughout this research. It is recognised that if findings infer that a given cause influences an effect, it cannot be said that the cause is the only factor influencing that effect, and that it can only be inferred that variation in the effect is related to variation in the cause. For each measure of association presented in this thesis, the predicted relationship between the variables is stated, identifying the independent (cause) and dependent (effect) variables, and the following procedure followed (Bryman and Cramer 1997): first, a relationship was identified between the constituent variables by demonstrating that the distribution of one variable was homogeneous with the distribution of the other and that the pattern was not randomly distributed; second, the relationship was demonstrated to be non-spurious (a spurious relationship is evident when the variations of both variables are affected by a common variable); and third, it could be proven that the cause preceded effect by establishing the time-order to the variables.

5.8 Replication

Fenton and Pfleeger (1997) define replication as 'repeating an experiment under identical conditions, rather than repeating measurements on the same experimental unit', and highlight the benefits as giving confidence in the results and enabling the mean effect of any experimental factor to be estimated. The benefits of replicating measurements in experimental settings are discussed by Altman (1991) who discusses the desirability of taking 'more than one reading in each combination of experimental condition as it gives greater precision for estimating the effects of interest'.

Decisions about replicating the evaluations carried out in this research were influenced by the time and resources available. However, the ability to replicate was a driving factor in ensuring that this research could be replicated by third parties from the information provided in this volume.

5.9 Summary

This chapter discussed issues relevant to the design and execution of the evaluations conducted as part of this research. Following the research process proposed by Bryman and Cramer (1997), section 5.2 discussed how theoretical considerations and third-party objectives lead to the employment of two categories of research agenda: pre-specified and curiosity. Section 5.3 touched on the issues taken into consideration during the formulation of the hypotheses while 5.4 looked at the research techniques used to operationalize the constituent concepts under investigation. Turning to the execution of the evaluations: section 5.5 expressed caution about generalising the research findings to wider populations; section 5.6 discussed the different categories of research design; and section 5.7 discussed some of the issues informing the analysis of the recorded data. Finally, section 5.8 discussed how time and resource constraints influenced decisions to replicate the work presented in the thesis. Informed by the principles put forward in this chapter, the next chapter presents the research agenda, conceptualization and operationalization processes that underpin the first evaluation of SCALPEL.

Chapter 6: The Research Agenda, Conceptualization and Operationalization

6.1 Introduction

The previous chapter introduced the evaluation methods and techniques that inform the empirical evaluations presented in this thesis. This chapter starts the evaluation presentation by introducing the research agenda, conceptualization and operationalization processes underpinning the primary evaluation of SCALPEL. Section 6.2 discusses the pre-specified and curiosity driven research agendas, introducing the constituent concepts of the primary research hypothesis and student-specific characteristics of interest. In section 6.3 these concepts are conceptualized to offer working definitions, and then operationalized to specify their working measures. Finally, relationships are forged in section 6.4 to specify conceptual definitions of research sub hypotheses for the pre-specified and curiosity driven research.

6.2 Defining the Evaluation Framework

Conventionally, empirical evaluations are categorised as being either formative or summative (Boyle 1997, Jacobs 1998, Laurillard 1993). Formative evaluation is employed during system development with the intention of using the information to influence improvements, while summative evaluation determines the system's overall success. Both phases evaluate the effectiveness and usability of a system, factors that Boyle (1997) believes to be 'crucial if the system is to operate as an effective learning environment'. He interprets the effectiveness of a system as depending on 'whether it achieves its learning objectives', and usability as the 'ease of learning and ease of use'.

Both formative and summative phases were employed for this evaluation. The pre-specified agenda of this research was chiefly summative. By addressing the effectiveness and usability of SCALPEL, the primary goal was to satisfy the medical faculty that replacing the traditional pathology practicals with courseware would not be detrimental to student learning. The curiosity driven research served the purpose of formative evaluation with data collected being used to influence improvements to the courseware environment.

6.2.1 Pre-Specified Research

The pre-specified research addressed the domain of computers in education with a focus on replacing traditional teaching with pedagogically valuable courseware. Within the context of this research: traditional teaching was conceptually defined as undergraduate pathology practicals in a traditional laboratory environment; pedagogic value was defined as students' education attainment and acceptance of a teaching environment; and courseware was taken to be the SCALPEL software. With these definitions, the primary research hypothesis was conceptually defined as:

> courseware could replace pathology laboratory practicals with improved undergraduate medical students' education attainment and acceptance of learning.

6.2.2 Curiosity Driven Research

Identifying student demographics and past experiences has long been recognised as an important exercise in the attempt to establish why some students perform better in education than others (Laurillard 1993, Ramsden 1992). Fraser (1997) expresses concern that 'failure to ascertain and take into consideration information about student specific characteristics can result in minimal student use of a product'. Past empirical research has shown that the use of courseware in medical education is not beneficial to all students (Jelovsek and Adebonojo 1993). Consequently a second research hypothesis was defined as:

> replacing pathology laboratory practicals with SCALPEL would not be pedagogically valuable to all students.

To address this hypothesis a decision was made to investigate what 'students bring to learning' (Laurillard 1993). Relevant student-specific characteristics were identified and categorised according to the following four broad types of social survey subject matter (Moser and Kalton 1979): demographics characteristics, socio-economic factors, behaviour and activities, and opinions and attitudes. The various dimensions of these concepts are introduced in the next section.

The primary objective of this research agenda was to identify relationships between relevant student-specific characteristics and the constituent dimensions of pedagogic value: education attainment and acceptance of SCALPEL. This formative evaluation was driven by the motivation to identify individuals or groups of students that faired better or worse with the courseware environment, with the intention of using the information to influence improvements to the courseware, and implement targeted preparation and support. A secondary consideration was the opportunity to investigate issues pertaining to the makeup of students and the use of computers in education. These issues are explored in sections 6.3.2 - 6.3.5, where the various dimensions of the four student-specific characteristic categories introduced above are discussed.

6.3 Conceptualization and Operationalization

Section 6.2 introduced abstract definitions for the concepts used to measure the pedagogic value of SCALPEL and student-specific characteristics. In this section pedagogic value and each of the four broad categories of student-specific characteristics are conceptualized to introduce working definitions that specify their qualities relevant to this research. These working definitions are then operationalized to specify the rules, procedures, and operations (Argyrous 1997) used to observe and measure the concepts in the population. As the education attainment dimension of Pedagogic Value cannot logically be pigeonholed into one of the four student-specific characteristic categories, this is addressed first. The second dimension of pedagogic value, Acceptance of SCALPEL, is conceptualized as an Opinion and Attitudinal trait.

6.3.1 Education Attainment

For the purposes of this research the conceptual definition of education attainment was:

> "the ability to recall and apply subject domain knowledge"

This concept was operationalized through the use of two examination methods: MCQ and short answer tests (see section 5.4.2).

The MCQ tests consisted of twenty questions. Negative marking principles applied with one mark given for a correct answer and one subtracted for a wrong answer. The short answer tests consisted of ten questions with one point awarded for a correct answer. Negative marking was not applied to the short answer tests.

6.3.2 Demographics Characteristics

The broad concept of demographics has many dimensions that reflect different underlying aspects. For the purpose of this study the following dimensions were determined to be of interest: Age group, gender and demographic area of post-primary education.

A different emphasis is being placed on the importance of demographic data within computer science. For example, within the hypertext community demographic data are used to identify patterns of interactions to inform adaptive systems for on-the-fly dynamic linking of domain material for individual users (Yan et al. 1996). With early research reporting on demographic differences within the population of computer users in favour of younger males, evidence of a more heterogeneous user base is welcomed in all quarters. Richmond (1996) investigated the impact of varying website functionality on the purchasing behaviour of online shoppers. Her sample frame consisted of regular visitors to a store chosen for their computer literacy (determined by having a personal computer in the home and current use of a mouse). Of the 75 subjects, 30 were men and 45 were women with a combined average age of 42. Can such a noticeable shift in the gender and age of users be generalised to differing populations, or is this empirical evidence, if not flawed by research methodology, specific to shopping? Are the concerns about gender and age difference of computer uses being allayed prematurely?

Voiskounsky (1998) recently looked at the characteristics of Russian Relcom users in an attempt to identify the sociological and psychological processes that influence network use. He found that more than half of the 305 respondents to his survey were aged between 21 and 25, and more than three-quarters between 21 and 35. Data were not given on the gender composition of the sample. Voiskounsky's study offers evidence in support of the hypothesis that computer users are predominately young. With this observation contradicting the age make-up of Richmond's sample, it is apparent that there is a continuing need for research into the age of computer users. The concept of age was therefore considered important to this study in an attempt to determine if such concerns would need to be addressed within the SCALPEL environment.

Concerns of gender inequalities in the computer classroom have long been observed in favour of a prominent masculine presence (Culley 1991). Recent case studies indicate that gender differences of computer-mediated communication (CMC) users show no sign of disappearing. Witmer and Katzman (1998) investigated gender differences in the use of graphic accents (the use of symbols e.g. ☺), challenging and hostile or abusive language during CMC. They found that out of their sample of 2599 publicly posted newsgroup messages, 83.6% originated from males. Smith et al. (1998) explored further the nature of offensive conduct and reproaches on Usenet and found 78% of their sample (n = 109) to be male. Examining the characteristics of users who engage in risky CMC on the Internet, Witmer (1998) also observed a predominately male (78.8%, n = 52) population. While Jones (1998) examined how members of a Usenet newsgroup valued and used news sources. In keeping with what seems to be a prevalent pattern, of the 128 respondents to his questionnaire, 85.5% were male. These findings are particularly pertinent to education technology, and therefore this research, given an increasing interest in CMC within this arena. As a consequence gender was considered an important demographic to address during this study.

Hawkridge (1991) discusses the need for computers in the schools of developing countries, and compares the rationales behind their introduction to developing and industrial countries. Magnus (1996) discusses the political, economic and social forces in sub-Saharan Africa that have affected formal and distance learning initiatives. The concept of students' demographic area of post-primary education was therefore considered necessary in an attempt to highlight differing attitudes towards the use of courseware. The data were also considered necessary in highlighting any culturally determined usability problems

(Bourges-Waldegg and Scrivener 1998) that may bias the successful use of SCALPEL for some students.

These demographics were operationalized through indicators on a questionnaire administered at the beginning of the study.

Age Group

Age group was measured at the ordinal level by asking students to tick one of the following categories: under 20, 20 – 25, 26 – 35, over 35.

Gender

Gender was measured at the dichotomous level by asking students to tick the appropriate box, male or female.

Demographic area of post-primary education

Demographic area of post-primary education was measured at the nominal level as the ordering of the categories was not applicable. A free text approach was used to ask students to indicate the country in which they went to school after the age of eleven. When the data was entered into SPSS, the responses were categorised as follows: Great Britain, Western Europe, Eastern Europe, Middle East, Africa, Asia, North America, Latin America, Australasia. Where a student indicated attending schools in more than one country, i.e. Sudan and England, the first of the indicated countries was taken to categorise the student for this variable.

6.3.3 Socio-economic Characteristics

The socio-economic characteristics of interest to this study were previous higher education and access to computers.

The concept of previous higher education was considered relevant to this research for the purpose of determining if higher education students differed in respect to the adoption or rejection of new technology in place of traditional teaching from those who had not previously undertaken higher education. Kirkwood (1996) expresses caution about making assumptions that learners have access to suitable equipment through highlighting the availability problem of 'media technologies suitable for study'. With the potential for SCALPEL to be made available as a distance learning medium, descriptive data on this concept was considered of interest to this research. In addition, the data served to determine if access to computers had a systematic effect on the success of SCALPEL.

These characteristics were operationalized through indicators on a questionnaire administered at the beginning of the study.

Previous Higher Education

A free text approach was used to determine students' previous higher education. When the data were entered into SPSS, the free text responses were coded onto an ordinal scale as follows: 0 = None, 1 = Diploma, 2 = Degree, 3 = Masters Degree, 4 = PhD. The subject studied was not of interest to this research, just the level of education. Given that the ordering of the categories was considered to be important, this concept was measured at the ordinal level. Students who indicated undertaking higher education were also asked to write down the country in which the education took place. When this data was entered into SPSS, it was interpreted into one of the following nominal categories: Great Britain, Western Europe, Eastern Europe, Middle East, Africa, Asia, North America, Latin America, or Australasia.

Access to Computers

An ordinal scale was specified to categorise students' access to computers at home and at university as follows:

1 = Students who have access to their own computer both at home and at university
2 = Students who have access to their own computer only at university
3 = Students who have access to their own computer only at home
4 = Students who only have access to university computers

A lower score on the scale indicated the most access to a computer. Five statements were used to correctly locate students on this ordinal scale. Students were asked to indicate 'yes' or 'no' to the following statements:

S1. I own a portable (laptop) computer

S2. I have a computer in my room at university

S3. I have a desktop computer at home

S4. I have access to a computer a home

S5. I do not have access to a computer other than the university workstations.

With S5 being mutually exclusive with the rest of the statements, all responses were screened to ensure that contradictions did not occur. Students were then located on the ordinal scale by applying the following rules:

'yes' to question S1 = 1

'yes' to question S2 only = 2

'yes' to questions S3 and/or S4 only = 3

'yes' to question S5 only = 4

All other combinations of 'yes' responses were coded on the ordinal scale as 1. e.g. a 'yes' response for statements S2 and S4 would indicate access to computers both at home and at university.

6.3.4 Behavioural Characteristics

Two underlying dimensions of student behaviour were determined to be relevant to this study: previous computer use and learning styles.

With the implementation of SCALPEL forcing a shift from the traditional laboratory to a courseware environment, students' previous computer use was considered a potentially important influencing factor on the success of SCALPEL. It is widely accepted that knowledge is not acquired, it is constructed (Mayes 1993), with students learning best when control of learning is negotiated between tutor and student. By taking responsibility, rather than being passive learners, students are in a position to explore the subject area according to their interests. Commentators believe this learning approach is augmented well by computers (Schank 1994, Schank and Cleary 1995). As a consequence, learning styles were considered a valuable area of assessment. To assess the impact of exposure to SCALPEL on learning styles, this measure was further employed as part of a 'before-after with control group' experiment.

Previous Computer Use

It was determined that the following categories would capture students' past exposure to the many dimensions of computer use: games, formal learning, Internet, word processing and work. For each of these categories, students were asked to indicate the frequency of computer use by ticking one of the following ordinal categories:

1 = Frequent (more than 1 x per week)
2 = Sometimes (less than 1 x per week)
3 = Never before.

This concept was operationalized through the above indicators on a questionnaire administered at the beginning of the study.

Learning Styles

To empirically test learning styles relative to pathology practicals, a Likert scale was employed to crudely measure the indicated strength of passive or exploratory learners. The scale indicates the strength of attitudes ranging from a strong appreciation of learning by exploration (constructing knowledge, taking responsibility) to a strong appreciation of passive learning (knowledge acquisition through rote memorisation of information primarily presented by course material). A low score on the scale indicates a strong appreciation of learning by exploration, while the higher scores represent a strong attitude towards passive learning.

This scale was operationalized by employing a multiple-item measure (see section 5.4.1.3) consisting of fourteen indicators on a questionnaire administered before and after exposure to SCALPEL. Recoding of negatively directed indicators (denoted by a minus sign in brackets after the indicator) was carried out for indicators 1, 2, 6, 8 and 11.

Indicator 1: I prefer to use handouts instead of textbooks (-)
Indicator 2: We should be told what sections of textbooks to read (-)
Indicator 3: Textbooks are helpful aids to fill in areas of learning (+)

Indicator 4: We should be left to find the information about subjects from textbooks (+)

Indicator 5: We should be allowed to explore subjects according to our interests (+)

Indicator 6: Lectures should tell us all we need to know (-)

Indicator 7: Lectures are unnecessary (+)

Indicator 8: Lectures should cover all I need to know about a subject (-)

Indicator 9: Lectures should provide me with a guide to additional information that I should know about a subject area (+)

Indicator 10: Lectures can be a helpful way of finding out about a subject area (+)

Indicator 11: Practicals should be demonstrations of what we need to know (-)

Indicator 12: Practicals should include some problem solving work (+)

Indicator 13: Practicals should require that we recognise things that we learnt in the course (+)

Indicator 14: Practicals should be about solving problems relating to areas that we have not covered (+)

6.3.5 Opinions and Attitudes

Students' opinions and attitudes determined to be of interest were acceptance of SCALPEL, attitudes towards the need for demonstrators and attitudes towards computers.

As a dimension of pedagogic value, acceptance of SCALPEL was considered important by offering an insight into attitudes towards SCALPEL and the traditional laboratory environments. Student attitudes towards computers were considered important given the shift from the traditional laboratory to a courseware environment. It was considered that if student-specific characteristics have an influence on students' attitudes towards computers, this could have a transitive effect on the success of SCALPEL. In addition, interested in establishing if exposure to SCALPEL would have an effect on these attitudes, the measure was employed as part of a 'before-after with control group' experiment. Having observed the key role of demonstrators in the traditional laboratory (see chapter 4), student attitudes towards the need for demonstrator assistance during exposure to SCALPEL were identified as a valuable area of assessment.

Attitudes towards SCALPEL and the traditional laboratory

Attitudes towards SCALPEL and the traditional laboratory were measured by using a multiple-item measure (see section 5.4.1.3) consisting of seventeen indicators on a questionnaire administered at the end of the study. Recoding of negatively directed indicators was carried out for indicators 4, 6, 7, 10, 13, 14 and 17. A low score on the scale represents a positive attitude towards SCALPEL, while higher scores represent attitudes in favour of the traditional laboratory.

Indicator 1: I prefer to use the computer application for pathology practicals (+)

Indicator 2: I learn more pathology using the computer application (+)

Indicator 3: I understand the teaching material better when it is presented by the computer application (+)

Indicator 4: I felt lost and confused using the computer application (-)

Indicator 5: I like the consistency of explanations provided by the computer application (+)

Indicator 6: I prefer laboratory teaching for pathology practicals (-)

Indicator 7: I understand the teaching material better when it is presented in the laboratory (-)

Indicator 8: I participate more in the practical using the computer application (+)

Indicator 9: I find the computer application a satisfying learning activity (+)

Indicator 10: I find more teaching material is provided in the laboratory (-)

Indicator 11: I do not feel that laboratory teaching is a satisfying learning activity (+)

Indicator 12: I like the environment the computer application provides (+)

Indicator 13: I participate more in the laboratory (-)

Indicator 14: I learn more pathology in the laboratory (-)

Indicator 15: I find more teaching material is provided by the computer application (+)

Indicator 16: Not enough teaching material is provided in the laboratory (+)

Indicator 17: I find practicals in the laboratory a satisfying learning activity (-)

Attitudes towards Computers

Six indicators were used to measure on a Likert scale (see section 5.4.1.3) the strength of student attitudes towards computers. These indicators were employed on questionnaires administered before and after exposure to SCALPEL to measure changes in attitudes. Recoding of negatively directed indicators was carried out for indicators 1, 2 and 6. A low score on the scale represents a positive attitude towards computers while higher scores represent more negative attitudes.

Indicator 1: Computers often make more problems than they solve (-)

Indicator 2: Computers can be useful, but I have never found them so (-)

Indicator 3: I sometimes find computers useful (+)

Indicator 4: I find computers useful and efficient tools (+)

Indicator 5: I could not function without computers (+)

Indicator 6: Computers make me uneasy (-)

The Need for Demonstrators

The need for demonstrators to be available during SCALPEL sessions was measured by a multiple-item measure (see section 5.4.1.3) consisting of four indicators on a questionnaire administered at the end of the study. Recoding of negatively directed indicators was carried out for indicator 1. A low score on the scale represents a stronger appreciation of demonstrator assistance than a higher score.

Indicator 1: I do not need the help of demonstrators when I use the computer application (-)

Indicator 2: I believe demonstrators should be available at set times in the future (+)

Indicator 3: I miss demonstrator input whilst using the computer application (+)

Indicator 4: I will always need the help of demonstrators, irrespective of the quality of knowledge given by the computer application (+)

6.4 Research Sub Hypotheses

This section introduces the research sub hypotheses put forward for enquiry to satisfy the agendas of the first empirical evaluation.

6.4.1 Pre-Specified Research - Assessing Pedagogic Value

Three sub hypotheses were formulated to address the pre-specified research agenda. Primarily concerned with satisfying the medical faculty of the pedagogic value of replacing the traditional pathology practicals with courseware, sub hypothesis 1 was authored to compare the education attainment of the two teaching environments.

Sub Hypothesis 1 - The education attainment of pathology on average differs between medical students who work through practicals in the traditional laboratory and those who use SCALPEL.

A further sub hypothesis was identified to assess the second dimension of pedagogic value, the attitudinal trait 'acceptance of SCALPEL'.

Sub Hypothesis 2 - Medical students find SCALPEL an acceptable pedagogic environment.

Finally, a relationship was forged between the two constituent concepts of pedagogic value: education attainment and acceptance of SCALPEL.

Sub Hypothesis 3 - Medical students' acceptance of SCALPEL influences education attainment.

6.4.2 Curiosity Driven Research - Assessing Student-specific Characteristics

Of the student-specific characteristics introduced in section 6.3, three were determined to be of particular interest to this study. Two were the attitudinal traits: The Need for Demonstrator Assistance, and Attitudes Towards Computers; the third was the behavioural characteristic, Learning Styles. The following sub hypotheses were shaped around pedagogic value and each of these primary concepts.

6.4.2.1 Pedagogic Value

Sub hypotheses 4 - 6 look at the relationships between pedagogic value and the three primary concepts introduced above, while sub hypotheses 7 - 9 address each of the student-specific categories identified in section 6.2. Note a sub hypothesis for the category of opinions and attitudes is not offered as the constituent concepts have already been addressed: acceptance of SCALPEL by sub hypothesis 3; the need for demonstrator assistance by sub hypothesis 4; and attitudes towards computers by sub hypothesis 5.

Sub Hypothesis 4 - Medical students' need for demonstrator assistance influences the pedagogic value of SCALPEL.

Sub Hypothesis 5 - Medical students' pre-SCALPEL attitudes towards computers influence the pedagogic value of SCALPEL.

Sub Hypothesis 6 - Medical students' pre-SCALPEL learning styles influence the pedagogic value of SCALPEL.

Sub Hypothesis 7 - Medical students' demographics influence the pedagogic value of SCALPEL.

Sub Hypothesis 8 - Medical students' social economic characteristics influence the pedagogic value of SCALPEL.

Sub Hypothesis 9 - Medical students' behavioural characteristics influences the pedagogic value of SCALPEL.

6.4.2.2 Demonstrator Assistance

Relationships were inferred between the attitudinal dimension 'the need for demonstrator assistance' and each of the student-specific characteristic categories.

Sub Hypothesis 10 - Medical students' demographics influence expressed need for demonstrator assistance.

Sub Hypothesis 11 - Medical students' socio-economic characteristics influence expressed need for demonstrator assistance.

Sub Hypothesis 12 - Medical students' behavioural characteristics influence expressed need for demonstrator assistance.

Sub Hypothesis 13 - Medical students' opinions and attitudes influence expressed need for demonstrator assistance.

6.4.2.3 Attitudes Towards Computers

Relationships were forged between the attitudinal dimension 'attitudes towards computers' and each of the student-specific characteristic categories. In addition, sub hypothesis 18 was conceptualized to assess the systematic effect of exposure to SCALPEL on students' attitudes towards computers.

Sub Hypothesis 14 - Medical students' demographics influence attitudes towards computers.

Sub Hypothesis 15 - Medical students' socio-economic characteristics influence attitudes towards computers.

Sub Hypothesis 16 - Medical students' behavioural characteristics influence attitudes towards computers.

Sub Hypothesis 17 - Medical students' opinions and attitudes influence attitudes towards computers.

Sub Hypothesis 18 - Exposure to SCALPEL changes medical students' attitudes towards computers.

6.4.2.4 Learning Styles

Relationships were inferred between the behavioural characteristic 'learning styles' and each of the student-specific characteristic categories. In addition, sub hypothesis 23 was conceptualized to assess the systematic effect of exposure to SCALPEL on students' learning styles.

Sub Hypothesis 19 - Medical students' demographics influence learning styles.

Sub Hypothesis 20 - Medical students' socio-economic characteristics influence learning styles.

Sub Hypothesis 21 - Medical students' behavioural characteristics influence learning styles.

Sub Hypothesis 22 - Medical students' opinions and attitudes influence learning styles.

Sub Hypothesis 23 - Exposure to SCALPEL changes medical students' attitudes towards learning styles.

6.5 Operational Statements

For clarity, the operational statement (alternative hypothesis H_1) and null hypothesis (H_0) for each research sub hypothesis is presented in the relevant chapters that address hypotheses testing.

6.6 Summary

This chapter presented the conceptualization and operationalization processes that underpinned the evaluation of SCALPEL. A discussion about the pre-specified research agenda introduced the conceptual definition of the primary research hypothesis. Addressing the curiosity driven research, student-specific characteristics were identified as being of interest to this study, and categorised into four broad types of social survey subject matter. Next, these concepts were conceptualized and operationalized to allow them to be observed and measured in the population. Finally, relationships were forged between these concepts to introduce the sub hypotheses. It is these sub hypotheses that are operationalized in chapters 7 and 8 and put forward for direct empirical testing. This begins in the next chapter with the first empirical study and the evaluation of the pedagogic value of SCALPEL.

Chapter 7: The First Empirical Study: Assessing the Pedagogic Value of SCALPEL

7.1 Introduction

This chapter presents the evaluation of the pedagogic value of SCALPEL. Chapter 5 went some way in introducing 'how' the empirical studies were conducted by discussing research methods and techniques. The previous chapter formally introduced 'what' was included in the first study and 'why'. This chapter presents the execution of the first study and the subsequent findings. Section 7.2 discusses the evaluation methodology specific to this phase of the research with an insight into the crossover design and problems encountered with random selection. Bryman and Cramer (1997) argue that 'once a concept has been operationally defined, in that a measure of it has been proposed, the ensuing measurement device should be both reliable and valid'. Section 7.3 addresses the reliability and validity of the measures used in this study, with a focus on the multiple-item measures operationalized in the previous chapter. Section 7.4 presents the demographic makeup of the sample population. Focused on addressing the pre-specified research agenda, the education attainment facet of pedagogic value is next assessed. This analysis begins with a look at the relationship between the MCQ and short answer tests (section 7.5.1) before moving on to discuss any statistical or practical differences between the mean examination results of students who had exposure to SCALPEL and those who worked in the traditional laboratory (section 7.5). Next, section 7.6 addresses acceptance of SCALPEL, the second dimension of pedagogic value, with a look at the distribution of attitudes towards SCALPEL and the traditional laboratory. Section 7.7 closes the pre-specified research agenda with a look at the relationship between the two constituent concepts of pedagogic value. Turning to the pre-specified research, sections 7.8 - 7.14 look at the relationships between the constituent concepts of pedagogic value and the student-specific characteristics identified in section 6.2.2 as being of interest. Moving away from the quantitative analysis, section 7.15 adds a qualitative dimension to this research by discussing the findings of participant observations while section 7.16 offers an insight into an informal focus group's discussion threads about the traditional laboratory and courseware environments. A discussion on all the statistical analyses presented in this chapter is offered in section 7.17. Section 7.18 concludes by highlighting the significant findings.

7.2 Method

A prospective longitudinal crossover experiment (see section 5.6) was designed to satisfy both the pre-specified and curiosity driven aspects of the research. A population of first year undergraduate medical students (N = 159) was split into two groups, 'A' and 'B'. Running for a total of six practicals (academic term 2), group 'A' used SCALPEL for the first three practicals, crossing over to the traditional laboratory for the final three, while group 'B' used the traditional laboratory for the first three, crossing over to use SCALPEL for the final three. This formal experiment is depicted graphically in Figure 16.

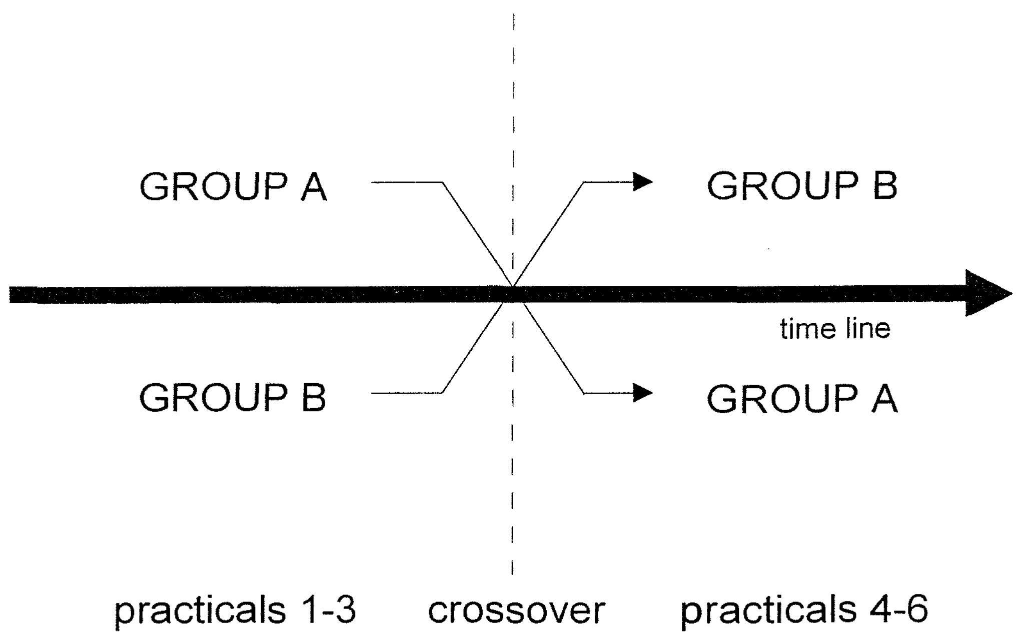


Figure 16 Crossover study design

To model the traditional laboratory environment as closely as possible (see section 4.2.5), and to ensure that students used the appropriate method of teaching, the SCALPEL application was only made available to students during the set practical times. Demonstrators were available during both the laboratory and SCALPEL sessions.

Altman (1991) highlights two consequences of using control groups to control for the random variation of extraneous variables:

1. Care is needed to make samples representative of the population.
2. In comparative studies care is needed in making groups similar with respect to known sources of variation.

Controlling for random variation by allocating students to the groups at random was not possible due to the logistical constraints of managing 159 students between a pathology laboratory, two computer suites and tutorials. Students were therefore assigned to group A or B according to their tutorial groups. Given that by virtue of being accepted as medical students all students in the population were considered to have achieved an equivalent education, it was determined that the potential of selection bias would be minimal. However, control techniques were used to control for extraneous variables both between-subject and within-subject. The latter technique controlled for extraneous variables by using students as their own controls (Altman 1991), therefore controlling for personal or background characteristics that may bias the between-subject analysis.

With the population of interest being undergraduate medical students, confidence can be given that the sample of first year undergraduate medical students used for the statistical analyses presented in this study is representative of the population of interest. However, caution is advised when generalising the inferences to a wider population.

7.2.1 Research Techniques

Students were required to sit an MCQ and short answer examination at crossover (after three practicals) and at the end of the study (after six practicals) (see section 5.4.2). A questionnaire was administered to students at the beginning of the study, after exposure to SCALPEL, and at the end of the study (see section 5.4.1). The first questionnaire was given to the students before the start of the first practical (in the traditional laboratory for group A and in the computer suite for group B). Students were given the second questionnaire to complete at the end of their last practical using SCALPEL (after practical 3 for group B and practical 6 for group A). The final questionnaire was given to students to complete before starting the MCQ and short answer examinations administered at the end of the study.

To promote anonymity, each student was assigned a number at the beginning of the study. Each questionnaire was stamped with a valid student number and a piece of paper with the corresponding student name was stapled to the front page. The questionnaires were then administered to the correct student at the appropriate time. Before returning a completed questionnaire students were instructed to remove the piece of paper with their name on it. The student numbers stamped on each of the questionnaires were then used to ensure that students were matched to the correct three questionnaires during the data processing stage. The questionnaires used during this study are presented in Appendix 9.

7.3 Reliability and Validity of Measures

This section addresses the reliability and validity of the measures used during this phase of the research (section 5.4.6).

7.3.1 Reliability

Cronbach's alpha was used to test the internal reliability (see section 5.4.6.1) of each multiple-item measure employed as part of this study. All other measures were determined to be reliable through establishing face validity (see section 7.3.2).

Learning Styles

Laurillard (1993) suggests that a student may adopt differing learning styles depending on the activity being undertaken. It is therefore recognised that the constituent items of the Learning Styles multiple-item measure (see section 6.3.4) are not mutually exclusive. Students may exhibit contradicting styles towards the underlying dimensions of this concept, i.e. a student who expresses a strong resolve for exploring the subject on one item, may indicate a preference for the acquisition of knowledge to be controlled by the course on another. The potential for contradictory learning styles has implications when testing the internal reliability (alpha coefficient) of this measure, which assumes that responses to the indicators representing the dimensions of the concept are related. As a consequence caution should be taken when interpreting the internal reliability of this multiple-item measure.

Cronbach's alpha was measured for students' pre (Alpha = 0.56, N of cases = 137, N of items = 14) and post (Alpha = 0.63, N of cases = 140, N of items = 14) SCALPEL learning styles. Both the coefficients were short of the 0.80 criterion (Bryman and Cramer 1997) suggesting that either: the measure is internally reliable with the low coefficients being a consequence of contradictory responses to the underlying facets of this concept; or the measure cannot be taken as being internally reliable. The former was taken to be the case given that this measure was determined to have face validity, that is, each indicator appeared to reflect a facet of the concept in question.

Attitudes towards SCALPEL

The alpha coefficient for the multiple-item measure that addressed students' attitudes towards SCALPEL (Alpha = 0.94, N of cases = 136, N of items = 17) suggested a very strong level of association between the constituent indicators, giving confidence in the internal reliability of the measure.

Attitudes towards Computers

Students' pre-SCALPEL attitudes towards computers were measured on a questionnaire administered before exposure to SCALPEL. With Alpha = 0.71 (N of cases = 94, N of items = 6) just below the recommended level of 0.80, this multiple-item measure was considered internally reliable. Post-SCALPEL attitudes towards computers were measured by using the same measure on a questionnaire administered after exposure to SCALPEL. The internal reliability alpha coefficient for this multiple-item measure (Alpha = 0.67, N of cases = 145, N of items = 6) indicated a moderate to high level of association between the constituent items. With the pre-SCALPEL reliability coefficient for this multiple-item measure indicating an acceptable level of internal reliability, confidence was given that the constituent indicators measured the same underlying concept.

The Need for Demonstrators

The alpha coefficient for the multiple-item measure (Alpha = 0.67, N of cases = 138, N of items = 4) authored to address students' need for demonstrators suggested a moderate to high level of association between the constituent items, giving confidence in the internal reliability of this measure.

7.3.2 Validity

All the measures used in this evaluation were determined to have face validity (see section 5.4.6.2). Concurrent validity could not be established given the difficulties associated with obtaining independent sources of information. For some measures construct validity was inferred by definition of their observed relationship with other measures.

7.4 Sample Frame Demographics

Introduced in section 7.2, a population of first year undergraduate medical students (N = 159) was used for this study. Demographic data were collected on the questionnaire administered to the population prior to exposure to SCALPEL. With two non-responses, the sample frame consisted of 157 students. Of these 82 (52.2%) were male and 75 (47.8%) female. The majority of the students (78.3%) were under the age of 20. 17.2% were aged between 20 and 25, with the remaining 4.5% aged between 26 and 35 (see Appendix 1, Table A1 - 1).

7.5 Education Attainment

This section looks for any statistical or practical differences between the mean examination results of students who had exposure to SCALPEL and those who worked in the traditional laboratory.

7.5.1 MCQ and Short Answer Tests

MCQ and short answer tests (see section 6.3.1) were employed to evaluate the education attainment facet of pedagogic value (see section 6.2.1). Both methods of assessment were administered to students at crossover (after three practicals) and at the end of the study (after six practicals). Each was designed by a pathology domain expert to only cover the pathology constructed by the three practicals prior to administering the examination, i.e. the examinations at crossover tested material covered in practicals 1-3, whilst those administered at the end of the study tested material covered in practicals 4-6.

7.5.2 The Relationship between MCQ and Short Answer Tests

An assumption was made that students who perform well in one method of assessment would perform well in the other. Scatter diagrams were employed to highlight any correlation between the two examination methods administered at crossover (after three practicals) and at the end of the study (after six practicals).

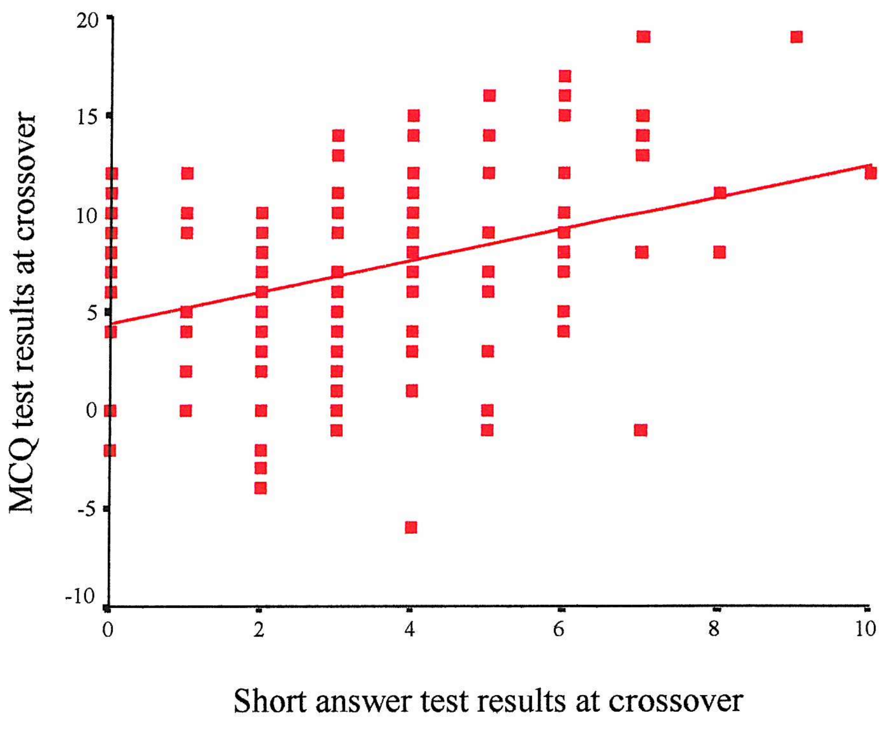


Figure 17 Relation between MCQ and short answer test results obtained at crossover.

Figure 17 presents the scatter diagram of MCQ by short answer test results at crossover.

The line of best fit suggested a positive relationship between the two methods of examination: an increase (decrease) in MCQ test results was correlated with an increase (decrease) in short answer test results. The slope of this regression line determined that for every increment of short answer test results, the MCQ results increased by 0.81. Pearson's r correlation coefficient $r = 0.36$ suggested a weak, low association between the results of these two examination methods. The coefficient of determination $r^2 = 0.13$ indicated that the regression line only explained approximately 13% of the variance in the data, therefore giving low confidence in the prediction made by the best fit line. However, a significant inferential statistic ($p < 0.0005$) confirmed that the association, albeit weak with low confidence, was attributed to the population.

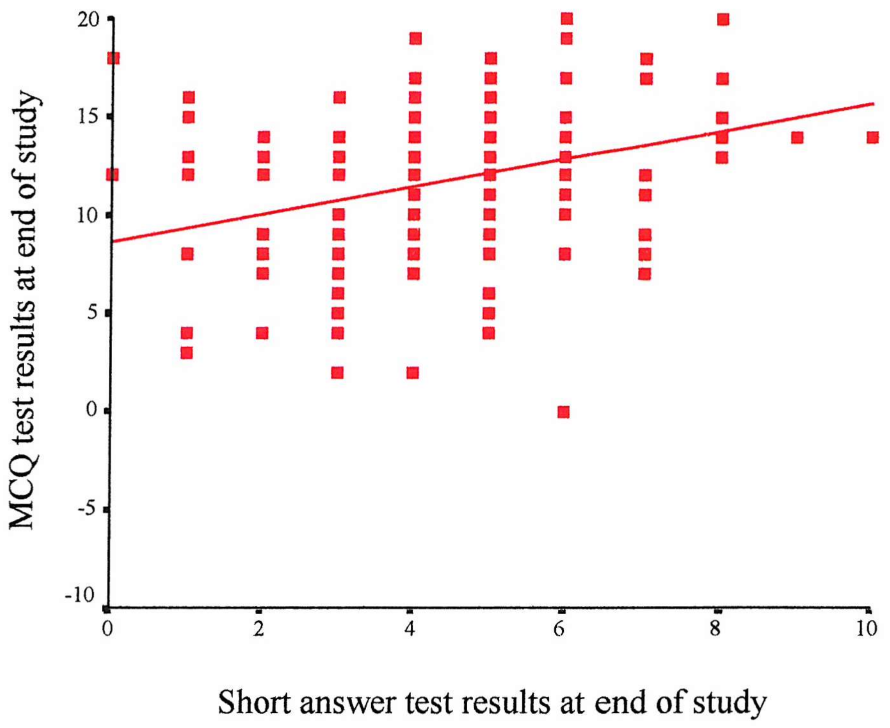


Figure 18 Relation between MCQ and short answer test results obtained at the end of the study.

Analysis of the MCQ test results at the end of the study (see Figure 18) again suggested a positive relationship between the two methods of examination. The line of best fit predicted that for every increment of short answer test results, the MCQ results increased by 0.70. Pearson's r correlation coefficient $r = 0.32$ suggested a weak, low association. The coefficient of determination $r^2 = 0.11$ indicated that the regression line explained a low proportion, approximately 11%, of the variance in the data, therefore giving low confidence in the prediction made by the best fit line. However, as with the results at crossover, the inferential statistic of $p < 0.0005$ was significant, indicating a weak association with low confidence.

These results made it inappropriate to take one method of assessment as a good indicator of the other: students who performed poorly in the MCQ tests may have performed well in the short answer tests. Conversely, students who performed well in MCQ tests may have performed relatively poorly in the short answer tests. This inference led to the decision to consider both MCQ and short answer test results when analysing the education attainment aspect of pedagogic value.

7.5.3 Operationalization of Sub Hypothesis

Sub hypothesis 1 (see section 6.4.1) was operationalized to address the education attainment dimension of pedagogic value by comparing the means of the two sample populations as follows:

H_0: the mean level of knowledge exhibited by medical students who use SCALPEL (μ_1), does not differ from the mean level exhibited by those who work in the traditional pathology laboratory (μ_2). $H_0 : \mu_1 = \mu_2$.

H_1: the mean level of knowledge exhibited by medical students who use SCALPEL (μ_1), differs from the mean level exhibited by those who work in the traditional pathology laboratory (μ_2). $H_1 : \mu_1 \neq \mu_2$.

Past empirical studies (Jelovsek and Adebonojo 1993) suggest that there is no reason to suspect a systematic difference in favour of courseware or traditional teaching, i.e. $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$, therefore the alternative hypothesis (H_1) was stated without specifying a directional difference.

7.5.3.1 Results of the Between-Subjects Analysis

Analysis of between-subjects was performed to test for variations in education attainment between the experiment and control groups. The former group of students were exposed to SCALPEL whilst the later worked through the practicals in the traditional laboratory. Independent samples t-test for the equality of means was employed to test for differences between the two groups. Four tests were performed, one for each of the two examination methods administered at crossover and at the end of the study.

For the equality of MCQ test means at crossover (see Appendix 1, Table A1 - 2) a difference of 0.70 marks was observed between the two sample means. The two-tailed significance ($p = 0.38$) gave confidence in accepting the null hypothesis of no difference between the MCQ test results of the two groups. Turning to the short answer test means at crossover, a -0.42 difference in marks between the two groups was observed (see Appendix 1, Table A1 - 3). With Levene's test for the equality of variances ($F = 3.96$, $p = 0.49$) suggesting a significant difference in the underlying variances of the two groups, the unequal variances two-tailed significance ($p = 0.28$) was chosen. Confidence was again given in accepting the null hypothesis, concluding that there was no difference between the mean level of education attainment of the two groups of students at crossover.

The same tests were applied to the MCQ and short answer tests results at the end of the study at which point the control and experiment groups had crossed-over. The equality of MCQ test means (see Appendix 1, Table A1 - 4) highlighted a non-significant ($p = 0.90$) -0.08 difference in marks between the two sample means. Finally, the -0.35 difference between the two short answer test sample means was also determined to be non-significant ($p = 0.25$) (see Appendix 1, Table A1 - 5), concluding that there was no difference in the mean level of education attainment between the two groups of students.

7.5.3.2 Results of the Within-Subjects Analysis

The between-subject method of analysis, where the samples were compared as a whole, had the disadvantage of extraneous variables, such as motivation, potentially biasing the results. By exposing students to both methods of teaching and using them as their own controls, the within-subject design controlled for such bias by holding all extraneous variables constant. Having identified two groups of students by their exposure order, LAB-SCALPEL and SCALPEL-LAB, paired-sample t-tests were employed to test for significant changes in individuals MCQ and short answer test results.

For the group with exposure order LAB-SCALPEL, the means of the same subjects' MCQ results obtained at crossover and at the end of the study were compared (see Appendix 1, Table A1 - 6). Short answer test results for the same group of students were also addressed (see Appendix 1, Table A1 - 7). Analyses of the MCQ (mean difference = -4.12, $p < 0.0005$) and short answer test (mean difference = -1.16, $p = 0.001$) results suggest that on average students achieved higher results after exposure to SCALPEL. Conversely, students with exposure order SCALPEL-LAB achieved higher MCQ (mean difference = -4.85, $p < 0.0005$) and short answer test results (mean difference = -1.18, $p = 0.002$) after participating in the laboratory (see Appendix 1, Table A1 - 8 and Table A1 - 9).

7.5.4 Discussion

In conclusion, the between-subjects statistical analyses give confidence that for each of the four examinations there was no significant difference between the mean results of students who used SCALPEL and those who worked in the traditional laboratory. This positive finding is backed-up by the between-subjects analyses. On average, students with exposure order LAB-SCALPEL performed better in both MCQ and short answer examinations after using SCALPEL. However, students with exposure order SCALPEL-LAB performed better in both examination methods after working in the laboratory environment. This suggests that time, and not exposure, was significant to education attainment. Therefore, it can be reasoned that irrespective of the teaching method, on average students performed better in both methods of examination at the end of the study.

Given this empirical evidence the null hypothesis was not rejected, concluding that: teaching pathology by delivering practicals in a laboratory or by using SCALPEL is equally educationally effective.

7.6 Acceptance of SCALPEL

This section addresses students' acceptance of SCALPEL (see section 6.3.5), the second concept defined to measure pedagogic value (see section 6.2.1).

7.6.1 Operationalization of Sub Hypotheses

Sub hypothesis 2 (see section 6.4.1) was operationalized to address students' acceptance of SCALPEL by looking at attitudes towards the two teaching environments, SCALPEL and the traditional pathology laboratory.

H_0: Medical students exhibit no more appreciation of SCALPEL than the traditional pathology laboratory.

H_1: Medical students exhibit more appreciation of SCALPEL than the traditional pathology laboratory.

7.6.2 The Distribution of Student Attitudes Towards SCALPEL

The distribution of the sample for the concept 'acceptance of SCALPEL' (see Appendix 1, Table A1 - 10) was taken as an indication of how well SCALPEL and the traditional laboratory environments were received by the students. The data for this measure were collected at the end of the study, at which point every student had exposure to the traditional laboratory and SCALPEL environments for 3 practicals. Presented graphically in Figure 19, only 23.4% of students expressed a positive attitude towards SCALPEL (strongly agree and agree categories). 50.4% were neutral, indicating no preference, with the remaining 26.2% expressing negative attitudes towards SCALPEL (disagree and strongly disagree categories), indicating a preference towards the traditional laboratory environment.

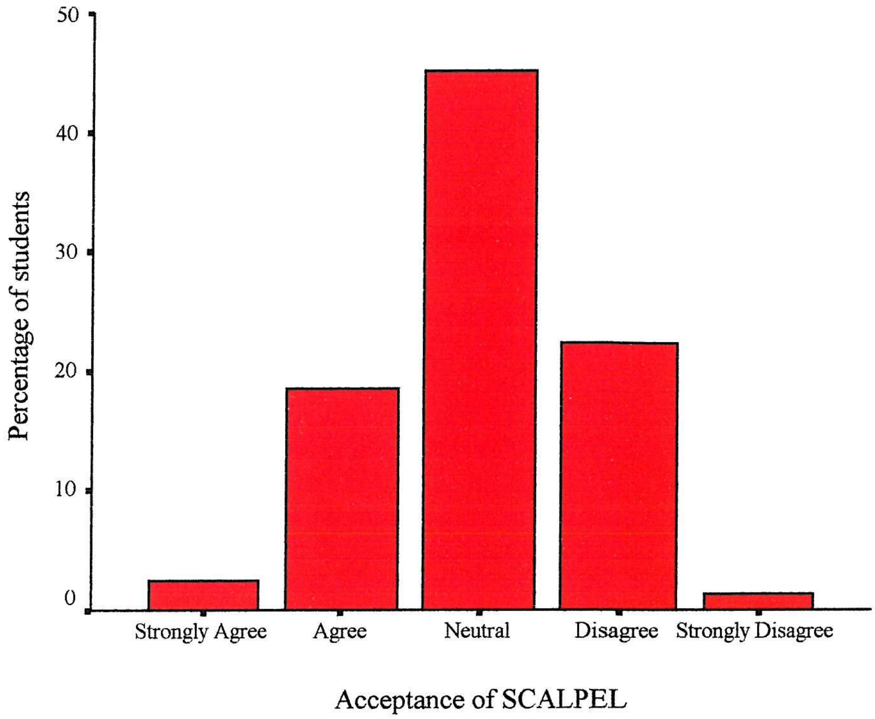


Figure 19 Bar chart on the distribution of students by acceptance of SCALPEL.

7.7 Acceptance of SCALPEL and Education Attainment

A prediction was made that a relationship would exist between students' acceptance of SCALPEL and education attainment, the two concepts used throughout the evaluation to measure the pedagogic value of SCALPEL.

7.7.1 Operationalization of Sub Hypotheses

Sub hypothesis 3 (see section 6.4.1) was operationalized to test the prediction that the distribution of students' acceptance of SCALPEL would be related to the distribution of education attainment.

H_0: the distribution of medical students' education attainment is not associated with the distribution of acceptance of SCALPEL.

H_1: the distribution of medical students' education attainment is associated with the distribution of acceptance of SCALPEL.

7.7.2 Results

A measure of association was performed between acceptance of SCALPEL and each of the examinations administered during the study. For each analysis, the sample population consisted of those students who used SCALPEL prior to the relevant examination. With acceptance of SCALPEL being measured at the ordinal level, the lowest level of measurement of the constituent variables, Spearman's rho was employed as the appropriate statistical test.

The rho coefficient for each analysis (see Appendix 1, Table A1 - 11) presented a non-significant negligible relationship. These results gave confidence to accept the null hypothesis that medical students' education attainment is not associated with acceptance of SCALPEL.

7.8 Pedagogic Value and Demonstrator Assistance

This section addresses the effect of the need for demonstrator assistance (see section 6.3.5) on the pedagogic value of SCALPEL.

7.8.1 The Distribution of Student Attitudes Towards the need for Demonstrators

The distribution of the sample for students' expressed need for demonstrators during SCALPEL sessions (see Appendix 1, Table A1 - 12) indicated an overwhelming dependence on demonstrators. The data for this measure were collected at the end of the study, at which point every student had exposure to the traditional laboratory and SCALPEL environments for 3 practicals. Presented graphically in Figure 20, 57.2% of students expressed a need for demonstrator assistance during SCALPEL sessions (strongly agree and agree categories). 37.7% were neutral, and only 5.1% expressed no need for demonstrators (disagree category).

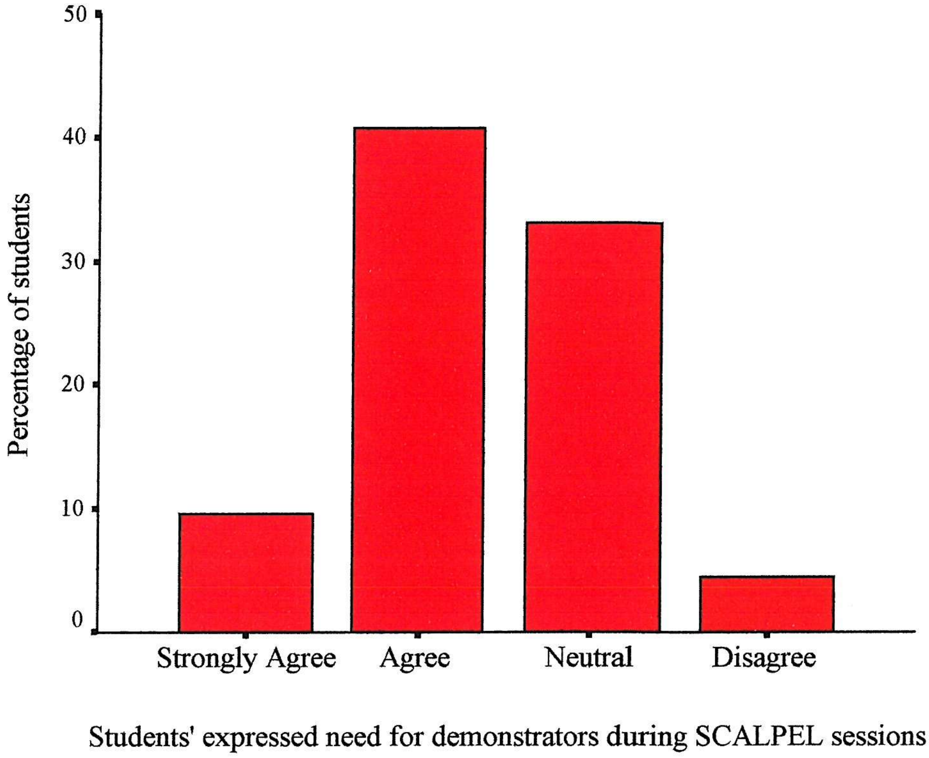


Figure 20 Bar chart on the distribution of students by expressed need for demonstrators during SCALPEL sessions.

7.8.2 Operationalization of Sub Hypotheses

Research sub hypothesis 4 (see section 6.4.2.1) was operationalized to test for associations between the need for demonstrator assistance and the constituent concepts of pedagogic value, education attainment and acceptance of SCALPEL, as follows:

H_0: education attainment is independent of a student's need for demonstrator assistance.
H_1: education attainment is dependent on a student's need for demonstrator assistance.

H_0: expressed need for demonstrator assistance is independent of a student's acceptance of SCALPEL.
H_1: expressed need for demonstrator assistance is dependent on a student's acceptance of SCALPEL.

7.8.2.1 Education Attainment and Demonstrator Assistance

Because no relationship was found between the results of the MCQ and short answer tests (see section 7.5.1), looking for a relationship between the need for demonstrator assistance and education attainment required one statistical test for each of the examinations administered at crossover and at the end of the study.

With students' expressed need for demonstrator assistance being measured at the ordinal level, the lowest level of measurement of the constituent variables, Spearman's rho was employed as the appropriate statistical test, and found no relationships with each of the examinations (see Appendix 1, Table A1 - 13).

7.8.2.2 Acceptance of SCALPEL and Demonstrator Assistance

Given that acceptance of SCALPEL was measured at the same time as expressed need for demonstrators, the independent (cause) and dependent (effect) variables could not be established by the time order of the two concepts. Theoretically the causal relationship between the two concepts can be reasoned either way, inferring that each concept is just as likely to be influenced by the other.

Table 1 presents Spearman's rho coefficients for acceptance of SCALPEL and the need for demonstrator assistance. A significant moderately negative relationship ($\rho = -0.40$, $p < 0.0005$) was found to exist. The null hypothesis was therefore rejected, inferring that: students who expressed a need for demonstrators to be available during SCALPEL sessions indicated a lower acceptance of SCALPEL than those who suggested a lesser need for demonstrators.

Table 1 Measure of association between students' expressed need for demonstrator assistance and acceptance of SCALPEL

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
acceptmi	136	-0.40**	0.000

** Correlation is significant at the 0.01 level (2-tailed)

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on the same questionnaire. Each sample (N) was made up of students who expressed an attitude towards the acceptance of SCALPEL and the need for demonstrator assistance. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

7.8.3 Discussion

Addressing students' expressed need for demonstrator assistance, no statistical or practical relationship was found with education attainment. However, a significant negative relationship was found to exist with acceptance of SCALPEL. With acceptance of SCALPEL already causing concern (see section 7.6.2), this relationship was considered significant to this research and subsequently informed the decision to investigate technologies that would allow SCALPEL to facilitate the role of demonstrators in the traditional laboratory.

7.9 Pedagogic Value and Attitudes Towards Computers

This section is concerned with the effect students' pre-SCALPEL attitudes towards computers (see section 6.3.5) might have had on the measured pedagogic value of SCALPEL. Given that pre-SCALPEL attitudes towards computers were measured before the constituent concepts of pedagogic value, the time order criteria of the causality model was satisfied. The dependent variables were determined to be acceptance of SCALPEL, and the MCQ and short answer tests administered at crossover and at the end of the study. The independent variable for each analysis was pre-SCALPEL attitudes towards computers.

7.9.1 Operationalization of Sub Hypotheses

Research sub hypothesis 5 (see section 6.4.2.1) was operationalized to test for associations between student attitudes towards computers and the two constituent concepts of pedagogic value, education attainment and acceptance of SCALPEL, as follows:

H_0: education attainment is independent of a student's pre-SCALPEL attitude towards computers.

H_1: education attainment is dependent on a student's pre-SCALPEL attitude towards computers.

H_0: acceptance of SCALPEL is independent of a student's pre-SCALPEL attitude towards computers.

H_1: acceptance of SCALPEL is dependent on a student's pre-SCALPEL attitude towards computers.

7.9.1.1 Education Attainment and Pre-SCALPEL Attitudes Towards Computers

With the absence of a relationship between the results of the MCQ and short answer tests (see section 7.5.1), looking for a relationship between students' pre-SCALPEL attitudes towards computers and education attainment required one statistical test for each of the examinations administered at crossover and at the end of the study.

This analysis addresses acceptance of SCALPEL and not the laboratory environment, as a consequence the sample population for this analysis constituted those students who had exposure to SCALPEL prior to the relevant examination. Employing Spearman's rho coefficient, no relationships were found between students' pre-SCALPEL attitudes towards computers and each of the examinations administered during the evaluation (see Appendix 1, Table A1 - 14).

7.9.1.2 Acceptance of SCALPEL and Pre-SCALPEL Attitudes Towards Computers

The entire sample population was used for this analysis since acceptance of SCALPEL was measured after all students had had exposure to SCALPEL. Spearman's rho coefficient ($\rho = 0.05$) highlighted a very weak, negligible relationship between pre-SCALPEL attitudes towards computers and acceptance of SCALPEL (see Appendix 1, Table A1 - 15). With the inference statistic ($p = 0.59$, $N = 127$) being non-significant, the null hypothesis could not be rejected, giving confidence that no relationship existed between pre-SCALPEL attitudes towards computers and acceptance of SCALPEL.

7.9.2 Discussion

Given the shift from a traditional laboratory to a courseware environment, attitudes towards computers exhibited by students before exposure to SCALPEL were considered important in that such attitudes could influence both education attainment and acceptance of a courseware learning environment. However, no statistical or practical relationships were observed between students' pre-SCALPEL attitudes towards computers and each of the constitute dimensions of pedagogic value: education attainment and acceptance of SCALPEL.

7.10 Pedagogic Value and Learning Styles

This section addresses the effect of student learning styles (see section 6.3.4) on the pedagogic value of SCALPEL. With students' pre-SCALPEL learning styles being measured before the constituent concepts of pedagogic value, the time order criteria of the causality model was satisfied. The dependent variables were therefore determined to be acceptance of SCALPEL, and the MCQ and short answer tests administered at crossover and at the end of the evaluation. The independent variable was pre-SCALPEL learning.

7.10.1 Operationalization of Sub Hypotheses

Research sub hypothesis 6 (see section 6.4.2.1) was operationalized to test for associations between learning styles and the constituent concepts of pedagogic value, education attainment and acceptance of SCALPEL, as follows:

H_0: no association exists between students' acceptance of SCALPEL and learning styles.
H_1: an association exists between students' acceptance of SCALPEL and learning styles.

H_0: education attainment is independent of a student's pre-SCALPEL learning style.
H_1: education attainment is dependent on a student's pre-SCALPEL learning style.

7.10.1.1 Education Attainment and Learning Styles

For each analysis, the sample population consisted of those students who used SCALPEL prior to each examination. Given that no association was found to exist between the results of the MCQ and short answer tests (see section 7.5.1), relationships between pre-SCALPEL learning styles and education attainment required one statistical test for each of the examinations administered at crossover and at the end of the study.

Employing Spearman's rho coefficient, no relationships were found to exist between students' pre-SCALPEL learning styles and each of the examinations administered during the study (see Appendix 1, Table A1 - 16).

7.10.1.2 Acceptance of SCALPEL and Learning Styles

The entire sample population was used for this analysis since acceptance of SCALPEL was measured after all students had exposure to SCALPEL. Spearman's rho correlation coefficient (see Appendix 1, Table A1 - 17) found no statistical or practical relationship between acceptance of SCALPEL and learning styles.

7.10.2 Discussion

No relationship was found between learning styles and acceptance of SCALPEL, and between learning styles and each of the four academic tests administered during the study. These results gave confidence in accepting the null hypothesis of no difference, concluding that the pedagogic quality of SCALPEL is not dependent on student learning styles.

7.11 Acceptance of SCALPEL and Demographic Characteristics

The causality model was employed to test predictions that demographic characteristics (independent variables) may have influenced students' acceptance of SCALPEL (dependent variable).

7.11.1 Operationalization of Sub Hypotheses

The concept of student demographics was operationalized in section 6.3.2. This process introduced the following dimensions: age group, gender and demographic area of post-primary education. To address these dimensions research sub hypothesis 7 (see section 6.4.2.1) was operationalized to produce the following sub hypotheses:

H_0: students' acceptance of SCALPEL is independent of age.
H_1: students' acceptance of SCALPEL is dependent on age.

H_0: students' acceptance of SCALPEL is independent of gender.
H_1: students' acceptance of SCALPEL is dependent on gender.

H_0: students' acceptance of SCALPEL is independent of the demographic area of post-primary education.
H_1: students' acceptance of SCALPEL is dependent on the demographic area of post-primary education.

7.11.2 Results

Spearman's rho showed no relationship between students' acceptance of SCALPEL and age (see Appendix 1, Table A1 - 18). Given the nominal level of measurement for gender and demographic area of post-primary education, the Goodman-Kruskal Tau PRE measure of association was used (see Appendix 1, Table A1 - 19). With students' acceptance of SCALPEL dependent, the asymmetric values for both demographics showed no significant relationship. Consequently the null hypotheses were accepted. Students' age, gender and demographic area of post-primary education does not influence their acceptance of

SCALPEL.

7.12 Acceptance of SCALPEL and Socio-economic Characteristics

The causality model was employed to test predictions that student socio-economic characteristics (independent variables) may have influenced acceptance of SCALPEL (dependent variable).

7.12.1 Operationalization of Sub Hypotheses

Research sub hypothesis 8 (see section 6.4.2.1) was operationalized to address the operational definition of social economic characteristics (see section 6.3.3) as follows:

H_0: students' acceptance of SCALPEL is independent of previous higher education.
H_1: students' acceptance of SCALPEL is dependent on previous higher education.

H_0: no association exists between students' acceptance of SCALPEL and previous access to computers.
H_1: an association exists between students' acceptance of SCALPEL and previous access to computers.

7.12.2 Results

For each analysis, Spearman's rho highlighted very weak, non significant relationships (see Appendix 1, Table A1 - 20). Consequently, the socio-economic null hypotheses were not rejected, concluding that relationships do not exist between acceptance of SCALPEL and previous higher education, and acceptance of SCALPEL and access to computers.

7.13 Acceptance of SCALPEL and Behavioural Characteristics

The causality model was employed to test for a relationship between acceptance of SCALPEL (dependent variable) and previous computer use (independent variables), an operationalized dimension of student behaviour (see section 6.3.4). Analysis of acceptance of SCALPEL and the other dimension of student behaviour, learning styles, is addressed in section 7.10.1.2.

7.13.1 Operationalization of Sub Hypotheses

Research sub hypothesis 9 (see section 6.4.2.1) was operationalized to address the previous computer use dimension of behavioural characteristics only.

H_0: no association exists between students' acceptance of SCALPEL and previous use of computers

H_1: an association exists between students' acceptance of SCALPEL and previous use of computers

7.13.2 Results

Spearman's rho correlation coefficients for acceptance of SCALPEL and previous computer use (see Appendix 1, Table A1 - 21) found no statistical or practical relationship. Therefore the null hypothesis was not rejected concluding that previous computer use does not influence students' acceptance of SCALPEL.

7.14 Acceptance of SCALPEL and Opinions and Attitudes

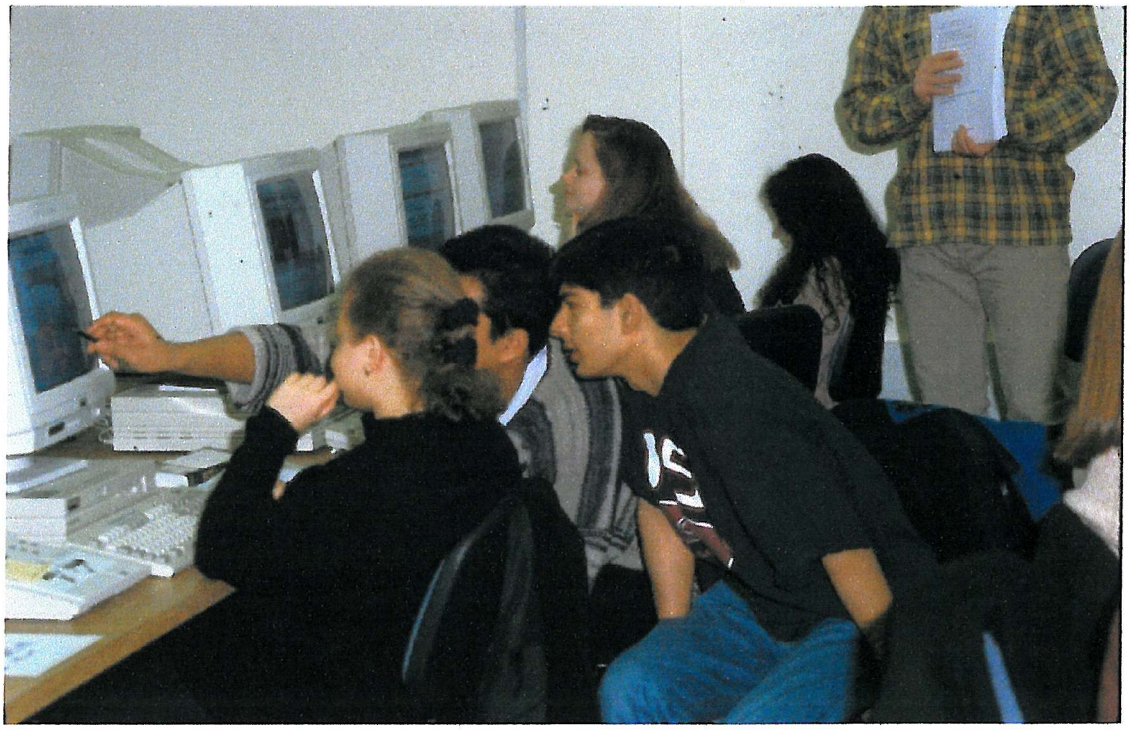
Statistical analysis between acceptance of SCALPEL and students' expressed need for demonstrator assistance is presented in section 7.8.2.2, and attitudes towards computers in section 7.9.1.2.

7.15 Participant Observations

Participant observation techniques (see section 5.4.3) were employed to add a qualitative dimension to the evaluation with a view of formulating a situated, holistic and well-rounded account (Bryman 1998) of the two teaching environments under investigation: the traditional laboratory and SCALPEL. Observations of students in the traditional laboratory have already been discussed in section 4.2.1. This section reports on the observations of students using the SCALPEL courseware whilst making comparisons with what was observed in the traditional laboratory.

Participant observations took place in a university computer suite of 27 workstations. The suite was closed to all students and staff other than the two demonstrators and 27 students scheduled to use SCALPEL for the pathology practical.

Asking for assistance from demonstrators or peers solved initial problems with the SCALPEL courseware. The majority of the problems with the application were often a result of a student's low appreciation of elementary 'windows' management techniques and were quickly overcome once they were told of the benefits of closing windows that were no longer relevant to the practical.



Picture 1 Peer dialogue was not as common during SCALPEL sessions.

Always in demand, the demonstrators were called upon to discuss the domain material with most of the discussions focusing on the pathology images. An occasional complaint was made that the images were too large to be appreciated via the computer and that having the whole image available, as was the case in the traditional laboratory, was more beneficial. Demonstrators were observed moving around repeating the same explanations about the domain material. This one-one student-demonstrator dialogue often meant that students were left waiting for a demonstrator to become available. Most students worked through the practical individually with the occasional group of not more than 3 students temporarily forming to engage in dialogue (see Picture 1): the overall noise level of the computer suite suggested that most students worked alone. This observation was in contrast with the patterns of interaction in the traditional laboratory where students working in groups of up to eight actively engaged in discussions about the domain material; and where demonstrator explanations and discussions were of benefit to more than one student. Relative to the traditional laboratory there was not much use of supplementary material in the way of textbooks and lecture notes. When asked, students indicated that this was due to lack of workspace in the computer suite (see Picture 1). These comments were consistent with the observation that where students did use textbooks, they were often kept on the floor.

In conclusion, the relaxed and studious atmosphere in the computer suite encouraged students to work through the practicals individually with peer dialogue kept to a minimum. Initial problems with the functionality of the courseware were soon overcome and students focused on working through the practical. Demonstrators were over-stretched as a result of students working alone. For some students, having to wait for demonstrators to become available was visibly frustration. Finally, participant observation highlighted the need to consider the effect of ergonomic factors (Preece et al. 1994) on students' attitudes towards SCALPEL as an effective method of participating in practicals. e.g. layout of computers in relation to each other, lack of workspace etc: ergonomics could be as influential on student attitudes to SCALPEL as the courseware itself.

7.16 SCALPEL Focus Group

A focus group held with a random selection of eight students engaged in interesting discussions about the SCALPEL environment. The process took the form of an informal exploratory interview (see section 5.4.4) with the discussion threads controlled by the participants.

Accepting SCALPEL as a valid teaching medium, the group stressed the continuing need for demonstrators to be available in order to discuss the pathology material. Having enforced this point, they focused on problems with SCALPEL and how the experience could be improved.

Although aware of the necessity to close windows no longer relevant to the current educational thread, the group was united in expressing difficulties navigating through the myriad of SCALPEL windows. Accessing the supplementary pathology material from the case studies meant that at least three windows needed to be referenced at any given time: namely the case study text, MCQ engine, and supplementary material.

An issue was raised concerning the area of an image visible at any given time. In the laboratory the entire image could be viewed. Limitations of the computer hardware used during the SCALPEL sessions rendered unavailable otherwise readily available functionality to zoom in or out of an image. This left small areas of images visible at any given time. This had the disadvantageous effect of requiring students to scroll around images in search of the relevant information. Conversely, students commented on the benefits of the histology within SCALPEL in that the viewing area of a slide is greater than that observed under a microscope. Feeling more involved by actually focusing and navigating the histology slides, one student commented on a preference to use a microscope. A request was also made for the images (including histology) to be augmented with text descriptions. The group received this suggestion with mixed feelings. The majority agreed, while a call to leave the images with no descriptions was argued from the standpoint that this encourages the learner to interpret what is observed: if text is provided it would be too easy just to read the text instead of interpreting the image. Other criticisms included poor video quality with a general consensus that the video segments offered no educational benefit and failed to keep the interest of the learner.

A discussion about introducing sound stirred enthusiasm within the group. It was suggested that sound would improve interaction and provide a realistic environment by delivering patient history and symptoms. It was also suggested that sound could be used to recreate the occasion of discussing a patient with a colleague or listening to a patient describing his/her symptoms and medical history. This led students to express a wish for more simulations like the ECG to be implemented in the application. They felt that such functionality went a long way to engaging the learner and offering a realistic environment.

The MCQ engine was found to provide helpful feedback on selecting correct as well as incorrect answers, with students discussing how they purposely selected wrong answers so to observe given explanations. However the MCQ engine came under much criticism. It was suggested that the interface was cumbersome, non-intuitive, and intruded on the main case history document and supplementary material.

On a social plain the group had mixed feelings regarding the preferred learning environment: the less sociable environment forced upon the learner by the SCALPEL application; and the sociable environment encouraged by traditional teaching in the laboratory. The former was preferred because of the studious atmosphere while the later was found to be more conducive to peer and demonstrator dialogue.

In conclusion, the group seemed generally satisfied with the quality of the application material. The main point was about the continued need for demonstrators to be available. Relative to the SCALPEL courseware, there were strong ill feelings about the MCQ engine with calls for this functionality to be refined, and for video quality to be improved.

7.17 Discussion

The statistical analyses presented in this chapter focused on the pedagogic value of

SCALPEL relative to the traditional pathology laboratory. Addressing the pre-specified research agenda, pedagogic value was assessed by measuring education attainment and acceptance of the courseware environment. The first statistical analysis looked for a relationship between the two methods of assessment used to measure education attainment: MCQ and short answer tests. Scatter diagrams with linear regression techniques found a weak correlation significant in the population. Given the weakness of this relationship, the decision was made that one method of assessment could not be taken as a reliable indicator for the other. Consequently all hypotheses testing involving education attainment required consideration of both MCQ and short answer test results.

The crossover research design allowed both between-subjects and within-subjects analyses of the samples' examination results. The between-subjects analyses found no statistical difference between the mean results of students who used SCALPEL and those who worked in the traditional laboratory. The within-subjects analysis strengthened the research design by controlling for extraneous variables. These statistics showed that, irrespective of assignment to SCALPEL or the laboratory, students performed better in both methods of assessment at the end of the study. Although the empirical evidence does not support the primary research hypothesis that courseware could replace pathology laboratory practicals with improved education attainment, the results do give confidence in concluding that pathology practicals are equally educationally effective, in terms of domain material, when delivered in the laboratory or by SCALPEL.

Attitudes towards SCALPEL were not so encouraging with only 23.4% of students expressing an agreeable attitude. 50.4% of the sample were indifferent, neither expressing an agreeable or disagreeable attitude, while the remaining 26.2% expressed dissatisfaction. It was assumed that students who subscribed to SCALPEL would learn more than those who expressed dissatisfaction. However, no evidence of a relationship was found between the two dimensions of pedagogic value: education attainment and acceptance of SCALPEL. This suggests that SCALPEL was an effective pedagogic environment even for students who didn't express an agreeable attitude towards working through pathology practicals using SCALPEL. It can also be inferred that, unexpectedly, SCALPEL did not improve the education of students who subscribed to its use. Qualitative data from participant observations and a focus group highlighted the need to consider the effect of ergonomic factors on students' attitudes towards SCALPEL with a notable decline in the use of supplementary material (i.e. textbooks) being attributed to lack of desk space; and frustration expressed about the inability to view whole images being a result of working on low specification computers.

Turning to the curiosity driven research, it was anticipated that SCALPEL would not be of pedagogic value to all students. Three concepts were identified as potentially important influencing factors. These were students' expressed need for demonstrator assistance, attitudes towards computers, and behavioural learning styles. No relationships were observed between the constituent concepts of pedagogic value and, learning styles and attitudes towards computers. In addition no relationship was found between students' expressed need for demonstrators during SCALPEL sessions and education attainment. However, a significant moderately negative relationship with acceptance of SCALPEL suggests that: students who expressed a need for demonstrators during exposure to SCALPEL indicated a lower acceptance of SCALPEL than those who suggested a lesser need. An explanation for this is that those students who called upon demonstrators during the practicals did so because the application didn't offer adequate domain material, therefore giving them low confidence in SCALPEL. This also explains those students who didn't require demonstrator assistance and expressed a higher opinion of SCALPEL, as the relationship also suggests. It can be inferred that these students found SCALPEL an acceptable medium for pathology practicals, with sufficient domain material to work through the practicals, and therefore had a lesser need for demonstrator assistance. Recognised as a failing of the research design, other than the qualitative data collected through the observational research, students' expressed need for demonstrator assistance in the traditional laboratory were not measured. With this omission of data, analyses to determine if students' dependence on demonstrators was greater in the laboratory than with SCALPEL, or visa versa, could not be assessed. However, the distribution of student attitudes towards demonstrators (see section 7.8.1), and the unanimous cries reverberated by the focus group, highlights the magnitude of student opinion. With the descriptive statistic showing such a high percentage (57.2%) of students expressing a need for demonstrator assistance, the significant relationship with acceptance of SCALPEL was considered to be consequential to this research (see section 7.8.2.2): not only did the majority of students express a need for demonstrator assistance, it was shown that those who did not held higher opinions of SCALPEL. This inferred that improving SCALPEL to address students' reasons for requiring demonstrators would have a positive impact by reducing their need for demonstrators to be available, with the subsequent effect of improving attitudes towards SCALPEL.

A look at the association between the constituent concepts of pedagogic value and student learning styles and attitudes towards computers found no significant relationships. With no mean difference in education attainment between the two pedagogic environments, further investigations involving education attainment were considered unnecessary. However, with an overwhelming negative attitude towards SCALPEL, further relationships were investigated involving the student-specific characteristics identified in section 6.2.2. The first analysis focused on acceptance of SCALPEL and the three dimensions of student demographics: age, gender and demographic area of post-primary education. However, no relationships were found, accepting the null hypotheses and concluding that: student's age, gender and demographic area of post-primary education did not influence their acceptance of SCALPEL. Turning to the dimensions of socio-economic characteristics, no relationships were found between acceptance of SCALPEL and previous higher education and access to computers. Finally, the relationship between acceptance of SCALPEL and students' previous computer use, a dimension of behavioural characteristics, was analysed. Once again, the absence of a relationship indicates that previous computer use did not influence students' acceptance of SCALPEL. Although caution is expressed in generalising these empirical findings to wider populations, they offer good news for computer practitioners in that individual characteristics are less of a factor in the successful use of computer applications, and that more heterogeneous user populations are being observed. This theme is explored further in chapter 8 with a look at the influence of student-specific characteristics determined to be of particular interest to this research: the need for demonstrator assistance, attitudes towards computers, and learning styles.

7.18 Conclusions

The data analysed in this chapter served to test the primary research hypothesis that courseware could replace pathology laboratory practicals with improved undergraduate medical students' education attainment and acceptance of learning (see section 6.2.1). Addressing first the education attainment element of this hypothesis, inferences gave confidence in concluding that pathology practicals were equally educationally effective, in terms of domain material, when delivered in the laboratory or by SCALPEL. Although this result was not in favour of accepting the research hypothesis, it was positive in that it showed that SCALPEL was at least no worse that the traditional laboratory environment. Addressing the curiosity driven research hypothesis that replacing pathology laboratory practicals with SCALPEL would not be pedagogically valuable to all students, there was no evidence to suggest that student demographics, socio-economic, and behavioural characteristics biased the pedagogic value of SCALPEL for individual or groups of students. It was observed however that students' negative acceptance of SCALPEL was related to their expressed need for demonstrator assistance during SCALPEL sessions. In conclusion, the empirical evidence presented in this chapter generated interesting avenues of enquiry for improving SCALPEL.

7.19 Summary

This chapter presented the evaluation findings of the pedagogic value of SCALPEL. The evaluation methodology was discussed before presenting the results of the quantitative statistical analyses and the qualitative findings of participant observations and an informal focus group. Finally a discussion was offered on the interpretation of the findings and their impact on the research focus. The next chapter continues with the first study's curiosity driven research agenda by looking at student-specific characteristics.

Chapter 8: The First Empirical Study: What Students Bring to Pathology Practicals

8.1 Introduction

The previous chapter focused on the pre-specified research agenda of the first empirical study. This chapter continues with the curiosity driven research agenda of the first empirical study by looking at the relationships between student-specific characteristics and concepts identified as being of particular interest to this research: students' expressed need for demonstrator assistance during SCALPEL sessions, attitudes towards computers, and learning styles. Sections 8.2 looks for cause and effect relationships between students' expressed need for demonstrator assistance and the conceptualized dimensions of the four types of social survey subject matter introduced in section 6.2. Sections 8.3 (attitudes towards computers) and 8.4 (learning styles) continue this theme with the added dimension of analysing changes in the respective attitudes after exposure to SCALPEL. Finally a discussion is offered in section 8.5 on all the statistical analyses presented in this chapter.

8.2 Attitudes towards demonstrator assistance during exposure to SCALPEL

Students' expressed need for demonstrator assistance has already been highlighted as a concept of significant importance to this research. This section focuses on establishing cause and effect relationships with student demographics, socio-economic, behavioural and attitudinal characteristics.

Data for students' expressed need for demonstrator assistance were collected after the other concepts, therefore establishing the time order and giving confidence to specify this as the dependent variable for each analysis.

8.2.1 Student Demographics

The causality model was employed to test the prediction that demographics (independent variables) influence students' attitudes towards demonstrator assistance during exposure to SCALPEL (dependent variable).

8.2.1.1 Operationalization of Sub Hypotheses

Research sub hypothesis 10 (see section 6.4.2.2) was operationalized to test for associations between students' expressed need for demonstrator assistance and the conceptualized dimensions of student demographics: age group, gender and demographic area of post-primary.

H_0: medical students' expressed need for demonstrator assistance is independent of age.
H_1: medical students' expressed need for demonstrator assistance is dependent on age.

H_0: medical students' expressed need for demonstrator assistance is independent of gender.
H_1: medical students' expressed need for demonstrator assistance is dependent on gender.

H_0: medical students' expressed need for demonstrator assistance is independent of the demographic area of post-primary education.
H_1: medical students' expressed need for demonstrator assistance is dependent on the demographic area of post-primary education.

8.2.1.2 Results

Spearman's rho found no relationship between students' expressed need for demonstrator assistance and age (see Appendix 2, Table A2 - 1). Given the nominal level of measurement for gender and demographic area of post-primary education the Goodman-Kruskal Tau PRE measure of association was used (see Appendix 2, Table A2 - 2). The statistics shown are asymmetric measures with students' expressed need for demonstrator assistance dependent. No relationship was found to exist with gender. However, a very weak (0.08) relationship with demographic area of post-primary education was found to be significant ($p = 0.002$) in the population. Although this relationship was statistically significant, it was regarded as too weak to be of any practical concern.

These results gave confidence in accepting the null hypotheses, concluding that students' attitudes towards the need for demonstrator assistance during SCALPEL sessions are not influenced by age, gender or demographic area of post-primary education.

8.2.2 Socio-economic Characteristics

This section investigates the prediction that social economic characteristics (independent variables) influence students' expressed need for demonstrator assistance during exposure to SCALPEL (dependent variable).

8.2.2.1 Operationalization of Sub Hypotheses

Research sub hypothesis 11 (see section 6.4.2.2) was operationalized to test for associations between students' expressed need for demonstrator assistance and the conceptualized dimensions of student social economic traits: previous higher education and access to computers.

H_0: medical students' expressed need for demonstrator assistance is independent of previous higher education.

H_1: medical students' expressed need for demonstrator assistance is dependent on previous higher education.

H_0: medical students' expressed need for demonstrator assistance is independent of previous access to computers.

H_1: medical students' expressed need for demonstrator assistance is dependent on previous access to computers.

8.2.2.2 Results

For each analysis, Spearman's rho highlighted very weak, non significant relationships (see Appendix 2, Table A2 - 3). Consequently, the socio-economic null hypotheses were not rejected, concluding that students' attitudes towards demonstrator assistance during

SCALPEL sessions are not biased by previous higher education or access to computers.

8.2.3 Behavioural Characteristics

This section looks at the relationship between behavioural concepts (independent variables) and students' expressed need for demonstrator assistance (dependent variable).

8.2.3.1 Operationalization of Sub Hypotheses

Research sub hypothesis 12 (see section 6.4.2.2) was operationalized to test for associations between students' expressed need for demonstrator assistance and the conceptualized dimensions of student behavioural characteristics: previous computer use and learning styles.

H_0: medical students' expressed need for demonstrator assistance is independent of previous computer use.

H_1: medical students' expressed need for demonstrator assistance is dependent on previous computer use.

H_0: medical students' expressed need for demonstrator assistance is independent of learning styles.

H_1: medical students' expressed need for demonstrator assistance is dependent on learning styles.

8.2.3.2 Results

Spearman's rho determined that a relationship did not exist between students' expressed need for demonstrator assistance and each of the variables operationalized to assess previous computer use (see Appendix 2, Table A2 - 4). Furthermore, Spearman's rho correlation coefficient for learning styles found no statistical or practical relationship (see Appendix 2, Table A2 -5). These results gave confidence to accept the null hypotheses: students' attitudes towards demonstrator assistance during SCALPEL sessions are not biased by previous computer or learning styles.

8.2.4 Opinions and Attitudes

The causality model was employed to test the prediction that student opinions and attitudes (independent variables) influence expressed need for demonstrator assistance (dependent variable).

8.2.4.1 Operationalization of Sub Hypotheses

For the purpose of this study the following student opinions and attitudes were determined to be of interest: acceptance of SCALPEL, expressed need for demonstrator assistance, and attitudes towards computer. With the relationship between students' expressed need for demonstrators and acceptance of SCALPEL being addressed in section 7.8, research sub hypothesis 13 (see section 6.4.2.2) was operationalized to test the relationship between students' expressed need for demonstrator assistance and attitudes towards computers.

H_0: medical students' expressed need for demonstrator assistance is independent of students' pre-SCALPEL attitude towards computers.

H_1: medical students' expressed need for demonstrator assistance is dependent on students' pre-SCALPEL attitude towards computers.

8.2.4.2 Results and Discussion

Spearman's rho found no relationship between students' attitudes towards demonstrator assistance during SCALPEL sessions and pre-SCALPEL attitudes towards computers (see Appendix 2, Table A2 - 6). The absence of a relationship gave confidence to accept the null hypothesis, concluding that students' expressed need for demonstrators is not influenced by pre-SCALPEL attitudes towards computers.

8.3 Attitudes Towards Computers

This section addresses the possibility that student demographic, socio-economic, behavioural, and attitudinal factors may influence student attitudes towards computers.

Because data for these concepts were collected at the same time the independent (cause) and dependent (effect) variables could not be determined by establishing time order. As a consequence causal relationships were determined by reasoning about the logical dependence for the individual relationships. Where there was a suspicion of a two-way dependence, a mutual causal relationship was assumed.

8.3.1 Student Demographics

The causality model was employed to test the prediction that various student demographics (independent variables) influence attitudes towards computers (dependent variable).

8.3.1.1 Operationalization of Sub Hypotheses

Research sub hypothesis 14 (see section 6.4.2.3) was operationalized to test for associations between student attitudes towards computers and the conceptualized dimensions of student demographics: age group, gender and demographic area of post-primary education.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of age.
H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on age.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of gender.
H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on gender.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of the demographic area of post-primary education.

H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on the demographic area of post-primary education

For each analysis, student attitudes towards computers was determined to be the dependent variable giving the impractical scenario that this concept could influence student demographics.

8.3.1.2 Results

Spearman's rho showed no relationship between students' pre-SCALPEL attitudes towards computers and age (see Appendix 2, Table A2 - 7). Given the nominal level of measurement for gender and demographic area of post-primary education, the Goodman-Kruskal Tau PRE measure of association was used. The statistics shown are asymmetric measures with student attitudes towards computers dependent (see Appendix 2, Table A2 - 8). No relationship was observed with demographic area of post-primary education. Turning to the relationship concerning gender, although the inference statistic ($p = 0.04$) indicated a significant relationship in the population, the measure of association (0.02) suggested this to be too weak to warrant any practical concern.

These results gave confidence in accepting the null hypotheses, concluding that students' attitudes computers (pre exposure to SCALPEL) are not biased by age, gender or demographic area of post-primary education.

8.3.2 Socio-economic Characteristics

The causality model was employed to test the prediction that student social economic characteristics (independent variables) influence attitudes towards computers (dependent variable).

8.3.2.1 Operationalization of Sub Hypotheses

Research sub hypothesis 15 (see section 6.4.2.3) was operationalized to test for associations between student attitudes towards computers and the conceptualized dimensions of student social economic traits: previous higher education and access to computers.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of previous higher education.

H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on previous higher education.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of previous access to computers.

H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on previous access to computers.

Previous higher education was determined to be an independent variable given the unlikely event that a student's attitude towards computers would influence previous higher education. Deciding on whether or not access to computers should constitute a dependent or independent variable was not so clear. Although reasonable to assume that exposure to computers influences attitudes towards computers, it is also just as likely that an agreeable attitude encourages more exposure. Consequently, a mutual causal relationship was reasoned between attitudes towards computers and access to computers.

8.3.2.2 Results

Spearman's rho for previous higher education highlighted a very weak, non significant relationship (see Appendix 2, Table A2 - 9). Analysis for access to computers highlighted a significant, albeit low, relationship ($\rho = 0.32$, $p < 0.0005$) between the two variables. Given that time order could not be established, a mutually causal relationship could only be assumed. It can be reasoned that accessibility to computers improves attitudes, which in turn inspires access to computers. Consequently, the null hypothesis was rejected with this in mind, concluding that a low relationship existed between medical students' pre-SCALPEL attitudes towards computers and access to computers.

8.3.3 Behavioural Characteristics

The causality model was employed to test the prediction that behaviour relevant to computers and learning (independent variables) influences attitudes towards computers (dependent variable).

8.3.3.1 Operationalization of Sub Hypotheses

Research sub hypothesis 16 (see section 6.4.2.3) was operationalized to test for associations between students' attitudes towards computers and the conceptualized dimensions of behavioural characteristics: previous computer use and learning styles.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of previous computer use.

H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on previous computer use.

H_0: medical students' pre-SCALPEL attitudes towards computers are independent of learning styles.

H_1: medical students' pre-SCALPEL attitudes towards computers are dependent on learning styles.

It was assumed that each behavioural concept was just as likely to have an effect on students' attitudes towards computers as the contrary. Computer use and learning styles may have influenced attitudes towards computers, however it may have been the case that an agreeable attitude towards computers encouraged more computer use and had an impact on learning styles. Consequently, a mutual causal relationship was assumed for each analysis.

8.3.3.2 Results

Spearman's rho correlation coefficient for previous computer use found no statistical or practical relationship for previous use of the Internet and computers at work (see Appendix 2, Table A2 - 10). However, weak significant relationships were found to exist for use of formal learning courseware ($\rho = 0.21$, $p = 0.01$) and use of computer games ($\rho = 0.24$, $p = 0.005$), and a moderate relationship was found for previous use of word processing packages ($\rho = 0.44$, $p < 0.0005$). In addition, Spearman's rho for learning styles (see Appendix 2, Table A2 - 11) highlighted a weak significant relationship ($\rho = 0.17$, $p = 0.04$), however the strength of this relationship was determined to be too weak to warrant any practical significance.

8.3.4 Opinions and Attitudes

Aside from attitudes towards computers, the other opinions and attitudes determined to be of interest to this study were acceptance of SCALPEL and expressed need for demonstrator assistance. Relationships between these concepts have been evaluated elsewhere in this thesis. Analysis of the relationship between attitudes towards computers and acceptance of SCALPEL is addressed in section 7.9, and between attitudes towards computers and the need for demonstrator assistance in section 8.2.4.

8.3.5 Changes in Attitudes Towards Computers after Exposure to SCALPEL

This section addresses the effect of exposure to SCALPEL on students' attitudes towards computers.

8.3.5.1 Operationalization of Sub Hypotheses

Research sub hypothesis 18 (see section 6.4.2.3) was operationalized to test for differences between student attitudes pre and post exposure to SCALPEL.

H_0: exposure to SCALPEL has no systematic effect on medical students' attitudes towards computers

H_1: exposure to SCALPEL has a systematic effect on medical students' attitudes towards computers

8.3.5.2 Results

Wilcoxon matched-pairs signed-ranks test for two related samples was employed to determine: of those students who express changes in attitudes towards computers after exposure to SCALPEL, do just as many change from positive to negative as change from negative to positive. Twenty-four students registered an increase in score, a negative change in attitude, while thirteen students expressed a positive change in attitude by registering a decrease in score. Ninety-four students expressed no change in attitude after using SCALPEL (see Appendix 2, Table A2 - 12). The two-tailed probability ($Z = -1.81$, $p = 0.07$, $N = 131$) on the sample indicated that the differences observed could have occurred due to sampling error when sampling from a population in which exposure to SCALPEL makes no difference in attitudes toward computers. This non-significant result gave confidence in accepting the null hypothesis, concluding that exposure to SCALPEL has no systematic effect on attitudes towards computers.

8.4 Learning Styles

This section looks at the potential influence of student demographic, socio-economic, behavioural, and attitudinal factors on student learning styles. Because data for these concepts were collected at the same time the independent (cause) and dependent (effect) variables could not be determined by establishing time order. As a consequence causal relationships were determined by reasoning about the logical dependence for the individual relationships. Where there was a suspicion of a two-way dependence, a mutual causal relationship was assumed.

8.4.1 Student Demographics

The causality model was employed to test the prediction that student demographics (independent variables) influence learning styles (dependent variable).

8.4.1.1 Operationalization of Sub Hypotheses

Research sub hypothesis 19 (see section 6.4.2.4) was operationalized to test for associations between student learning styles and the conceptualized dimensions of student demographics: age group, gender and demographic area of post-primary education.

H_0: medical students' learning styles are independent of age.
H_1: medical students' learning styles are dependent on age.

H_0: medical students' learning styles are independent of gender.
H_1: medical students' learning styles are dependent on gender.

H_0: medical students' learning styles are independent of the demographic area of post-primary education.

H_1: medical students' learning styles are dependent on the demographic area of post-primary education

It was assumed that each demographic was just as likely to have an effect on students' learning styles as the contrary. Consequently, a mutual causal relationship was inferred for each analysis.

8.4.1.2 Results

Spearman's rho showed no relationship between students' pre-SCALPEL learning styles and age (see Appendix 2, Table A2 -13). Given the nominal level of measurement for gender and demographic area of post-primary education, the Goodman-Kruskal Tau PRE measure of association was used. The statistics shown (see Appendix 2, Table A2 - 14) are asymmetric measures with student learning styles dependent. The measure of association for each of these demographics showed no relationship with learning styles. As a consequence of these results the null hypotheses were accepted leading to the conclusion that students' age, gender and demographic area of post-primary education does not influence learning styles.

8.4.2 Socio-economic characteristics

The causality model was employed to test the prediction that social economic characteristics (independent variables) influence student learning styles (dependent variable).

8.4.2.1 Operationalization of Sub Hypotheses

Research sub hypothesis 20 (see section 6.4.2.4) was operationalized to test for associations between student learning styles and the conceptualized dimensions of student social economic traits: previous higher education and access to computers.

H_0: medical students' learning styles are independent of previous higher education.
H_1: medical students' learning styles are dependent on previous higher education.

H_0: medical students' learning styles are independent of previous access to computers.

H_1: medical students' learning styles are dependent on previous access to computers.

It was assumed that each socio-economic factor was just as likely to have an effect on a student's learning styles as the contrary. Consequently, a mutual causal relationship was inferred for each analysis.

8.4.2.2 Results

For each analysis, Spearman's rho highlighted very weak, non significant relationships (see Appendix 2, Table A2 - 15). Consequently, the socio-economic null hypotheses were not rejected, concluding that students' learning styles are not biased by previous higher education or access to computers.

8.4.3 Behavioural Characteristics

The causality model was employed to test the prediction that behaviour relevant to computers and learning (independent variables) influences learning styles (dependent variable).

8.4.3.1 Operationalization of Sub Hypotheses

Research sub hypothesis 21 (see section 6.4.2.4) was operationalized to look at the association between the two conceptualized dimensions of behavioural characteristics: previous computer use and learning styles.

H_0: no association exists between student learning styles and previous use of computers
H_1: an association exists between student learning styles and previous use of computers

It was assumed that previous use of computers was just as likely to have an effect on students learning styles as the contrary. Consequently, a mutual causal relationship was inferred for this analysis.

8.4.3.2 Results

Spearman's rho correlation coefficients found no relationship between learning styles and previous computer use for formal learning, games, word processing and work (see Appendix 2, Table A2 - 16). However, a significantly weak negative relationship ($\rho = -0.17$, $p = 0.04$) with previous use of the Internet was observed but was considered too weak to be of any practical significance.

These results gave confidence in accepting the null hypotheses and concluding that students' pre-SCALPEL learning styles are not influenced by previous computer use.

8.4.4 Opinions and Attitudes

For the purpose of this study the following student opinions and attitudes were determined to be of interest: acceptance of SCALPEL, expressed need for demonstrator assistance, and attitudes towards computers. Relationships between learning styles and these concepts have been evaluated elsewhere in this thesis. Analysis of the relationship between learning styles and acceptance of SCALPEL is addressed in section 7.10, between learning styles and the need for demonstrator assistance in section 8.2.3, and between learning styles and attitudes towards computers in section 8.3.3.

8.4.5 Changes in Learning Styles after Exposure to SCALPEL

This section addresses the effect of exposure to SCALPEL on student learning styles.

8.4.5.1 Operationalization of Sub Hypotheses

Research sub hypothesis 23 (see section 6.4.2.4) was operationalized to test for differences between student learning styles pre and post exposure to SCALPEL.

H_0: exposure to SCALPEL has no systematic effect on medical students' learning styles
H_1: exposure to SCALPEL has a systematic effect on medical students' learning styles

8.4.5.2 Results

Wilcoxon matched-pairs signed-ranks test for two related samples was employed to determine: of those students who express changes in learning styles after exposure to SCALPEL, do just as many change from exploratory to passive as changed from passive to exploratory. Nine students registered an increase in-score, reflecting a change towards a more passive learning style. Twenty-four students recorded a decrease in score, exhibiting a change towards a more explorative style of learning while ninety-nine students recorded no change in learning styles after exposure to SCALPEL (see Appendix 2, Table A2 - 17). The two-tailed probability on the samples ($Z = -2.61$, $p = 0.01$, $N = 132$) indicated a significant shift in students' learning styles after exposure to SCALPEL.

This significant result gave confidence to reject the null hypothesis concluding that exposure to SCALPEL has a systematic effect on learning styles in favour of exploratory learning.

8.5 Discussion

This chapter looked at the cause and effect relationships between the conceptualized dimensions of the four types of social survey subject matter and each of the concepts identified in section 6.4.2 as being of particular interest: students' expressed need for demonstrator assistance during exposure to SCALPEL, attitudes towards computers, and learning styles.

Beginning with a look at students' expressed attitudes towards demonstrator assistance, no relationships were found with the demographic dimensions: age and gender; while a significant relationship with demographic area of post-primary area was considered to be too weak to be of any practical concern. Relationships with the socio-economic concepts of previous higher education and access to computers were also weak and non significant, as were the behavioural concepts of previous computer use and learning styles. Finally, no relationship was observed with the concept of attitudes towards computers. The absence of relationships here suggest that attitudes towards demonstrator assistance are formed during and/or after working through the practicals, and are not biased by characteristics brought by students to the pathology practicals. This inference strengthens the observed relationship between a need for demonstrators and acceptance of SCALPEL (section 7.8) by indicating that the relationship is not spurious and is as a consequence of the learning experience.

Turning to the concept of students' attitudes towards computers, no relationships were found with the demographic dimensions: age, and area of post-primary education; while a significant but weak relationship with gender was decided to be too weak to be of no practical significance. The socio-economic dimension of previous higher education was found to have no influence on attitudes towards computers. However, the relationship with access to computers was significant, with a low association found to exist in the population. A problem in establishing the time order of two concepts led to the assumption of a mutual causal relationship, inferring that accessibility to computers improves attitudes, which in turn inspires access to computers. Turning to students' behavioural characteristics, no relationships were found with two dimensions of previous computer use: the Internet and computers at work. However, weak significant relationships were found with formal learning courseware and use of computer games; and a moderate relationship was found with previous use of word processing packages. The final cause and effect relationship for the concept of attitudes towards computers was with learning styles. Although a weak significant relationship was observed, it was determined too weak to be of any practical significance. Collecting data about attitudes towards computers pre and post exposure to SCALPEL allowed the effect of SCALPEL on these attitudes to be analysed. Although differences in attitudes were observed, the inference statistic found these to be non significant, leading to the conclusion that exposure to SCALPEL had no systematic effect on attitudes towards computers.

The final primary concept addressed by this chapter was learning styles. No relationships were observed between this concept and the demographic dimensions: age, gender and demographic area of post-primary education. The same was true for the socio-economic dimensions: previous higher education and access to computers. With learning styles itself categorised as a behavioural characteristic, the next analyses looked at the relationship with previous computer use. No relationship was found between previous use of computers for formal learning, games, word processing and work. For use of the Internet, a significantly negative relationship was determined too weak to be of any practical significance. Finally, a look at the effect of SCALPEL on student learning styles indicated a systematic effect in favour of exploratory learning.

8.6 Conclusions

The analyses in this chapter concluded that the student characteristics identified to be of interest to this research had no effect on student attitudes towards the need for demonstrator assistance during SCALPEL sessions. This finding was determined to be positive to this research as it showed that the need for demonstrators is not determined by characteristics brought by students to the practical sessions. If then, such attitudes are formed as a consequence of exposure to SCALPEL, an opportunity arises to positively affect these attitudes by implementing functionality to address students' need for demonstrator assistance. The findings of the effect of student characteristics on attitudes towards computers provided evidence of a more heterogeneous population of satisfied computer users. Furthermore, the data showed that students with greater accessibility to computers were more receptive towards their use, and that those students who had more exposure to formal learning courseware, computer games, and word processing packages tended to express positive attitudes towards computers. It can be argued that these observations amount to nothing more than confirming common sense or 'conventional wisdom' (Fenton and Pfleeger 1997). However, in the context of this research, these empirical observations helped to determine the makeup of the population and establish construct validity of the constituent measures (see section 5.4.6.2). The final observation made in this chapter was that exposure to SCALPEL had a positive effect on students' learning styles: more students indicated a shift towards a more exploratory learning style than from an exploratory to a passive learning style.

8.7 Summary

This chapter looked at the cause and effect relationships between the conceptualized dimensions of the four categories of social survey subject matter and each of the primary concepts identified as being of particular interest to this research. Finally a discussion was offered on the interpretation of the findings and their impact on the research focus. Influenced by, and subsequently designed to complement the evaluation findings of this and the previous chapter, the next chapter presents the second empirical evaluation of SCALPEL.

Chapter 9: The Second Empirical Study

9.1 Introduction

The previous two chapters introduced the evaluation findings of the first empirical evaluation of SCALPEL. This chapter continues with the evaluation theme by presenting a second empirical study. The sections of this chapter follow closely the structure of chapters 6 and 7. Section 9.2 starts the evaluation process (see section 5.1) by introducing a framework influenced by, and subsequently designed to complement, the first empirical evaluation. Section 9.3 discusses the conceptualization and operationalization of the concepts underpinning this phase of the research. Some were brought forward from the previous evaluation intact, while others were revised and new ones introduced to satisfy the pre-specified and curiosity driven research agendas. Sub hypotheses that address the pre-specified research agenda are introduced in section 9.4 while section 9.5 considers their operational definitions. Next, the evaluation methodology specific to this phase of the research is considered in section 9.6 together with the way in which the measurement instrument was administered. Section 9.7 addresses the reliability and validity of the constituent measures of this instrument. The data analysis begins in section 9.8 with a look at the demographic makeup of the sample population. Section 9.9 takes the first look at the pre-specified research agenda and the concept of student attitudes towards SCALPEL. Section 9.10 continues by assessing students' expressed need for demonstrator availability during SCALPEL sessions. The pre-specified agenda concludes with section 9.11 assessing the relationship between these two concepts. Section 9.12 begins to address the curiosity driven agenda by presenting the factual data on reasons prompting students to ask for demonstrator assistance during SCALPEL sessions. Continuing with the focus on problems encountered by students, section 9.13 looks at the methods used by students to solve problems in the absence of demonstrators, and section 9.14 discusses the importance students placed on the proposed improvements to SCALPEL. Finally, the work presented in this chapter is discussed in chapter 9.15 and the significant findings highlighted in section 9.16.

9.2 Defining the Evaluation Framework

In determining the evaluation framework for this phase of the research two significant differences from the first study needed to be taken into consideration. First, the series of pathology practicals under evaluation (the same six practicals used in the first study, see section 7.2) were to be taught entirely using SCALPEL. Second, demonstrators would not be available during the SCALPEL sessions.

9.2.1 Pre-Specified Research

The empirical evaluation presented in chapter 7 satisfied the 'education attainment' facet of the primary research hypothesis (see section 6.2) by concluding that pathology practicals are equally educationally effective, in terms of domain material, when delivered in the laboratory or by SCALPEL. Consideration was given to the benefits of replicating (section 5.8) the evaluation performed in chapter 7. Unfortunately, practical constraints of executing a controlled experiment with a large cohort of students in a live education environment, coupled with the decision not to use the traditional laboratory environment, made replication of the education attainment results unfeasible. Having satisfied the education attainment research objective, further investigations into this facet of the first evaluation's primary research hypothesis were considered unnecessary. However the negative attitudes expressed by students towards SCALPEL, the second concept of the primary research hypothesis put forward in section 6.2.1, required further investigation (see section 7.6). In addition, the curiosity driven research raised concerns about the distribution of students' expressed need for demonstrator assistance during SCALPEL sessions (see section 7.8.1). It was the observed relationship between these two concepts (see sections 7.8.2 and 7.17) that influenced the pre-specified agenda for this evaluation phase, with a primary research hypothesis authored to address the relationship as follows:

> students who express a need for demonstrator assistance during SCALPEL sessions express negative attitudes towards SCALPEL relative to those who express less of a need for demonstrators

Therefore, the primary objective of this summative evaluation was the success of SCALPEL in terms of students' attitudes towards using SCALPEL to deliver pathology practicals and the role of demonstrators.

9.2.2 Curiosity Driven Research

The curiosity driven research agenda was driven by the continued need to improve the SCALPEL courseware.

Given that demonstrators were unavailable for future SCALPEL sessions, their role during the pathology practical sessions needed to be addressed. In addition to the attitudinal data on the role of demonstrators during SCALPEL sessions, a descriptive (see section 5.4.1.2) understanding of their role was necessary to satisfy two objectives. First, factual data might highlight reasons for the distribution of attitudes towards demonstrators, acceptance of SCALPEL, and any relationship observed between these two variables. Second, the data would be used to inform the research and development of improved SCALPEL functionality, i.e. methods that would replicate the traditional role of demonstrator and peer dialogue during pathology practicals. Factual data were therefore required about the reasons prompting students to call upon demonstrators. In addition it was determined that data explaining the frequency of students requiring demonstrators, and the importance placed on demonstrators to be available to address any given query would offer valuable viewpoints. Finally, given the absence of demonstrators, an understanding of the methods used by students to solve problems during SCALPEL sessions was considered necessary.

The final objective of this research agenda was to have an understanding of the ways in which students believed SCALPEL could be improved. It was considered that this data could inform the research and development of enhanced SCALPEL functionality by providing an insight into what the students wanted most.

Hypotheses were not generated to satisfy these descriptive investigations as the statements about the research objectives were considered sufficient in describing the issues under investigation (see section 5.3).

9.3 Conceptualization and Operationalization

In keeping with the research methodology employed by the first study (see chapters 5 and 6), this section offers conceptual and operational definitions of the concepts addressed during this phase of the research. Some of the measures operationalized in chapter 6 are employed here with no changes, while others have been revised and new ones introduced to reflect the shift in research focus. All measures for this evaluation phase were operationalized on a questionnaire administered to students after using SCALPEL for six practicals.

9.3.1 Demographic Characteristics

The concepts 'age' and 'gender' were employed as part of this phase of the research to offer an insight into the demographic makeup of the sample population. No changes were made to the operational definitions introduced in section 6.3.2.

9.3.2 Opinions and Attitudes

Students' opinions and attitudes determined to be of interest to this phase of the research were 'acceptance of SCALPEL' and 'attitudes towards the need for demonstrators'.

No alterations were made to the operational definition of 'attitudes towards demonstrators' (see section 6.3.5). However, the concept 'acceptance of SCALPEL' was revised due to a shift in research focus. The majority of indicators that formed this measure in the first study considered both the traditional laboratory and SCALPEL environments. With the traditional laboratory no longer a consideration, indicators referring to this method of teaching were removed, leaving those considering only SCALPEL. Attitudes towards SCALPEL were therefore measured by using a multiple-item measure consisting of four indicators. Recoding of negatively directed indicators was carried out for indicator 2. A low score on the scale represents a positive attitude towards SCALPEL.

Indicator 1: I find the computer application a satisfying learning activity (+)

Indicator 2: I felt lost and confused using the computer application (-)

Indicator 3: I like the consistency of explanations provided by the computer application (+)

Indicator 4: I like the environment the computer application provides (+)

9.3.3 Factual Data

Given the difficulties associated with coding qualitative answers into categories (see section 5.4.1.2), open questions were avoided in favour of pre-coded questions for the capture of factual data during this study.

9.3.3.1 Modelling the Role of Demonstrators

For an understanding of the role of demonstrators during SCALPEL sessions, three entities were determined to be particularly important: reasons for demonstrator assistance, frequency of reasons, and importance to student learning. These entities were conceptualized respectively as students' reasons for requiring demonstrator assistance, frequency demonstrator assistance for each reason, and the importance placed by students on the reasons for demonstrator availability.

Addressing the reasons for students requiring demonstrator assistance, it was determined that a limited but well established list of alternative categories was required to represent the questions typically directed at demonstrators during pathology practicals in the traditional laboratory and SCALPEL sessions. Qualitative (participant observation and focus group) data from the first study (see chapter 7) highlighted categories representative of issues pertaining to the pathology material (RD1, RD3, RD4, RD5 and RD6) or the operation of the SCALPEL courseware (RD2, RD7 and RD8).

RD1 - Discuss pathology material being taught other than Histology
RD2 - Explain how to use the multiple choice window
RD3 - Discuss Histology
RD4 - Discuss background information about any aspect of the material being taught
RD5 - Discuss how the material being taught relates to other areas of medicine
RD6 - Clarify aims and objectives of the practical session

RD7 - Explain how to start the application

RD8 - Explain how to use the application features (i.e. video, image zoom)

These categories were used to operationally realise the concepts 'frequency of reasons' and 'importance to student learning'. Students were asked to indicate how many times on average in each practical they would like to have consulted a demonstrator, and to rank using a scale of 1 (important) to 10 (least important) the importance they placed on demonstrators to be available for each of the categories. Students were also asked to only rank those categories for which they had indicated the need for demonstrator assistance. Acknowledging that the list of pre-coded alternatives would not be exhaustive, students were also given the opportunity to record reasons for requiring demonstrators not covered by the categories operationalized above.

9.3.3.2 Solving Problems in the Absence of Demonstrators

To model how students solved problems in the absence of demonstrators, a conceptual definition was put forward as 'methods used by students to solve problems during practicals'. Using the list of common reasons for requiring demonstrator assistance (see section 9.3.3.1), a measure was operationalized asking students to indicate the methods used to solve the problems encountered during the pathology practicals. Qualitative data from the first study (see chapter 7) were used to collate the following list of methods commonly used by students to solve problems encountered during pathology practicals:

a) text books

b) help provided by the pathology application

c) asked in tutorial

d) asked a member of staff outside of tutorial

e) asked another student

f) never found the answer

g) read lecture notes / handouts

h) found the answer in the application

An introduction to this measure informed students that each solution could be used more than once, and that more than one solution could be used for each reason. Opportunity was given for students to indicate 'reasons for requiring demonstrators' and 'solutions used to solve problems' not covered by the pre-specified lists.

9.3.3.3 Looking at Ways to Improve SCALPEL

To assess what students considered to be good improvements to SCALPEL, a measure was conceptually defined as 'ways in which students believe SCALPEL could be improved'. This measure was operationalized by asking students to rank on a scale of 1 (most important) to 10 (least important) the importance for each of the following proposed improvements in terms of what they would most like to see incorporated into the SCALPEL courseware.

- Glossary of terms
- On-line text books
- Ability to ask questions and receive answers about pathology material
- On-line notepad
- More links to background information about pathology material
- Optional practicals
- More animation like the ECG
- Introduction of sound to listen to patient history etc.

Students were also given the opportunity to include, and rank, proposed improvements not covered by the pre-specified list.

9.4 Research Sub Hypotheses

Consideration was given to ensure that the sub hypotheses derived for this evaluation phase satisfied conditions introduced in section 5.3.

9.4.1 Pre-Specified Research Agenda

Three research sub hypotheses were formulated to address student attitudes towards SCALPEL and the need for demonstrator assistance. Sub hypothesis 2 from the first study (see section 6.4.1) was revised to address student attitudes towards SCALPEL.

Sub Hypothesis 1 - Medical students find SCALPEL an acceptable learning environment.

A further sub hypothesis was conceptualized to assess attitudes towards the need for demonstrator assistance during SCALPEL sessions.

Sub Hypothesis 2 - Medical students rely on demonstrator assistance during SCALPEL sessions.

Finally, sub hypothesis 4 from the first study (see section 6.4.2) was revised to reflect the prediction that students who express a need for demonstrator assistance during SCALPEL sessions are more negative towards SCALPEL than those who have a lesser need for demonstrators.

Sub Hypothesis 3 - Medical students' need for demonstrator assistance influences acceptance of SCALPEL.

9.4.2 Curiosity Driven Research

Hypotheses were not generated to satisfy the descriptive investigations that make up the curiosity driven research agenda. This decision was made on the basis that the statements about the research objectives were considered sufficient in describing the issues under investigation (see section 9.2.2).

9.5 Operational Statements

For clarity, the operational statement (alternative hypothesis H_1) and null hypothesis (H_0) for each research sub hypothesis is presented in the relevant sections of this chapter that analyse the data.

9.6 Method

This evaluation phase constituted a prospective cross-sectional observational study (see section 5.6). A population of first year undergraduate medical students (N = 160) used SCALPEL to work through the same series of pathology practicals used in the first study (see section 7.2). The traditional laboratory environment was not used and demonstrators were not available for any of these practical sessions. Students were left to work through the practicals in their own time, with the only brief being that they needed to complete each practical before the relevant tutorial (see section 4.2.1).

9.6.1 Research Techniques

A self-administered questionnaire was given to the students at the start of the last lecture of the pathology course and collected after the lecture (see Appendix 10 for presentation of questionnaire). Students who wished to take the questionnaire away for completion at a later date were given the opportunity to do so and were instructed to return them to the pathology office or their pathology tutor. The anonymity of respondents was again promoted in an effort to encourage responses (see section 5.4.1.1). Of the 160 students in the sample population, 97 students returned completed questionnaires leaving 63 non-responses. It is believed that the high non-response rate was due to a mixture of refusals and non-contacts: students had no incentive to return a completed questionnaire and not all the students were present during the last lecture. With the population of interest being undergraduate medical students, confidence can be given that the sample of first year undergraduate medical students used for the statistical analyses presented in this study are representative of the population of interest. However, caution is advised when generalising the inferences to a wider population.

9.7 Reliability and Validity of Measures

This section addresses the reliability and validity (section 5.4.6) of the measures used during this phase of the research.

9.7.1 Reliability

Cronbach's alpha was used to test the internal reliability (see section 5.4.6.1) of each multiple-item measure employed as part of this study. All other measures were determined to be reliable through establishing face validity (see section 9.7.2).

9.7.1.1 Behavioural Characteristics

Attitudes towards SCALPEL

The alpha coefficient for the multiple-item measure that addressed students' attitudes towards SCALPEL (Alpha = 0.73, N of cases = 95, N of items = 4) suggested a moderate to high level of association between the constituent indicators, giving confidence in the internal reliability of the measure.

The Need for Demonstrators

The alpha coefficient for the multiple-item measure (Alpha = 0.81, N of cases = 93, N of items = 4) authored to address students' need for demonstrators suggested a moderate to high level of association between the constituent items, giving confidence in the internal reliability of this measure.

9.7.2 Validity

All the measures used in this evaluation were determined to have face validity (see section 5.4.6.2). Concurrent validity could not be established given the difficulties associated with obtaining independent sources of information. For some measures construct validity was inferred by definition of their observed relationship with other measures.

9.8 Sample Frame Demographics

Introduced in section 9.6, a population of first year undergraduate medical students (N = 160) was used for this study. Demographic data were collected on a questionnaire administered to the population during the last lecture of the pathology course. With 63 non-responses, the sample frame consisted of 97 students. Of these 41 (42.3%) were male and 56 (57.7%) female. The majority of the students (70.1%) were under the age of 20. 25.8% were aged between 20 and 25, with the remaining 4.1% aged between 26 and 35 (see Appendix 3, Table A3 - 1).

9.9 Acceptance of SCALPEL

This section looks at the distribution of students' acceptance of SCALPEL.

9.9.1 Operationalization of Sub Hypothesis

Sub hypothesis 1 (see section 9.4) was operationalized to address students' acceptance of SCALPEL.

H_0: The distribution of student attitudes towards acceptance of SCALPEL is normal.

H_1: The distribution of student attitudes towards acceptance of SCALPEL is skewed in favour of positive attitudes.

9.9.2 The Distribution of Student Attitudes Towards SCALPEL

The distribution of the sample for the concept 'acceptance of SCALPEL' was taken as an indication of how well SCALPEL was received by the students (see Appendix 3, Table A3 - 2). In addition to presenting graphically the distribution for this study, Figure 21 offers the distribution of attitudes from the previous study (see section 7.6). A positive difference in attitudes was observed with this sample. 47.4% (previous sample 23.4%) of students expressed a positive attitude towards SCALPEL (strongly agree and agree categories) while only 8.4% (previous sample 26.2%) exhibited negative attitudes. The remaining 44.2% of students (50.4% in the first study) fell into the 'neutral' category on the scale. The distribution is clearly skewed in favour of more positive attitudes towards SCALPEL, giving confidence in rejecting the null hypothesis of a normal distribution.

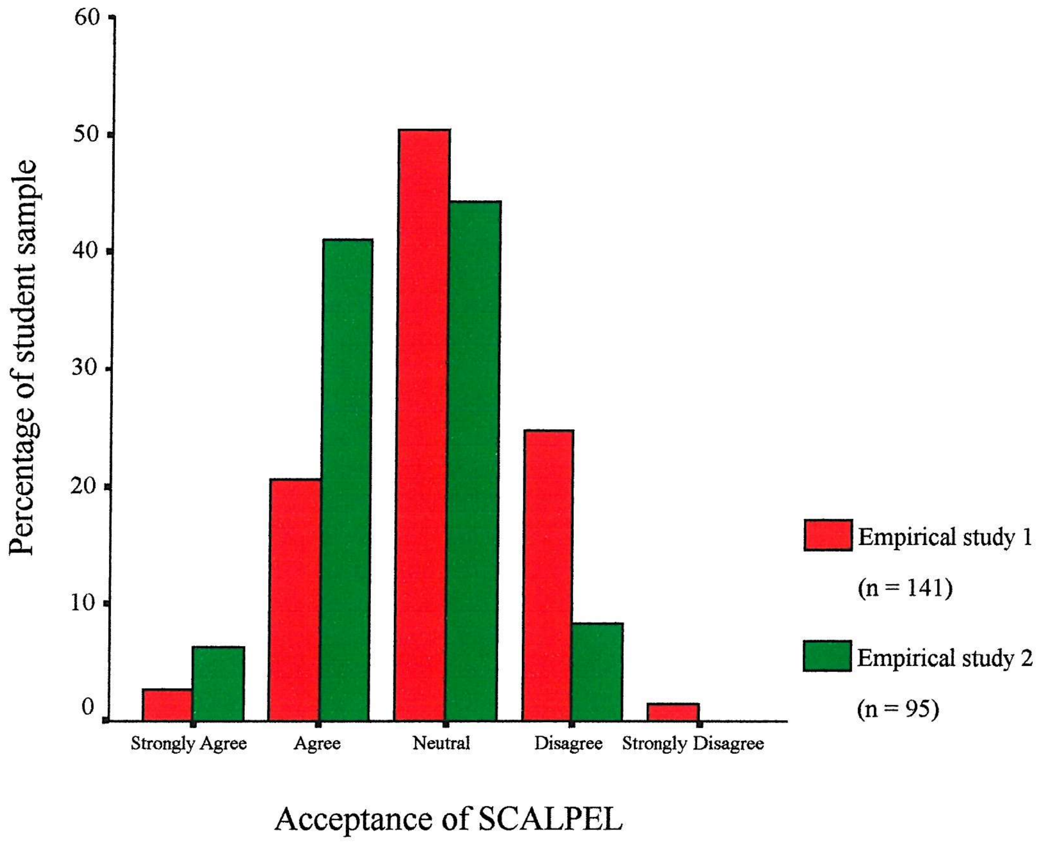


Figure 21 Bar chart on the distribution of students by acceptance of SCALPEL for empirical studies 1 and 2.

9.10 Assessing the Need for Demonstrator Assistance

This section looks at the distribution of students' expressed need for demonstrator assistance during SCALPEL sessions.

9.10.1 Operationalization of Sub Hypothesis

Sub hypothesis 2 (see section 9.4) was operationalized to address students' attitudes towards demonstrators during SCALPEL sessions.

H_0: The distribution of student attitudes towards the need for demonstrators is normal.

H_1: The distribution of student attitudes towards the need for demonstrators is positively skewed.

9.10.2 The Distribution of Student Attitudes Towards the need for Demonstrators

The distribution of the sample for students' expressed need for demonstrators during SCALPEL sessions indicated a continuing need for demonstrator assistance during SCALPEL sessions (see Appendix 3, Table A3 - 3).

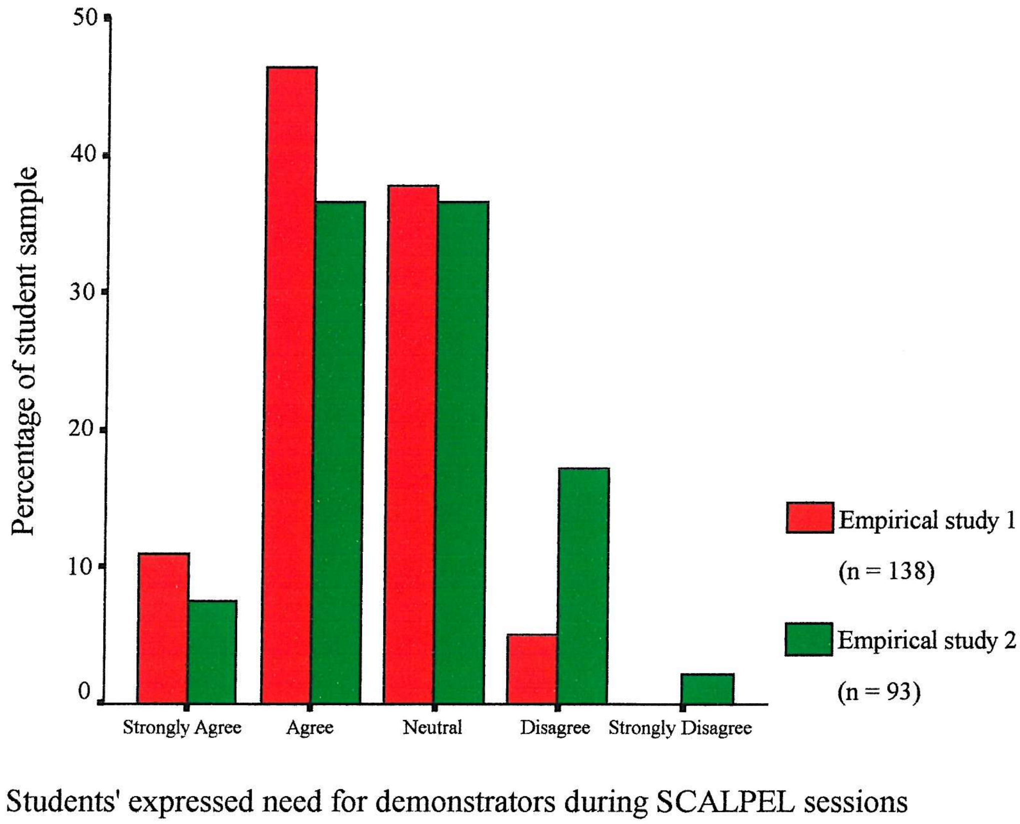


Figure 22 Bar chart on the distribution of students by expressed need for demonstrators during SCALPEL sessions for empirical studies 1 and 2.

Figure 22 presents the distribution of attitudes towards demonstrators for this and the previous study (see section 7.8). Although the distribution gives confidence in rejecting the null hypothesis, an improved difference in student attitudes was observed with the distribution of the first empirical study. 41.1% (previous sample 57.2%) of the sample expressed a need for demonstrator assistance during SCALPEL sessions (strongly agree and agree categories). 19.4% of students (previous sample 5.1%) exhibited less of a need for demonstrators, with the remaining (36.6% this time, 37.7% in the first study) being measured in the 'neutral' category of the scale.

9.11 Acceptance of SCALPEL and Demonstrator Assistance

This section addresses the effect of the need for demonstrator assistance on students' acceptance of SCALPEL.

9.11.1 Operationalization of Sub Hypotheses

Research sub hypothesis 3 (see section 9.4) was operationalized to test for associations between the need for demonstrator assistance and acceptance of SCALPEL as follows.

H_0: expressed need for demonstrator assistance is independent of a student's acceptance of SCALPEL.

H_1: expressed need for demonstrator assistance is dependent on a student's acceptance of SCALPEL.

9.11.2 Results

Because acceptance of SCALPEL was measured at the same time as expressed need for demonstrators, the independent (cause) and dependent (effect) variables could not be established by the time order of the two concepts. Theoretically the causal relationship between the two concepts can be reasoned either way, inferring that each concept is just as likely to have an effect on the other.

Table 2 presents Spearman's rho coefficients for acceptance of SCALPEL and the need for demonstrator assistance. A significant moderately negative relationship ($\rho = -0.42$, $p < 0.0005$) was found to exist. The null hypothesis was therefore rejected, inferring that: students who expressed a need for demonstrators to be available during SCALPEL sessions indicated a lower acceptance of SCALPEL than those who suggested a lesser need for demonstrators. This inference confirms and strengthens the results from the first empirical study (see section 7.8.2.2) where an identical observation was made.

Table 2 Measure of association between students' expressed need for demonstrator assistance and acceptance of SCALPEL

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
acceptmi	93	-0.42**	0.000

** Correlation is significant at the 0.01 level (2-tailed)

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered during the last lecture of the pathology course. Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on the same questionnaire. The sample (N) was made up of students who expressed an attitude towards the acceptance of SCALPEL and the need for demonstrator assistance. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

9.12 Modelling the Role of Demonstrators

This section looks at the role of demonstrators, as seen by students, during pathology practicals. Of the 97 students in the sample frame, 66 indicated on returned questionnaires the frequency of reasons for requiring demonstrator assistance. Although given the opportunity to do so, none of the students indicated reasons not covered by the categories operationalized in section 9.3.3.1. It is acknowledged that students omitting to indicate additional reasons may not have been due to an exhaustive list of alternatives being presented, but more probably because of a bias inherent in pre-coded questions (see section 5.4.1.2): students had little motivation to reach a considered view by reasoning about what was being asked and simply chose from the alternatives offered.

The data showed that if available, the population of 66 students would have required 557 demonstrator consults, an average of 8 per student (see Appendix 3, Table A3 - 4). Presented graphically in Figure 23, 94.4% of reasons for demonstrator assistance were driven by the need to discuss with demonstrators the pathology material (RD3, RD1, RD4, RD5 and RD6). The remaining 5.6% (RD2, RD7 and RD8) of required demonstrator consults were to address problems with the operation of the SCALPEL courseware. This data suggests that the majority of students' reasons for requiring demonstrator assistance are because of misconceptions with the pathology material, and were not as a consequence of failings with the operability of the SCALPEL courseware.

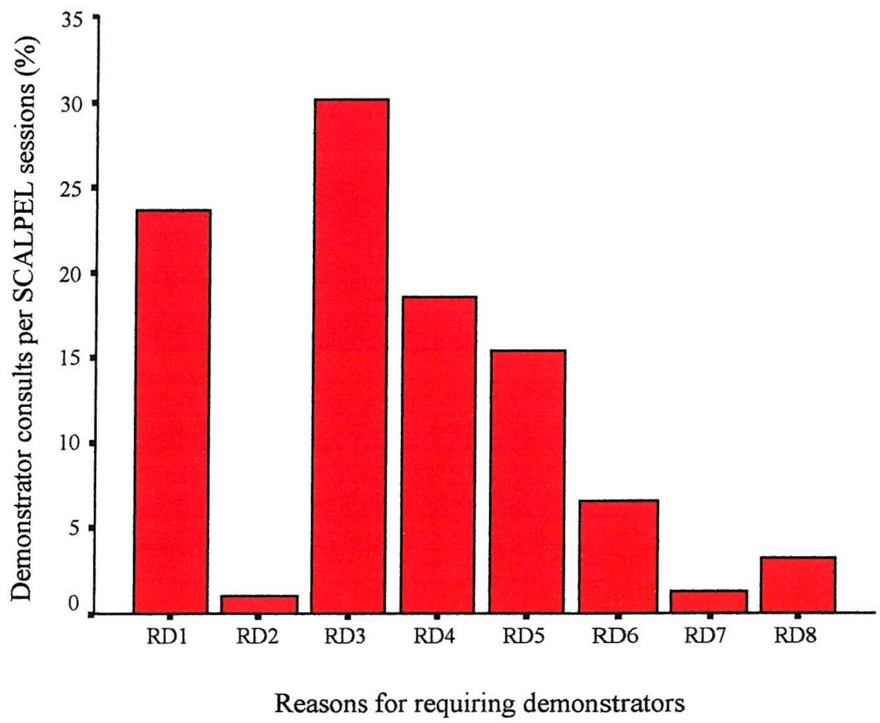


Figure 23 Bar chart on the mean distribution of required demonstrator consults during a SCALPEL session.

This observation was supported by data collected on the importance students placed on demonstrators to be available. The rounded mean score of the sum of all ranks was taken to reflect on a scale of 1 (most important) to 10 (least important) a category's position relative to the other categories (see Appendix 3, Table A3 - 5). Ordering the categories by their mean importance ranked those representing domain material (RD3, RD1, RD4 and RD5) on the lower portion of the scale (most important), and those representing issues regarding application functionality (RD8, RD2 and RD7) at the higher end (least important).

9.13 Solving Problems in the Absence of Demonstrators

The previous section showed that the sample of 66 students would like to have consulted demonstrators an average of 557 times during a SCALPEL session and that 94.4% of these consults would have been to discuss the domain material. This section looks at the methods used by students to solve problems in the absence of demonstrators. An understanding of this formative data was considered important to the research and development of improvements to SCALPEL in that methods used by students to solve problems could be automated or augmented by the SCALPEL courseware. None of the students offered additional methods to the alternatives categorised in section 9.3.3.2.

Of the 97 students in the sample frame, 80 indicated on returned questionnaires the methods used to solve problems in the absence of demonstrators. Figure 24 presents the number of students who solved problems during SCALPEL sessions by using the methods identified in section 9.3.3.2. Spikes in the graph indicate that text books, tutorial discussions and peer dialogue were the more favourable methods employed by students. Furthermore, it is clear that pathology problems (RD1, RD3, RD4 and RD5) were mainly solved by these methods, while problems with SCALPEL functionality (RD2, RD7 and RD8) were addressed by the SCALPEL courseware and peer dialogue.

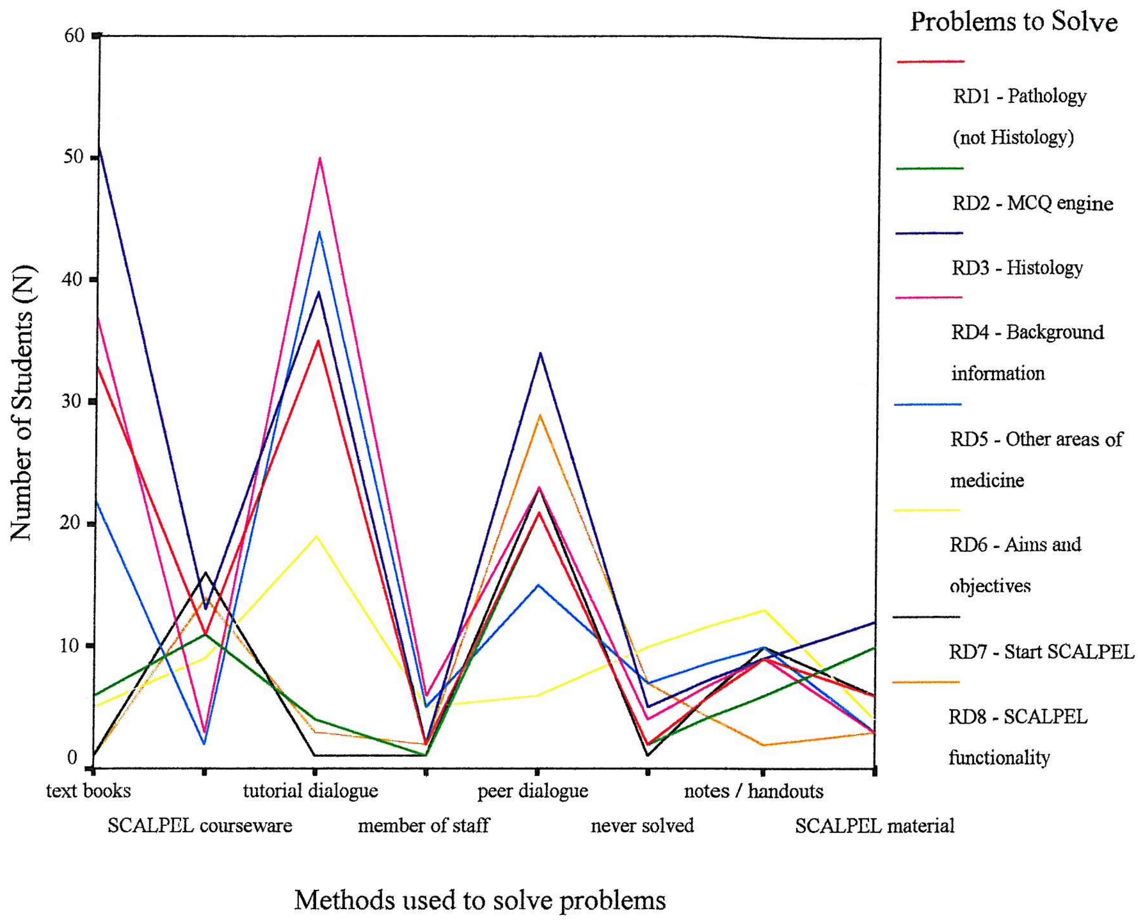


Figure 24 Number of students who used methods to solve problems encountered during SCALPEL sessions (for data see Appendix 3, Table A3 - 6).

An encouraging observation from these data was that very few student problems were never solved. Of most concern was that 15.2% of the student sample frame (N = 66) on at least one occasion experienced trouble in finding help to address the aims and objectives of the course (RD6). Focusing on problems with the domain material, the biggest failing was that 10.6% of students expressed that on at least one occasion their need to discuss how the material relates to other areas of medicine (RD5) was not satisfied. In addition, 7.6% of the students did not solve problems with Histology (RD3), followed by 6.1% who needed to discuss background information relative to the domain material (RD4) and 3% who never solved misconceptions about the pathology material other than Histology (RD1). Turning to problems relative to the SCALPEL functionality, 10.6% of students never solved problems encountered with the courseware feature, i.e. video, image zoom etc. (RD8), 3% never managed to solve problems with the MCQ engine (RD2), and 1.5% could not solve problems with starting the application (RD7).

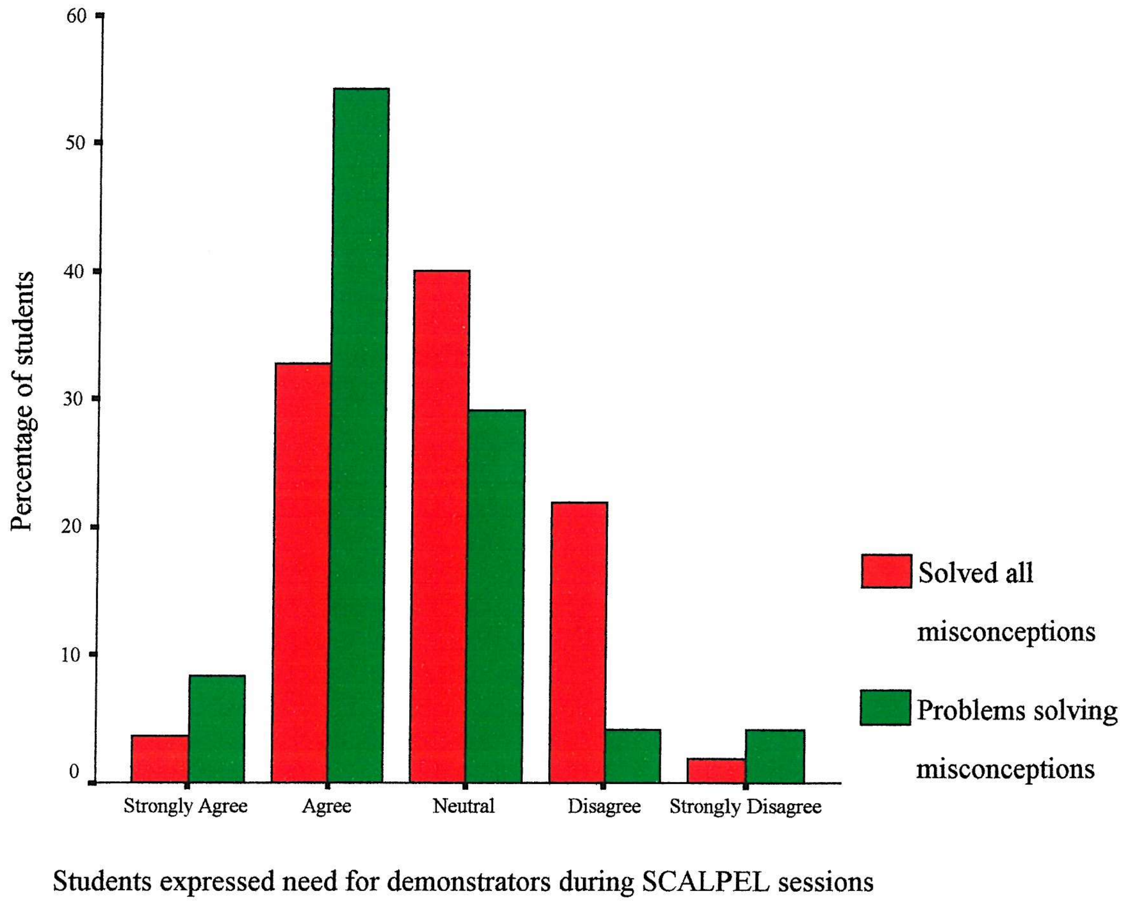


Figure 25 Bar chart on the distribution of students by expressed need for demonstrators during SCALPEL sessions for those students who had difficulty in solving misconceptions and those who did not.

Reasoning that students who on at least on occasion never found the answer to a problem experienced during SCALPEL sessions would exhibit more of a need for demonstrators than those who solved all problems, Figure 25 looks at the distribution of attitudes towards demonstrators with a focus on this nominal categorisation. It can be seen that those students who solved all encountered problems expressed less of a need for demonstrators than those who never solved at least one problem. This observation was enforced by the Mann-Whitney U test for unrelated samples which found a significant difference ($U = 464.5$, $Z = -2.21$, $p = 0.03$) between the mean ranking of attitudes towards demonstrators for these two cohorts of students.

A significant difference ($U = 458.5$, $Z = -2.62$, $p = 0.009$) in the distribution of student attitudes towards SCALPEL was also found between students who solved all encountered problems and those who did not. Figure 26 delineates this difference as students who solved problems exhibiting a higher degree of acceptance of SCALPEL than those who experienced solving misconceptions with the pathology practicals.

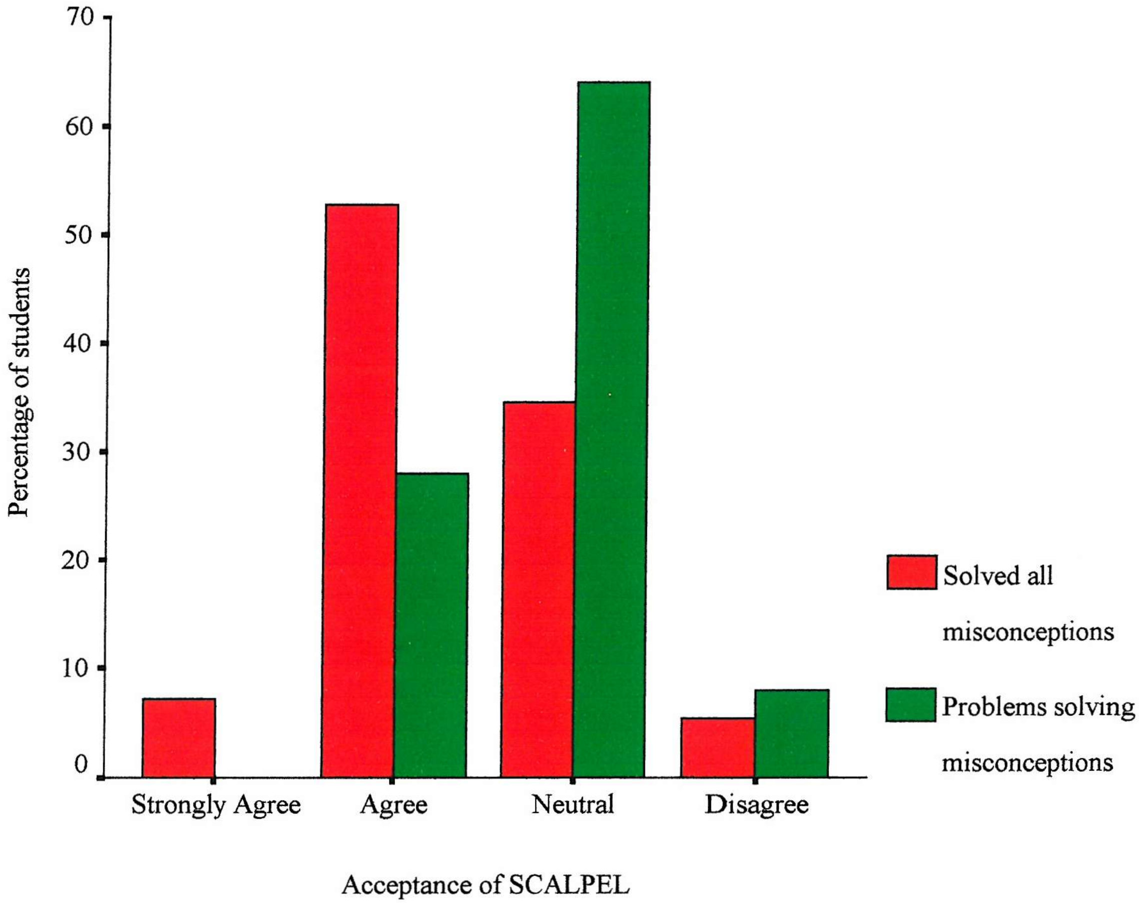


Figure 26 Bar chart on the distribution of students by acceptance of SCALPEL for those students who had difficulty in solving misconceptions and those who did not.

9.14 Possible Application Improvements - what the students want

This section looks at what students believed to be the most important of the proposed improvements to the SCALPEL courseware.

Although given the opportunity to do so, none of the students indicated improvements to the courseware not covered by the list of alternatives operationalized in section 9.3.3.3. The rounded mean score of the sum of all ranks was taken to reflect on a scale of 1 (most important) to 10 (least important) the importance of an improvement relative to the other proposals (see Appendix 3, Table A3 - 7). Functionality offering on-line text books, the ability to ask questions and receive answers about pathology material, a glossary of terms, and more links to background information about the pathology material were considered by students to be the most important improvements to SCALPEL. Consistent with the patterns observed in the data describing the role of demonstrators (section 9.12) and the methods used to solve problems in the absence of demonstrators (section 9.13), students placed more importance on including functionality designed to address the pathology material and not SCALPEL functionality. This finding was encouraging in that it demonstrated that students were not biased by the novelty of the 'bright lights' of the courseware and concentrated on the serious pedagogic value of the experience.

9.15 Discussion

The empirical evaluation presented in this chapter builds on the findings of the first evaluation presented in chapters 7 and 8. Beginning with a look at student attitudes towards SCALPEL, a positive difference was observed between the attitudes exhibited in this and the previous evaluations. 47.4% (previous sample 23.4%) of students expressed a positive attitude towards SCALPEL while only 8.4% (previous sample 26.2%) exhibited negative attitudes. The remaining 44.2% of students (50.4% in the first study) fell into the 'neutral' category on the scale. Considering that demonstrators were not available, and the previous evaluation's inference that students who express a need for demonstrators are more negative towards SCALPEL that those who have less of a need (see sections 7.8.2 and 7.17), attitudes presented in this chapter were expected to be negative relative to the previous evaluation. A look at the distribution of attitudes towards demonstrators, and the relationship between these attitudinal concepts, was required to reason about this unexpected positive shift in acceptance of SCALPEL. First, fewer students in this evaluation (41.1%) expressed a need for demonstrators relative to the previous evaluation (57.2%). Having observed no significant change in the neutral category (36.6% this time, 37.7% in the previous study), this difference is reconciled by more students (19.4%) expressing less of a need for demonstrators than previously (5.1%). Second, the relationship between the two concepts ($\rho = -0.42$, $p < 0.0005$) was found to be consistent with what was observed in the previous evaluation ($\rho = -0.40$, $p < 0.0005$). The distribution of attitudes towards demonstrators, and the relationship with students' acceptance of SCALPEL explains the improved attitudes towards SCALPEL. However, since the same courseware and delivery platform were used during both evaluations, the improved attitudes towards the need for demonstrator assistance relative to the previous evaluation was not so obvious.

One explanation for students exhibiting less of a need for demonstrator assistance during this study is that the results of the first study were biased by the subconscious effects of knowing that the traditional laboratory, SCALPEL, and the role of demonstrators were under evaluation. This bias is documented by Altman (1991) who recommends employing a procedure called blinding, a technique where the variables under investigation are unknown to the respondents. Unfortunately blinding could not be applied to this research since it was obvious to the respondents that SCALPEL and the role of demonstrators were being assessed. Another explanation is that attitudes of students in the first evaluation may have been biased by the availability of demonstrators. It may have been easier for these students to call for demonstrator assistance just because they were available, rather than attempt to solve their misconceptions by other available means. It can be reasoned that these students formulated their opinions towards demonstrators having received valuable assistance during the SCALPEL sessions, and faced with the prospect of demonstrators being withdrawn from future practical sessions, it is reasonable to expect their attitudes to be biased accordingly. Contrary to this, students in this evaluation never had the opportunity to call for demonstrator assistance at any time during exposure to SCALPEL.

prompting them to be pro active in finding solutions elsewhere. This could have influenced these students to exhibit less of a need for demonstrators since they would not miss what was never available in the first instance. If this were the case, students who at least on one occasion never solved a problem with the pathology practical would be expected to express more of a need for demonstrator assistance on the attitudinal scale. The descriptive factual data on the methods used to solve problems during SCALPEL sessions provided empirical evidence to support this assumption. A look at the distribution of students' expressed need for demonstrator assistance for those students who solved all problems and those who did not showed that the later cohort expressed more of a need for demonstrator assistance. This observation was underpinned by the significant difference ($U = 464.5$, $Z = -2.21$, $p = 0.03$) found by the Mann-Whitney *U* test for two unrelated samples. In addition, a significant difference ($U = 458.5$, $Z = -2.62$, $p = 0.009$) was found in students' acceptance of SCALPEL for these two groups of students, with students who managed to solve problems expressing a higher degree of acceptance of SCALPEL.

The descriptive factual data on the role of demonstrators during SCALPEL sessions indicated that the 66 students in the sample frame would liked to have consulted demonstrators an average of 557 time during a single SCALPEL session (an average of 8 consults per student). An observation brought to light by this quantitative data was that the majority (94.4%) of consults were driven by a need to discuss the pathology material, and not to address problems with the SCALPEL courseware. This was encouraging as it offered evidence in support of SCALPEL as a platform to deliver pathology practicals: if students were experiencing problems with SCALPEL functionality, a higher proportion of problems in the categories operationalized to address such problems would be expected. This observation was supported by students placing more importance on demonstrators to be available to discuss problems with the domain material than to address issues pertaining to the functionality of SCALPEL. Coupled with the observation that the majority of all problems with the domain material were solved by students being pro active in securing alternative solutions, it can be reasoned that students coped well in working through the pathology practicals without demonstrator assistance. However, the data also emphasised the relatively ineffective help offered by the SCALPEL courseware in solving problems compared to the more favourable methods of text books, tutorial and peer dialogue.

Concluding that although students continued to express a need for demonstrator assistance during exposure to SCALPEL, relatively few problems went unsolved as a consequence of students employing methods outside of the courseware domain. These data fuelled the research opinion that SCALPEL functionality that subscribes to a dialogue model would be influential in improving the efficiency of SCALPEL.

Although functionality to replicate demonstrator and peer dialogue was already being considered as an improvement to the SCALPEL courseware, students were asked what they considered to be the most important improvements. Responses focused on improvements relative to the quantity and quality of the domain material with on-line text books, the ability to ask questions and receive answers about the pathology material, a glossary of terms, and more links to background information being the more favourable of the proposed improvements. Improvements relative to the presentation of the material, e.g. more animation and the use of sound, were considered to be not so important.

9.16 Conclusions

Concerned with student attitudes towards SCALPEL and the need for demonstrator assistance, the data collected and analysed to satisfy this evaluation phase highlighted improved attitudes relative to the previous evaluation: students expressed a higher degree of acceptance of SCALPEL and less of a need for demonstrator assistance. Furthermore, statistical analyses of the data pertaining to the primary research hypothesis found a significant relationship between these two attitudinal concepts. This evidence led to the rejection of the null hypothesis in favour of the alternative: students who express a need for demonstrator assistance during SCALPEL sessions express negative attitudes towards SCALPEL relative to students who express less of a need for demonstrators.

In searching for explanations about the marked improvement in these attitudinal concepts relative to the previous study, statistical differences in attitudes expressed by students who had difficulty in solving problems and those who did not were observed. In the case of attitudes towards demonstrators, students who managed to solve all problems encountered during the pathology practicals expressed less of a need for demonstrators than those students who experienced unsolved misconceptions. For attitudes towards computers, a statistical difference was highlighted between these categories of students, with the distribution suggesting that students who solved all problems expressed a higher degree of acceptance of SCALPEL. Coupled with the overall low percentage of unsolved problems, it can be inferred that the improved student attitudes were as a consequence of students being pro active in solving problems encountered during the practicals. The data also showed that the SCALPEL courseware was a relatively ineffective method of solving problems with students opting for text books and tutorial/peer dialogue. Furthermore, students placed more importance on improvements to SCALPEL that would replicate the methods they most frequently used during this evaluation, namely, on-line text books and the ability to ask questions and receive answers about the pathology material. This empirical evidence underlined the need for improved SCALPEL functionality that would take away from students the onus of having to look outside of the courseware domain for solutions to pathology problems.

In conclusion, the insight into the role of demonstrators, and the methods used to solve problems proved to be valuable to this research. Addressing student needs for demonstrator assistance by offering effective SCALPEL functionality could result in a marked improvement in students' acceptance of this method of delivery.

9.17 Summary

This chapter presented the second evaluation of SCALPEL. The entire research process was discussed relative to this phase of the research before presenting the results of the analytical and descriptive analysis. Finally a discussion was offered on the interpretation of the findings and their impact on the research focus. Having identified the need for improved SCALPEL functionality based on the role of demonstrators in the traditional laboratory, the next chapter discusses a dialogue model that addresses students' misconceptions by mediating dialogue with domain experts within SCALPEL.

Chapter 10: A Model for Student-Demonstrator and Peer Dialogue

10.1 Introduction

The previous chapter highlighted the need to augment SCALPEL with functionality that would take away from students the onus of having to look outside of the courseware domain for solutions to pathology problems. This chapter looks at ways of offering functionality for discourse and domain expert advice within the SCALPEL environment. The limitations of the first implementation of SCALPEL are presented in section 10.2. Next section 10.3 defines the considered requirements of student-demonstrator and peer dialogue within SCALPEL before section 10.4 looks at computer mediated communication and examples of systems that facilitate the use of dialogues in learning environments. Informed by the approaches used by these example systems, section 10.5 offers a conceptual definition of the approach employed by SCALPEL (see also, Kemp et al. 1999). Finally section 10.6 discusses the implementation of dialogue functionality within SCALPEL and offers an operational overview.

10.2 The Limitations of the First Implementation of SCALPEL

The first implementation of SCALPEL was limited in the way it offered help to students when faced with misconceptions, or offered help to reinforce students' correct understanding of concepts. Within SCALPEL the only means of offering students help in resolving misconceptions was provided by SToMP's MCQ engine (see section 3.5.1.4 and 4.3.4). This was achieved by offering relevant supplementary domain knowledge by utilising Microcosm's show links feature (see section 3.4.2.3).

The traditional laboratory environment facilitates the resolution of misconceptions by encouraging student-tutor and peer dialogue about the subject domain (see section 4.2.2). The implementation of SCALPEL and the subsequent withdrawal of demonstrators from students' pathology practical experience removed dialogue from the learning process.

Clearly, removing demonstrators also removes the opportunity of student-demonstrator dialogue, but in addition, the very nature of computers encourages a closed working environment that offers no or little support for collaborative work. This was highlighted through participant observations. Most students using SCALPEL worked through the practical individually with the occasional group of not more than 3 students temporarily forming to engage in dialogue (see section 7.15). This observation was in contrast with the patterns of interaction in the traditional laboratory where students working in groups of up to eight actively engaged in discussions about the domain material, and where demonstrator explanations and discussions were observed to be beneficial to more than one student.

The absence of demonstrators during the second evaluation forced students to look at alternative means of resolving problems encountered during the pathology practicals. Textbooks, tutorial and peer dialogue were the most common methods used by the students (see section 9.15). When asked about how SCALPEL could be improved, students placed a greater emphasis on the need for on-line textbooks and the ability to ask questions and receive answers about the pathology material. Furthermore, students' attitudes towards the need for demonstrators to be available during SCALPEL sessions were strong, and were shown to be a significant influence on attitudes towards SCALPEL.

The literature (see section 3.2) suggests that learning can be improved by encouraging dialogue with peers and teachers. Indeed the empirical evidence presented in this thesis (see chapter 9) strengthens this claim. Consequently, functionality that mediates demonstrator assistance, as typified by dialogue in the traditional laboratory, was considered to be a necessary augmentation to the SCALPEL learning environment. The task of improving SCALPEL was therefore essentially seen as one of mediating student-demonstrator and peer dialogue within the SCALPEL environment.

10.3 Defining Student-Demonstrator and Peer Dialogue Functionality

Informed by the literature (see chapters 2 and 3), and the findings of the evaluations presented in chapters 7, 8 and 9, this section offers a description of what was determined to be the requirements, from the students' perspective, of student-demonstrator and peer dialogue within SCALPEL.

Fundamentally, improvements were seen as providing two distinct but complementary roles. As discussed in section 10.2 SCALPEL primarily needed to mediate student-demonstrator and peer dialogue about context sensitive concepts of any knowledge domain authored into the learning environment. A secondary, but none the less important, consideration was that raising a new line of enquiry should be the last resort in resolving problems during a pathology practical. Reasoning that previous discussion threads and/or domain material could already be available to satisfy any given query, it was considered advantageous from the perspective of both students and demonstrators for SCALPEL to attempt to address misconceptions before a new, and potentially repetitive, dialogue thread is instigated by a student.

An interface was therefore required that would allow students to tap into a resource of previously asked questions relevant to the concepts being addressed by the student at any given time. If, having exhausted existing relevant questions and answers, and having determined that any supplementary material presented by SCALPEL does not satisfy a query, students would then be able to post a new line of enquiry for the attention of a domain expert. To this end, SCALPEL needed to continually identify the concepts being addressed by an individual in order to present existing relevant dialogues, and/or mediate new lines of enquiry in the correct discussion threads.

10.4 Looking at ways Forward

The potential for collaborative dialogue and resource sharing falls under an area of research known as computer mediated communication (CMC). Whalley (1998) comments that an important part of CMC has been the development of tools for collaboration in the work place, an area of research known as computer supported cooperative work (CSCW), and highlights the interesting point that 'as the shared tasks have become more conceptual and less task oriented, the "cooperative work" of CSCW has evolved into the "collaborative learning" of CSCL' (Computer Supported Collective Learning). From an educational perspective, the primary push behind research into CMC is the opportunities that lie behind distance learning, not only in delivering course material (see, for example, Masterton 1998; Petre et al. 1998; Stratfold 1998), but also in providing systems that address traditional support services typically offered in a bricks and mortar university, e.g. counselling service, course selection advice etc (Scott and Phillips 1998). Distance learning, which aims to engage the student in a 'community of learning' (Petre et al. 1998), is viewed as a means of alleviating the continuing problem of increased staff-student ratios and is perceived as offering the potential to promote lifelong learning (Summer and Taylor 1998). What underpins all of these research domains is the necessity to mediate communication, whether it be text, video or voice, between users in different localities and at different times.

Technology affords two modes of communication, synchronous and asynchronous. Synchronous dialogue entails real-time discussions and, in the context of SCALPEL, would require the co-ordination of students and demonstrators to use the system at the same time (for examples in education see, Petre et al. 1998; Scott and Eisenstadt 1998; Whalley 1998). Asynchronous dialogue on the over hand is not real-time, e.g. Email. Given that a benefit of SCALPEL was communicated as offering students the opportunity to work through practicals at their own pace, whenever it suited them (see section 4.2.3), the synchronous mode of communication was determined to be inappropriate within the existing SCALPEL ethos. As a consequence the review of the literature primarily focused on asynchronous communication. The rest of this section looks at exemplar systems that facilitate the use of asynchronous dialogues in learning environments.

10.4.1 Knowledge Tree

Advocates of dialogue in the learning process, Brailsford et al. (1996) believe that it is 'one of the most fundamental aspects of education'. Consequently they developed Knowledge Tree (KT), a software system designed to facilitate discourse in computer-based learning. Brailsford et al. (1996) were inspired by the work carried out by Ackerman and Malone (1990) on the 'Answer Garden' approach to learning. Originally intended to distribute an organisation's knowledge, Answer Garden presents users with a branching network of questions about problem domains. Questions previously asked offer immediate answers while new questions for which solutions are required are emailed to a relevant domain expert. Mayes (1993, 1995) first highlighted the potential of the Answer Garden approach in educational, describing it as 'evolving courseware' that mediates concept-specific questions, answers and discussions about subject domain material. Brailsford et al. (1996) found that the problem with the original work on Answer Garden was that the users were isolated from each other. They could see questions posed in the past, but had no means of engaging in dialogue with other users. To this end they built on the Answer Garden approach by providing peer-peer as well as student-tutor dialogue.

Underpinning the KT model are two components: a hierarchical concept-oriented knowledgebase and a threaded bulletin board called a Forum, a place where students can ask questions and participate in ongoing debates. When a student raises a new question it is automatically emailed to the relevant domain expert. Once answered, questions and answers are archived into the subject-specific hypermedia knowledgebase for future reference. Students use a Concept Browser to navigate the knowledgebase. The Concept Browser is essentially a folding outline of the concept hierarchy and represents the structure of the concepts addressed by the domain. In the Forum the Concept Browser is used to attach a concept to a question.

Using a KT implementation of a domain within an education setting, students are able to engage in bulletin board style discussions about the subject while a database of frequently asked questions allows them to find existing answers about the knowledge domain. Students are provided with functionality not only to view questions their peers have asked, but also to comment on or even answer them.

Brailsford et al. (1996) used KT as ancillary support to lectures with three different cohorts of students but offer no empirical evidence regarding its effectiveness.

10.4.2 The Extended Bulletin Board System

Stratfold (1998) categorises online interacting and collaboration in three broad areas: unstructured/guest book; real time and near real-time/chat boards; and conferencing. The guest book approach is structured as a linear list of comments. New comments are added to the end of the list. The list displays the text and titles in one page while structure cues, e.g. 're', in the title constitute the only means of tying together entries on the list. Chat boards are structured similar to a guest book but differ in that the pages are automatically reloaded at regular intervals to present new messages and allow a more synchronous conversation. Typically, messages posted to chat boards are not archived and only remain available for a short period of time. Conferencing has more structure than a guest book in that it provides mechanisms that allow users to see how messages interrelate. In a conferencing system messages are viewed by selecting a subject heading from a structured list, or concept hierarchy, of headings. Stratfold (1998) believes that the structure inherent in conferencing systems provides 'a better mechanism for linking in longer messages and supporting an asynchronous discussion'.

Stratfold (1998) identifies key characteristics that differentiate dialogue systems for education from systems that are used in other contexts. These characteristics include focused discussions, the use of tutors, and course management: discussions are likely to be focused around particular issues; there is likely to be an authoritative role by means of a tutor or demonstrator; and the system ought to provide some mechanism for guiding learners through a course. Based on these characteristics, Stratfold (1998) designed the Extended Bulletin Board System (EBBS), a World Wide Web-based asynchronous computer mediated conferencing system.

EBBS adopts a hierarchical tree to represent the message/reply discussion structure. Each subsection of the tree represents a series of interchanges about a particular topic and is referred to as a thread. Various threads converge at the root of the tree, which represents the starting point for a particular conference. In order to communicate ancestry, and therefore context, a discussion is presented as an outline of headings. If a heading is indented relative to the heading above then it is a reply to that message. The hierarchical structure contains three pieces of information that are determined to be essential to the contributing dialogue: attribution (who said what), sequence (in what order), and relationship (to whom). This information allows a learner who is returning to a conversation after a period of absence to catch-up on new messages and the direction of the conversation.

EBBS functionality supports different approaches to finding messages. Learners can view messages by directly navigating the outline, obtaining a list of new messages, or searching on content, heading or author.

EBBS builds on traditional bulletin board systems by recognising individual learners. This allows the system to show individuals new messages and areas of activity since they last visited. This is achieved by keeping track of what messages the learner has seen and by presenting to the learner a value that reflects the number of unread messages in a thread. Returning to the system after a period of time, a learner can elect to view a list of new messages in the conference or for a particular thread. However, the context of the message, normally inferred by presenting the hierarchical ancestry, is not known to the learner. Stratfold (1998) reports that evaluations have shown that learners find this problematic and consequently prefer to navigate using the hierarchical outline as this allows them to be informed of the ensuing dialogue context.

The only evaluation evidence offered is anecdotal. Stratfold (1998) suggest that having implemented the EBBS in two courses, they found that: one third of students never used the system, one third read the messages but did not add much, and one third were active in reading and posting.

10.4.3 Virtual Participant

Masterton (1998) highlights repeated discussion threads and the subsequent duplication of effort as problems with online discussion forums. These problems are of particular interest within the domain of educational technology given the pace at which the Web is being used to deliver distance learning for large-population courses. Masterton (1998) observed that online conferencing forms a reusable resource of problems and solutions, and that it allows distance learning students to communicate more freely with each other and tutors, resulting in an increase of staff time and resources. On the basis of these observations Masterton (1998) developed the Virtual Participant (VP), an agent designed to support staff and assist students by identifying common themes and bring relevant past discussions to the attention of all involved in a distance learning course.

Within VP, conferences are presented to students as icons on their desktop, with each icon representing a course. When a conference is opened students are presented with a list of current messages with the most recent first. This flat list presentation is in contrast to the hierarchical structure of both the Knowledge Tree (section 10.4.1) and EBBS (section 10.4.2) systems.

The Virtual Participant adopts a case-based (see section 3.2.4) approach by identifying problems experienced by past students and presenting them as cases to current students in the right context and at the right time. Each individual case represents one previously experienced problem and the key discussions surrounding its resolution. Using previous years' discussions VP builds a case-base of past interactions through a four-step process. First messages are grouped into threads. Threads which discuss the same concept are then grouped into topics. Next keywords and phrases are identified for indexing before topics which discuss similar things are grouped together to generate a story.

VP continually interacts with the conferences, accumulating keywords from current discussion threads. Identified keywords are then used by a pattern matching retrieval algorithm to match cases/stories from the case-base and provide students with contextually relevant cases. When a case is retrieved an overview message is posted explaining that the VP has identified a thread of conversation from the previous year that may be relevant to the current discussion. The message gives the original context of the case and an overview of the constituent problem. Students are then given the opportunity to retrieve further information about the case through a number of fixed questions that can be mailed back to the conference.

VP was evaluated (Masterton 1998) by using it in The Open University Business School's 'Creative Management' course. Masterton (1998) reports that VP coped in the opening stages of the course because the problems experienced with the domain material, and the terms used by the students to describe them, were 'quite constrained'. But as the course progressed and the problems experienced became more diverse, VP's pattern matching retrieval algorithm became 'inappropriate', with visual inspection of the conferences revealing discussions for which VP failed to retrieve relevant stories. However, positively, those stories that were retrieved were deemed relevant to the context.

Of 104 students that were surveyed 40 returned questionnaires, of which 2 claimed to have never read any VP messages. Relative to the content and context of the cases presented by VP, Masterton (1998) offers the following student comments which he feels accurately summarises the students' and tutors' feelings about VP's contributions: 'Some of the contributions are appropriate, others are a tad on the side, but this may be useful as it may be providing a view from another angle'; 'I felt UB helped structure/ensure good quality information in the conference. One of my criticisms of conferencing is that tutors (as they would in a real tutor group) do not control poor quality student input'.

10.5 The Conceptual Design of SCALPEL's Student-Demonstrator and Peer Dialogue Functionality

There were two principal problems in realising a conceptual model of student-demonstrator and peer dialogue functionality that meet the requirements defined in section 10.3. The first was one of concept and context representation. How would dialogue threads be represented within the system? And what mechanism could be used to identify concepts, and their context? The second problem was navigational in nature, and centred around SCALPEL being able to determine the concept being considered by a student, and in what context.

10.5.1 Concept Representation

Whereas the KT (section 10.4.1) and the EBBS (section 10.4.2) used concept hierarchies to model the underlying domain concepts of a given course, the approach employed to realise contextualised discussion threads within SCALEPL uses a concept hierarchy to model the structure of the pathology course and practicals. Within SCALPEL, every entity that makes up the structure of the pathology course is considered to be a concept (represented as a node in the hierarchy), i.e. the pathology course (concept) is made up of a number of practicals/case histories (concept) which in turn are modelled around a number of questions (concept) (see Figure 27). In addition, using a system of upward inheritance, every entity is also considered to go some way in representing the context of the concepts represented by the nodes in its subtrees. For example, in Figure 27 'Term 2' is a concept, but it also contextualises each of its constituent practicals and questions. The higher the level of any given concept, the more focused the context, e.g. leaf nodes represent questions in the practicals and therefore specific concepts.

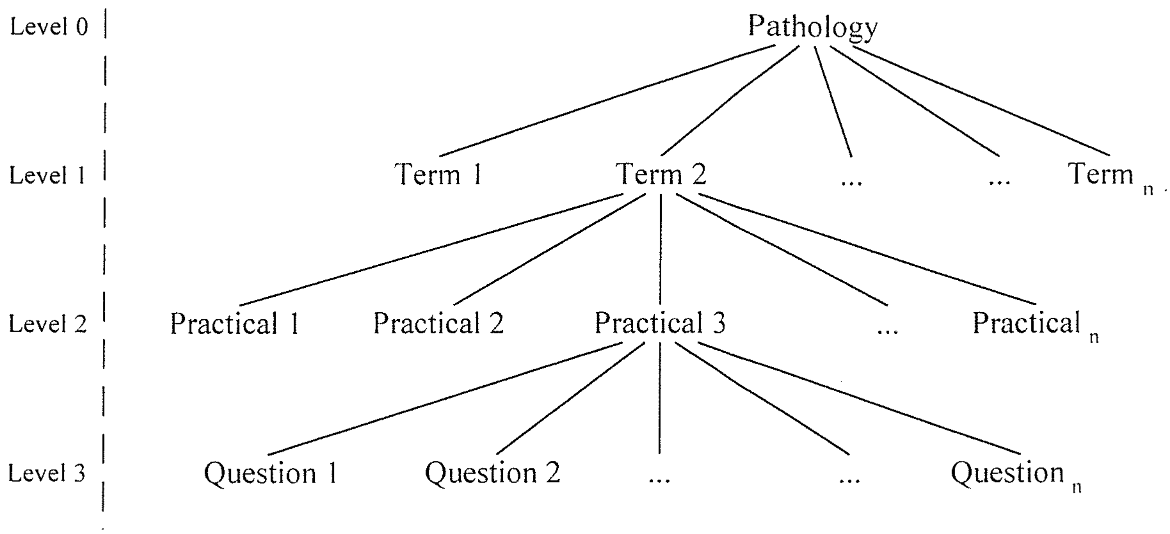


Figure 27 The hierarchical structure of pathology practicals

Each concept is uniquely defined by a numerical 'concept ID'. Because each concept also represents the context of the concepts in its subtrees, each concept is also given a 'context ID'. The technique used to determined the context ID for any given concept is analogous to the way symbolic domain names are used to represent Internet (IP) addresses (see, for example, Coulouris et al. 1994; Halsall 1992). A concept/node's context ID is built by appending the concept ID of each successive node in the path from the root node to the relevant concept/node in the hierarchy (see Figure 28), delimiting each concept ID with the period character ('.'). For example the context ID for Question 2 in Figure 28 is '1.3.9.13', i.e. the path is: Pathology (1) - Term 2 (3) - Practical 3 (9) - Question 2 (13). However, the context for Question 2 is '1.3.9' because the context ID attached to a concept does not represent that concept's context, but the context of its child concepts. Within KT and EBBS, no scope is offered to associate concepts that may hang from different nodes on the hierarchy: concepts that are closely related and therefore potentially relevant to a given discussion thread. This representation tackles this shortcoming.

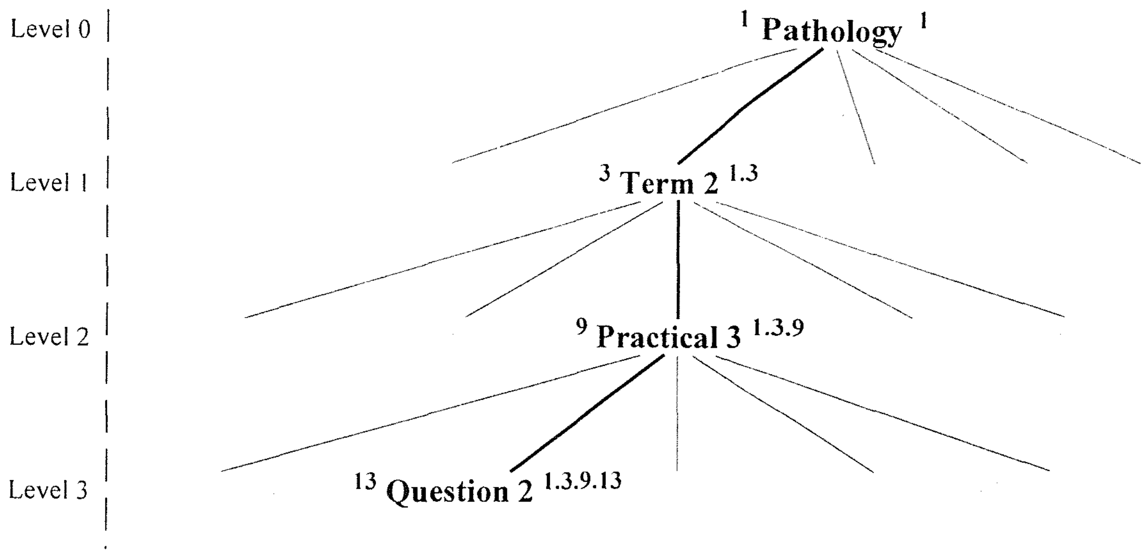


Figure 28 The indexing scheme employing to represent concepts within SCALPEL. The number before each concept heading is the concept's unique ID, while the superscript numbers after are used to determine the context of concepts further down the tree.

Having both a unique concept and context ID facilitates a many-to-many relationship between concepts and contexts. This allows a concept to be represented in multiple contexts, e.g. Question 2 may be presented in two different practicals under different contexts. From an operational perspective, students can view discussion threads and/or any supplementary material about concepts that may be presented in different contexts in other pathology practicals/cases. This representation is not concerned with what, for example, Term 2, Practical 3 and Question 2 are about, it does however offer a framework for facilitating contextualised dialogue within SCALPEL. The way this is achieved is discussed in the next section.

10.5.2 Concept Navigation

Providing functionality to mediate contextualised dialogue threads within SCALPEL is achieved as a result of the structure of the pathology practicals and the way students work through them. As discussed in section 4.2.1, practicals are structured around fictitious patient case histories and relevant questions. Questions are authored by domain expert in an attempt to prompt students to reason about the case histories and their underlying concepts. It is at the point of answering a question that students realise that they hold a misconception or require their understanding to be reinforced. In the traditional laboratory this was tackled through dialogue with peers and demonstrators (see section 4.2.2). Clearly, within the traditional environment, the concept that a student wishes to discuss, and indeed the context, is visually apparent to both student and demonstrator given the synchronous nature of the dialogue and the presence of the pathology material.

In both the KT and EBBS, and with bulletin boards in general, the onus is left on the user to identify the discussion thread that they believe best represents the concept and context of what they wish to discuss. This problem manifests itself in navigating the concept hierarchy. Problems can arise navigating the concept hierarchy when the same concept is discussed in differing contexts, i.e. within different discussion threads (different points on the concept hierarchy), or where there is the potential of differing understandings of the same concept being held by students. This problem is acknowledged by Mayes (1997), who comments: "Asking one's own question, and having it answered by a tutor sensitive to the context in which it has been asked, is valuable. How valuable, though, will that same answer be to a future questioner whose learning context many be subtly different? More important, how useful would a database of questions and answers be to a learner who cannot yet pose a question at all?"

When the onus is left with the student to identify a relevant discussion thread by navigating a concept hierarchy, the potential arises for related discussion threads to take place simultaneously at different points on the concept hierarchy as a result of students categorising their conceptions differently. Two advantages were seen in taking the onus of identify relevant discussion threads off students: first it would increase the likelihood of an existing discussion thread being found to address a student query before a new line of enquiry is embarked upon; and second, it would minimise the potential for duplicate discussion threads. Consequently a mechanism was required within SCALPEL to present students with existing discussion threads relevant to the context of the concept under consideration (i.e. the practical question being attempted), and to ensure mutual understanding between student and tutor, and/or peers, about the concept and context being addressed in a new line of enquiry.

Essentially, for SCALPEL, the problem was one of being able to determine the MCQ question, the smallest structural entity of a pathology practical, being addressed by a student. Using the hierarchical structure introduced in the previous section, if the question can be identified then so can the context. The problem of determining the MCQ question being considered by a student is dealt with by turning to the MCQ engine, an integral component of SCALPEL (see section 4.3.4). By tying together the MCQ and dialogue functionality, SCALPEL is able to identify the question currently being addressed and present to the student the relevant discussion thread. For example, if the MCQ engine is currently set to deal with question 2 in term 2's practical 3 (see Figure 28), SCALPEL can retrieve the discussion thread evolving around the concept ID '13' in the context identified by the context ID '1.3.9'. As a student moves through the practical, and by definition, the MCQ engine, SCALPEL is continually able to identify the concept currently under consideration and in what context. Students can then be confident that the dialogue thread (list of questions and answers) presented is directly relevant to the concept/MCQ question they are currently addressing. In addition, when they generate a new MCQ question they do not need to be concerned about traversing the concept hierarchy to identify and attach an appropriate concept. This is also of benefit to domain experts who will automatically know the concept and context of any given question posted to them.

10.6 Implementing Student-Demonstrator Dialogue Functionality

A proof-of-concept application was developed to evaluate the proposed model for student-demonstrator dialogue. Given that the dialogue functionality needed to work in association with the MCQ engine, and considering presentational issues about overcrowding the desktop with too many windows (see section 4.3.3), a decision was made to implement the dialogue functionality as part of a new MCQ/Dialogue agent.

10.6.1 Concept Hierarchy

The concept hierarchy was implemented as a table in a relational database management system (RDBMS). For the purposes of implementing the proof-of-concept system Microsoft Access was determined to be sufficient.

A single table 'Concept_Contexts' was designed to represent the concept hierarchy (see Appendix 8, section A8.1). Each tuple in the table represents a concept in the hierarchy and is identified by a composite primary key consisting of the 'ConceptID' and 'ContextID' fields. This allows multiple concepts to be represented in multiple contexts.

A second table 'Questions' was designed to represent contextualised dialogue (see Appendix 8, section A8.2). Each tuple in this table represents a single student question and domain expert answer, and is related to relevant context in the 'Concept_Context' table by means of the foreign key field 'ContextID'.

10.6.2 MCQ/Dialogue Agent

The MCQ /Dialogue agent was developed as a Microcosm aware viewer (see section 3.4.2.3) using Visual Basic. The fundamental workings of SToMP's MCQ engine (see section 3.5.1.4 and 4.3.4) were augmented by functionality to mediate contextualised dialogue, and present students with relevant supplementary material. The MCQ/Dialogue agent as first viewed by students can be seen in Figure 29.

During execution, the viewer is configured by reading a configuration script (see Appendix 7) designed to be backward compatible with the format employed by SToMP's MCQ engine. A single change was made to SToMP's original configuration script to allow the MCQ/Dialogue agent to tie together any MCQ question with the corresponding discussion thread on the concept hierarchy. By simply tagging each question in the configuration script with its valid context ID, the MCQ/Dialogue agent is able to interrogate the concept hierarchy for any dialogue that is evolving around that particular concept in the correct context. For example, in Appendix 7, the context ID of Question 1 in term 2's practical 6 is '1.2.5.6'. This is explicitly expressed by the 'ContextID' tag within the configuration script. The dialogue functionality within the agent can then select all tuples in the 'Questions' table (see section 10.6.1) where the 'ContextID' field = 1.2.5.6, thus offering students with a discussion thread that is relevant to the MCQ question they are currently attempting. At the same time, if a student wishes to engage in a new line of enquiry, the correct context ID is attached to the question, and therefore it is correctly contextualised within the concept hierarchy.

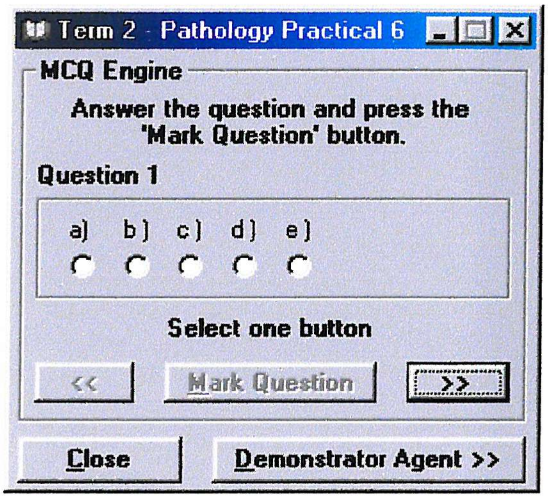


Figure 29 SCALPEL's MCQ / Demonstrator Agent engine

Presenting discussion threads about a given concept in different contexts is also catered for within this implementation. By knowing the context ID of a given concept, which is explicitly stored in the agent's configuration script, the concept ID can be determined from the concept hierarchy database. The concept ID can then be used to select all other contexts from the 'Concept_Contexts' table that share the same concept ID. Discussion threads from the 'Questions' table can then be retrieved and presented to the student.

In addition to the functionality that supports contexualised dialogue, the agent builds on the way SToMP's MCQ engine presented relevant supplementary domain material to students (see section 4.3.4). Microcosm's show links feature is again used by the agent, but instead of requiring keywords to be explicitly identified and mapped to supplementary material by a domain expert, an entire text is passed to the show links feature for resolution. The text passed is either feedback provided by the MCQ engine in response to a student's attempt at answering a question, or a question and answer pair from a discussion thread. The text that is passed is dependent on what the student is currently doing within the agent, i.e. answering an MCQ question, or viewing discussion threads.

10.6.3 Authoring a Pathology Practical

When a new practical is authored, two components of the MCQ/Dialogue agent need to be configured. A new configuration script needs to be authored and the concept hierarchy database needs to be amended to reflect the addition of new concepts.

Because each practical and its constituent questions are recognised as concepts (see section 10.5.1) the concept hierarchy needs to be updated to reflect the new additions. The practical is allocated a unique concept ID and assigned a context ID that reflects its position/context within the hierarchy (see section 10.5.1). A new tuple in the 'Concept_Context' table is added to reflect this new concept. Next a tuple is required for each of the pathology questions authored in the practical. Again a unique concept ID is assigned to each question and the context ID determined by its position within the hierarchy. When the configuration script is authored the context ID for the practical, and each of the MCQ questions is included in the script. This ties the MCQ questions in the MCQ/Dialogue agent viewer with the discussion threads in the concept hierarchy database.

Crucial to the success of the MCQ/Dialogue agent, is the consistency of context IDs between the concept hierarchy database and MCQ/Dialogue agent configuration script. For the purposes of the proof-of-concept implementation both the concept hierarchy and the configuration script were authored manually. Initial population of the concept hierarchy was achieved by using a domain expert to identify common questions posed during practical sessions and author appropriate answers. The resultant discussion threads were then manually entered into the concept hierarchy database.

10.6.4 Operational Overview

As students work through the practicals, they answer the practical questions in much the same way as using SToMP's MCQ engine (see section 4.3.4). For a given MCQ question feedback is authored by domain expert to be informative irrespective of whether it is answered correctly. Through Microcosm's show links functionality, this feedback is useful in attempting to find domain material that may be relevant to the MCQ question. Students can look for supplementary material by activating the 'Show Links' button on the feedback dialogue box (see Figure 30). This is particularly useful if students need to reinforce their understanding of the concepts addressed by the MCQ questions. The effectiveness of this functionality is dependent however on domain experts providing sufficient supplementary material for authoring into SCALPEL.

Students can activate the dialogue functionality by pressing the 'Demonstrator Agent' button. This name was chosen as it was considered to be a metaphor that best describes the purpose of the functionality. When activated, the agent opens up to present a hierarchical outline of questions previously asked in respect of the MCQ question (see Figure 31). The hierarchy presents two root nodes. The first represents all previously asked questions, while the second focuses on questions asked by the specific student. When a question is selected the details are presented to the student. In addition to presenting the question itself, the student is informed of when the question was raised and who raised it.

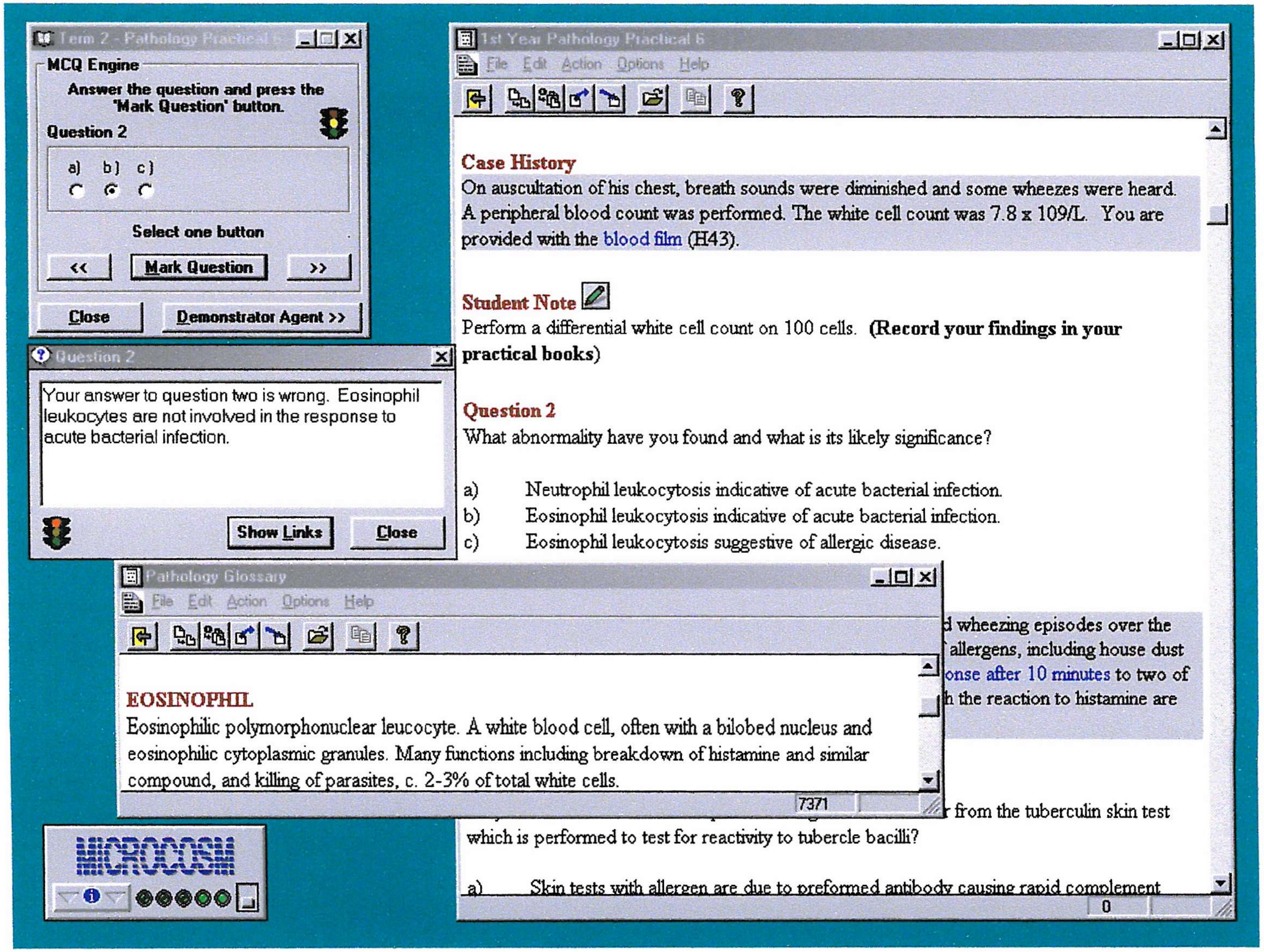


Figure 30 Screenshot of SCALPEL's MCQ/Dialogue agent using Microcosm's show links functionality

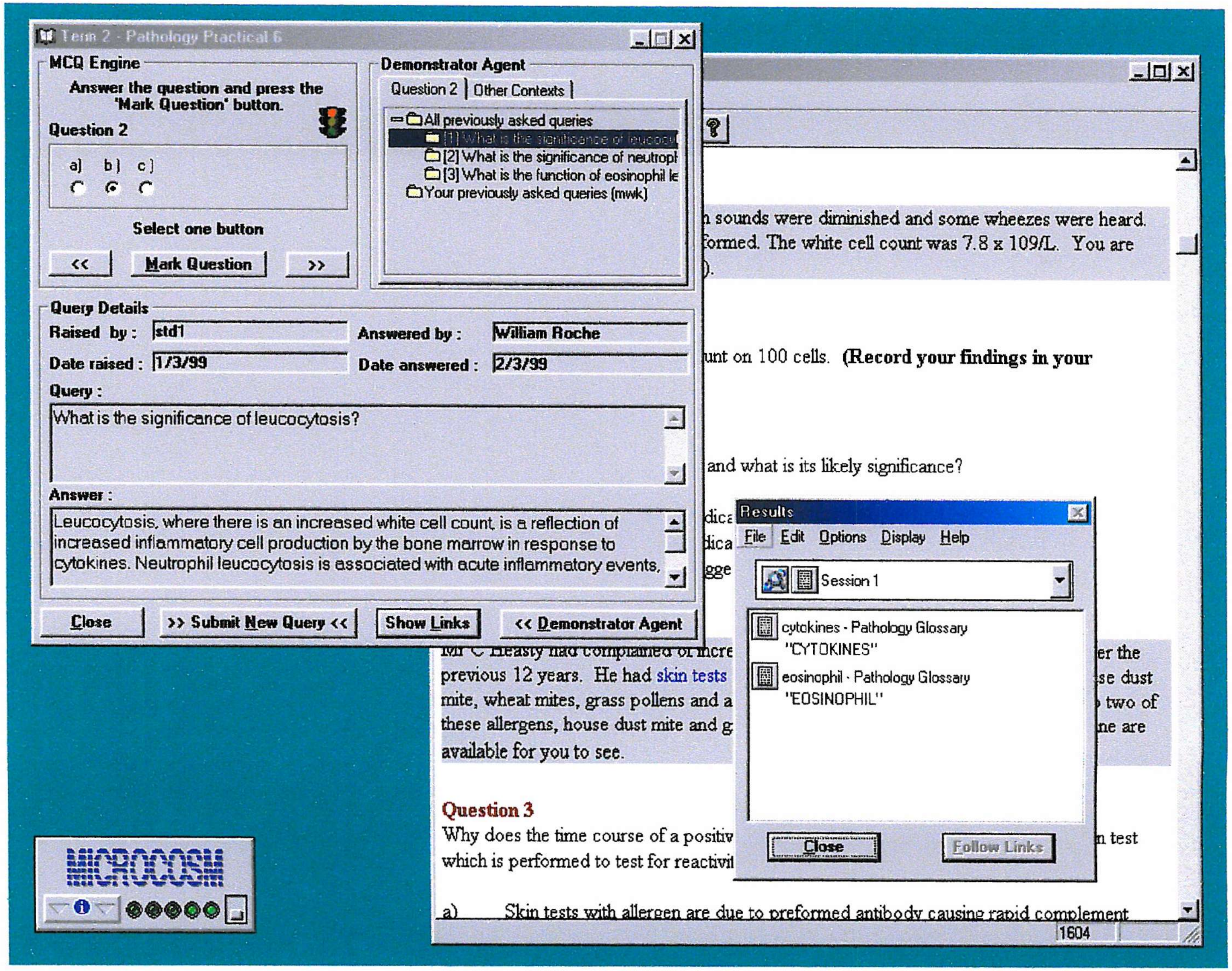


Figure 31 Screenshot of SCALPEL's MCQ/Dialogue agent using Microcosm's show links functionality

If the question has been answered, the answer is presented along with who answered it and when. Students can elect to raise a new line of enquiry by pressing the '>> Submit New Query Button <<'. All new enquiries are posted to the concept hierarchy where, via a different interface, a domain expert authors replies. Students can also elect to view discussion threads about the MCQ question under consideration in different contexts by selecting the 'Other Contexts' tab. In addition, an attempt can be made to identify any supplementary material by pressing the 'Show Links' button. In this mode the question and answer text of the currently selected question are passed through Microcosm's show links functionality, e.g. Figure 31 shows two relevant links (Results dialogue box) that SCALPEL believes may be relevant to the question in focus.

As students navigate through the MCQ questions ("<<" and ">>" buttons) the concept hierarchy refreshes to present the contextualised discussion thread for that particular question. This ensures that students are always presented with relevant discussion threads, thus maximising their chances of having their problems addressed by past dialogue before embarking on a new line of enquiring. In addition, all new lines of enquiry are contextualised by the agent thus ensuring consistency through the concept hierarchy.

10.7 Conclusion

This chapter introduced a model of discourse designed to address the limitations with the first implementation of SCALPEL. Having identified the requirements, a conceptual overview of a new MCQ/Dialogue agent was offered before a discussion on its implementation and operational functionality. The next chapter communicates the execution of a third empirical study designed to determine the effectiveness of this new functionality.

Chapter 11: The Third Empirical Study

11.1 Introduction

The previous chapter discussed the design and implementation of the MCQ/Dialogue agent. This chapter presents the empirical study of this new SCALPEL functionality with a focus on its effectiveness relative to the first implementation of SCALPEL (see chapter 4). The sections in this chapter are modelled on the structure of the second empirical study presented in chapter 9. Section 11.2 begins with the introduction of an evaluation framework that builds on the empirical evidence of the previous evaluations presented in this thesis. An insight into the evaluation process underpinning this phase of the research continues with section 11.3 discussing the conceptualization and operationalization of the concepts identified for this evaluation. Next, the sub hypotheses that address the pre-specified research agenda are introduced in section 11.4. The evaluation methodology specific to this phase of the research is considered in section 11.5, while section 11.6 assesses the reliability and validity of the constituent measures of the research instrument. Descriptive statistics describe the demographic makeup of the sample population in section 11.7. The pre-specified research agenda begins in section 11.8 with a look at students' acceptance of SCALPEL and continues with section 11.9 examining students' expressed need for demonstrator assistance. Section 11.10 concludes the pre-specified agenda with a look at the relationship between these two concepts. Turning to the curiosity driven research agenda, section 11.11 assesses the factual data on the reasons prompting students to ask for demonstrator assistance during SCALPEL sessions. Section 11.12 focuses on the methods used by students to solve problems encountered during the pathology practicals in the absence of demonstrators. Section 11.13 discusses the improvements students would most like to incorporate into SCALPEL. Finally, the work presented in this chapter is discussed in section 11.14 and the significant findings highlighted in section 11.15.

11.2 Defining the Evaluation Framework

Designed to complement the previous empirical evaluation, the evaluation framework for this phase of the research differed in two significant aspects. First, the evaluation techniques were designed to focus on a single practical, practical 6, and not the entire series of practicals used in the previous evaluations (see section 9.6). Second, the SCALPEL courseware was amended to include the MCQ/Dialogue agent functionality (see chapter 10) for use with an experimental group of students.

11.2.1 Pre-Specified Research

The second empirical evaluation (see chapter 9) inferred that students' attitudes towards SCALPEL would improve by providing methods to solve problems encountered with the pathology material. The primary objective of this phase of the research was to assess the effectiveness of the MCQ/Dialogue agent introduced in chapter 10. The attitudinal concepts 'acceptance of SCALPEL' and 'attitudes towards the need for demonstrators', and their relationship, continued to underpin this research. It was a predicted difference in the distribution of these attitudes relative to students who used the original and improved SCALPEL environments that influenced this pre-specified research agenda. The primary research was subsequently conceptually defined as:

> students who use SCALPEL courseware designed to solve misconceptions exhibit more agreeable attitudes towards the SCALPEL environment than those students who use the SCALPEL with this functionality unavailable.

11.2.2 Curiosity Driven Research

The primary objective of the curiosity driven research agenda was to highlight possible reasons for the distribution of attitudes towards SCALPEL, acceptance of SCALPEL, and their observed relationship. A secondary objective was the continuing need to improve the SCALPEL courseware. This agenda was essentially a replication (see section 5.8) of the curiosity driven research agenda of the previous evaluation (see section 9.2.2).

Hypotheses were not generated for this descriptive research since the statements about the research objectives were considered sufficient in describing the issues under investigation (see section 5.3).

11.3 Conceptualization and Operationalization

In keeping with the research methodology employed by the previous studies (see chapters 5 and 6), this section offers conceptual and operational definitions for concepts addressed during this phase of the research. Some of the measures operationalized for use in the previous evaluations (see chapters 6 and 9) are employed here with no changes, while others have been revised to reflect the focus on practical 6. All measures for this evaluation phase were operationalized on a questionnaire administered to students after working through practical 6.

11.3.1 Demographic Characteristics

The concepts 'age' and 'gender' were employed as part of this phase of the research to offer an insight into the demographic makeup of the sample population. No changes were made to the operational definitions introduced in section 6.3.2.

11.3.2 Opinions and Attitudes

Students' opinions and attitudes determined to be of interest to this phase of the research were 'acceptance of SCALPEL' and 'attitudes towards the need for demonstrators'. Multiple-item measure techniques (see section 5.4.1.3) were employed to operationally define these concepts.

Attitudes towards SCALPEL

Four indicators were used to stand for the concept of 'acceptance of SCALPEL'. Recoding of negatively directed indicators was carried out for indicator 3. A low score on the scale represents a positive attitude towards SCALPEL.

Indicator 1: I feel confident using a computer to work through practical 6 (+)
Indicator 2: I find working through practical 6 a satisfying learning activity (+)
Indicator 3: I found the practical 6 computer application confusing to use (-)
Indicator 4: I like the consistency of explanations offered by practical 6 (+)

The Need for Demonstrators

A multiple-item measure consisting of four indicators was employed to stand for the concept of 'attitudes towards demonstrators'. Recoding of negatively directed indicators was carried out for indicator 1. A low score on the scale represents a stronger appreciation of demonstrator assistance than a higher score.

Indicator 1: I did not need the help of demonstrators whilst working through practical 6 (-)

Indicator 2: I believe demonstrators should be available at set times for practical 6 (+)

Indicator 3: I missed demonstrator input whilst working through practical 6 (+)

Indicator 4: Having worked through practical 6, I will always need the help of demonstrators, irrespective of the quality of knowledge offered by the computer application (+)

11.3.3 Factual Data

No alterations were made to the factual data measures operationalized (see section 9.3.3) to asses the role of demonstrators, the methods used by students to solve problems in the absence of demonstrators, and the ways in which students believe SCALPEL should be improved. Minor changes were made on the questionnaire to the introductions to each of these measures to reflect the evaluation's focus on practical 6.

11.4 Research Sub Hypotheses

Consideration was given to ensure that the sub hypotheses derived for this evaluation phase satisfied conditions introduced in section 5.3. For clarity, the operational statement (alternative hypothesis H_1) and null hypothesis (H_0) for each research sub hypothesis is presented in the relevant sections of this chapter.

11.4.1 Pre-Specified Research Agenda

Three research sub hypotheses were formulated to assess the improved SCALPEL functionality. Sub hypotheses 1 and 2 were authored to look at the difference in attitudes for students who had exposure to the enhanced functionality within SCALPEL and those who did not.

Sub Hypothesis 1 - Medical students who use SCALPEL with enhanced functionality exhibit more agreeable attitudes towards SCALPEL than those who use SCALPEL with the functionality unavailable.

Sub Hypothesis 2 - Medical students who use SCALPEL with enhanced functionality express less of a need for demonstrator assistance than those who use SCALPEL with the functionality unavailable.

The final hypothesis was put forward to assess the relationship between students' acceptance of SCALPEL and their expressed need for demonstrator assistance.

Sub Hypothesis 3 - Medical students' need for demonstrator assistance influences acceptance of SCALPEL.

11.4.2 Curiosity Driven Research

Hypotheses were not generated to satisfy the descriptive investigations that make up the curiosity driven research agenda since statements about the research objectives were considered sufficient in describing the issues under investigation (see section 11.2.2).

11.5 Method

A prospective cross-sectional experiment (see section 5.6) was designed to satisfy the research agendas of this evaluation phase. A population of first year undergraduate medical students (N = 165) worked through practical 6 using SCALPEL. From this population, 40 volunteers used SCALPEL with the MCQ/Dialogue agent functionality, while the remaining 125 students represented a control group by using SCALPEL with the functionality turned off. To ensure that students used the appropriate version of SCALPEL, the MCQ/Dialogue agent was made available on a closed cluster of workstations, and admission was only permitted to subjects of the experimental group. Students in the control group were left to work through the practical in their own time, with the only brief being that they needed to complete the practical before the associated tutorial (see section 4.2). Apart from the MCQ/Dialogue agent, all other aspects of the courseware, e.g. domain material, remained constant. The traditional laboratory environment was not used and demonstrators were not available during this study.

11.5.1 Research Techniques

A questionnaire was administered to the control group at the start of the last lecture of the pathology course and collected after the lecture (see Appendix 11 for presentation of questionnaire). The same questionnaire was administered to the students in the experimental group after completing practical 6 in the controlled environment. Questionnaires from the experimental group were flagged so they could be coded accordingly when entered onto SPSS. Students who wished to take the questionnaire away for completion at a later date were given the opportunity to do so and were instructed to return them to the pathology office or their pathology tutor.

To encourage a respectful return rate a £50 cash incentive (see section 5.4.1.1) was offered to two students randomly selected (names in a hat) from the cohort of all students who returned a completed questionnaire with a removable name label. In addition, the anonymity of respondents was again respected in the hope that this would improve the response rate: all name labels were removed from the returned questionnaires for the cash incentive so there was no way of identifying respondents. The cash incentive notwithstanding, of the 165 students in the sample population only 74 (N = 28 for the experimental group, N = 46 control group) students returned completed questionnaires (91 non-responses). With the population of interest being undergraduate medical students, confidence was given that the sample of first year undergraduate medical students used for the statistical analyses presented in this study were representative of the population of interest. However, caution is advised when generalising the inferences to a wider population.

11.6 Reliability and Validity of Measures

This section addresses the reliability and validity (section 5.4.6) of the measures used during this phase of the research.

11.6.1 Reliability

Cronbach's alpha was used to test the internal reliability (see section 5.4.6.1) of each multiple-item measure employed as part of this study. All other measures were determined to be reliable through establishing face validity (see section 11.6.2).

11.6.1.1 Behavioural Characteristics

Attitudes towards SCALPEL

The alpha coefficient for the multiple-item measure that addressed students' attitudes towards SCALPEL (Alpha = 0.58, N of cases = 74, N of items = 4) suggested a moderate level of association between the constituent indicators.

The Need for Demonstrators

The alpha coefficient for the multiple-item measure (Alpha = 0.69, N of cases = 71, N of items = 4) authored to address students' need for demonstrators suggested a moderate to high level of association between the constituent items, giving confidence in the internal reliability of this measure.

11.6.2 Validity

All the measures used in this evaluation were determined to have face validity (see section 5.4.6.2). Concurrent validity could not be established given the difficulties associated with obtaining independent sources of information. For some measures construct validity was inferred by definition of their observed relationship with other measures.

11.7 Sample Frame Demographics

Introduced in section 11.5, a population of first year undergraduate medical students (N = 165) was used for this study. Demographic data were collected on a questionnaire administered to the population during the last lecture of the pathology course. With 91 non-responses, the sample frame consisted of 74 students. Of these 32 (43.2%) were male and 42 (56.8%) female. The majority of the students (79.7%) were under the age of 20. 16.2% were aged between 20 and 25, with the remaining 4.1% aged between 26 and 35 (see Appendix 4, Table A4 - 1). Although a disappointing response rate, these demographics give confidence that the sample frame was at least consistent with the two previous samples (see sections 7.4 and 9.8) with respect to these concepts, inferring that age or gender did not bias non responses.

11.8 Acceptance of SCALPEL

This section looks at the distribution of students' acceptance of SCALPEL for the entire sample, and the experimental and control groups.

11.8.1 Operationalization of Sub Hypothesis

Sub hypothesis 1 (see section 11.4) was operationalized to address the difference between attitudes towards SCALPEL for students in the control and experiment groups.

H_0: There is no difference between the control and experiment groups in the mean ranking of attitudes towards acceptance of SCALPEL.

H_1: There is a difference between the control and experiment groups in the mean ranking of attitudes towards acceptance of SCALPEL.

11.8.2 The Distribution of Student Attitudes Towards SCALPEL

The distribution of the sample for the concept 'acceptance of SCALPEL' was taken as an indication of how well SCALPEL was received by students (see Appendix 4, Table A4 - 2).

Figure 32 offers a graphical presentation of the distribution of student attitudes towards SCALPEL for this, and the previous two evaluations. Irrespective of the version of SCALPEL used to work through practical 6 during this phase of the research, a positive difference was observed relative to the previous samples. None of the sample expressed negative attitudes and a reduction in the percentage of students expressing neutrality resulted in an increase of positive attitudes.

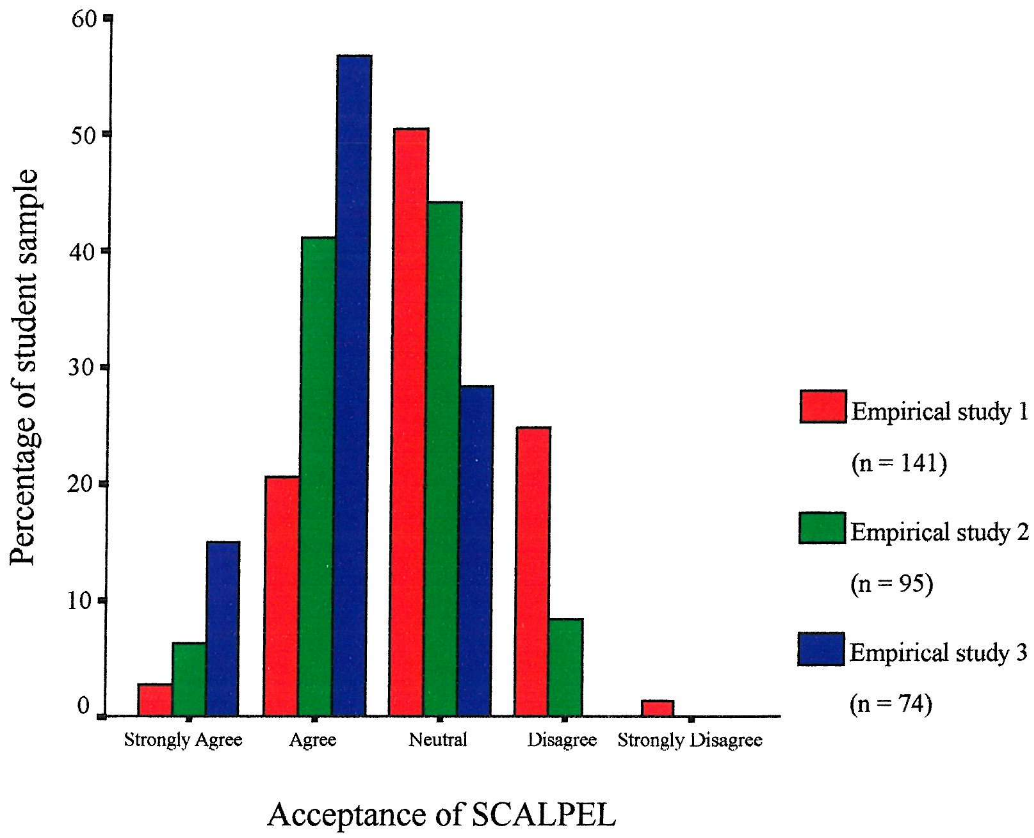


Figure 32 Bar chart on the distribution of students by acceptance of SCALPEL for empirical studies 1, 2 and 3.

Analysis of the breakdown of the sample into the control and experiment groups highlighted a difference in the distribution of attitudes (see Figure 33). A higher percentage of the experiment group (89.3%) expressed positive attitudes (agree and strongly agree categories) than the control group (60.9%), with the observed difference explained by the percentage of students in the neutral category (10.7% for the experiment group and 39.1% for the control group). In support of this observation the Mann-Whitney U test for two unrelated samples found the difference to be significant in the population ($U = 492.5$, $Z = -1.90$, $p = 0.05$).

The unanimous positive attitudes of the 46 students in the control group were unexpected given that the students of this group used the same version of SCALPEL as the previous samples, with all other variables held constant, e.g. the workstation specification, domain material, level of demonstrator availability etc.

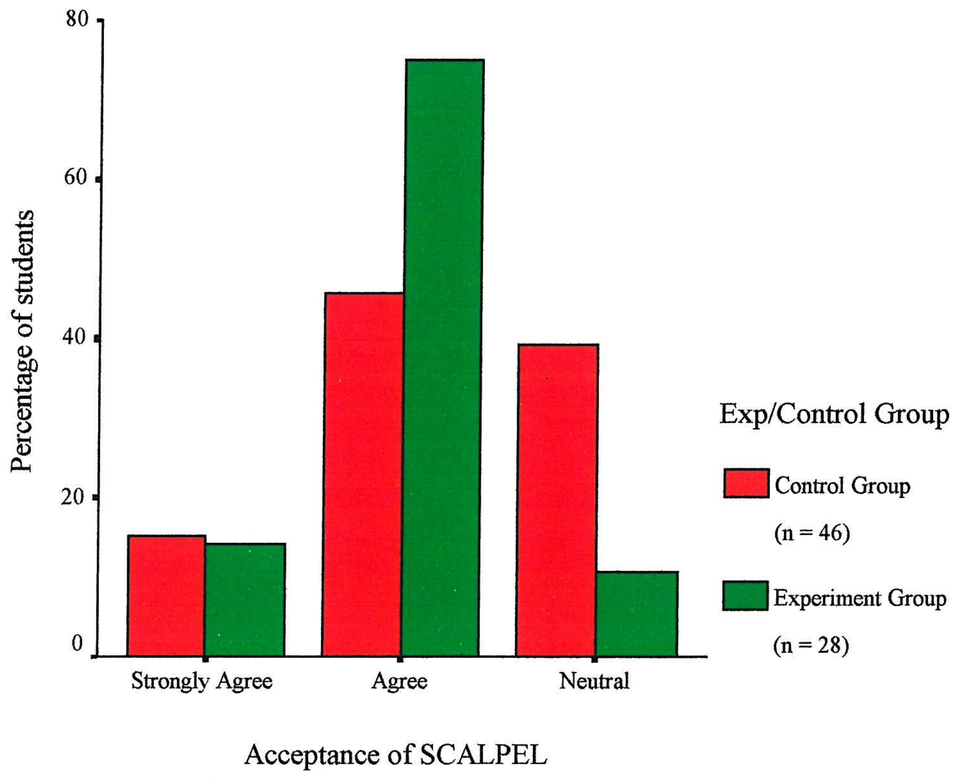


Figure 33 Bar chart on the distribution of students by acceptance of SCALPEL for the control and experimental groups.

11.9 Assessing the Need for Demonstrator Assistance

This section looks at the distribution of students' attitudes towards demonstrators for the entire sample, and the experimental and control groups.

11.9.1 Operationalization of Sub Hypothesis

Sub hypothesis 2 (see section 11.4) was operationalized to address the difference between students' expressed need for demonstrator assistance during practical 6 for students in the control and experiment groups.

H_0: There is no difference between the control and experiment groups in the mean ranking of attitudes towards demonstrator assistance.

H_1: There is a difference between the control and experiment groups in the mean ranking of attitudes towards demonstrator assistance.

11.9.2 The Distribution of Student Attitudes Towards the need for Demonstrators

The distribution of students' attitudes towards demonstrators (see Appendix 4, Table A4 -3) was taken as an indication of students' expressed need for demonstrator assistance during practical 6.

Presented graphically in Figure 34, the distribution of students' attitudes towards demonstrators improved relative to the previous studies, with students expressing less of a need for demonstrator assistance during practical 6. Only a small percentage expressed a need for demonstrators (8.5% in the agree category) and a notable increase was observed in the disagree (36.6%) and strongly disagree (11.3%) categories. Taking into consideration that the control group used the same version of SCALPEL as the samples from the previous evaluations, a higher percentage of the students in the control group were expected to express a stronger need for demonstrator assistance.

The Mann-Whitney *U* test for two unrelated samples found no significant difference between the attitudes of students in the control and experiment groups ($U = 591.0$, Z = -

0.14, p = 0.89). However, a look at one of the indicators of the multiple-item measure revealed an interesting distribution of the sample population (see Figure 35).

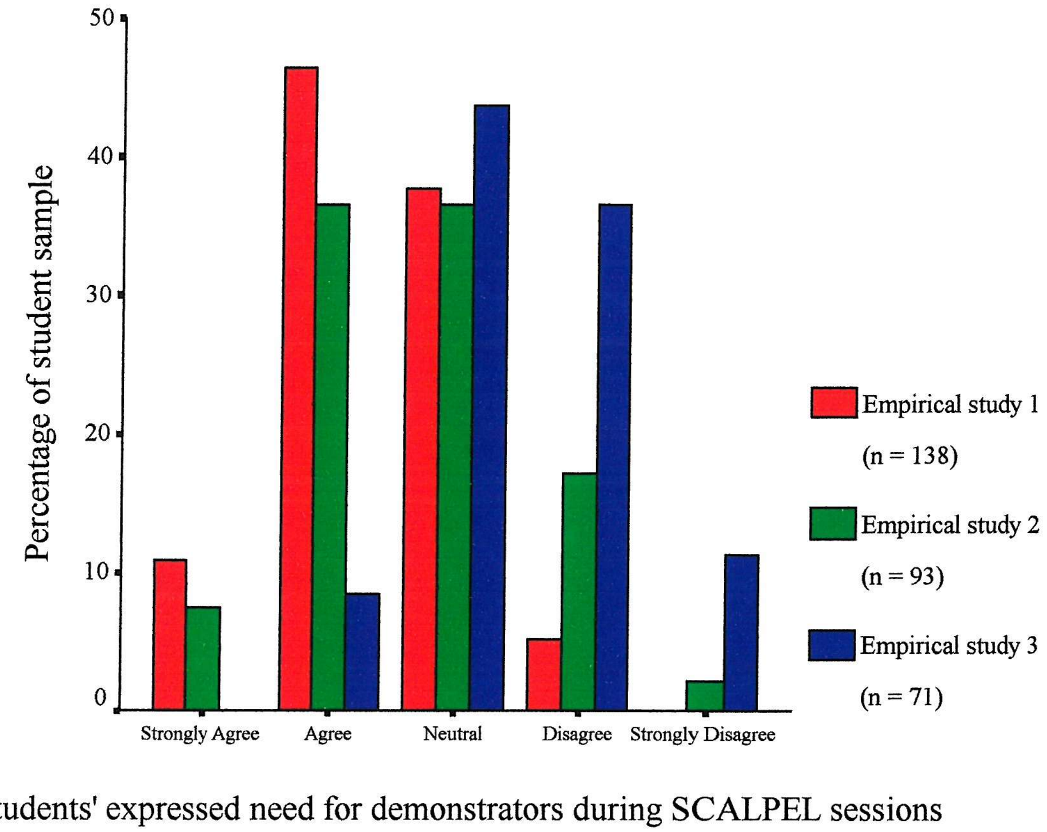


Figure 34 Bar chart on the distribution of students by expressed need for demonstrators during SCALPEL sessions for empirical studies 1, 2 and 3.

A higher percentage of the control group (31.1%) agreed to missing demonstrators whilst working through practical 6 than the experiment group (10.7%). With an almost identical percentage in both of the groups expressing neutrality, the difference is explained by a higher percentage of students in the experiment group (57.1%) than the control group (31.1%) suggesting that they did not miss demonstrator assistance. Furthermore, the significance test for two unrelated samples (U = 475.0, Z = -1.845, p = 0.65) fell just short of an alpha level (α) of 0.05, inferring a degree of confidence in a significant difference. One explanation for this observation is that although students in both groups expressed an equal need for demonstrator assistance whilst working though practical 6, those in the experiment group didn't miss demonstrators as much as a result of their problems being addressed by the improved SCALPEL functionality available to them.

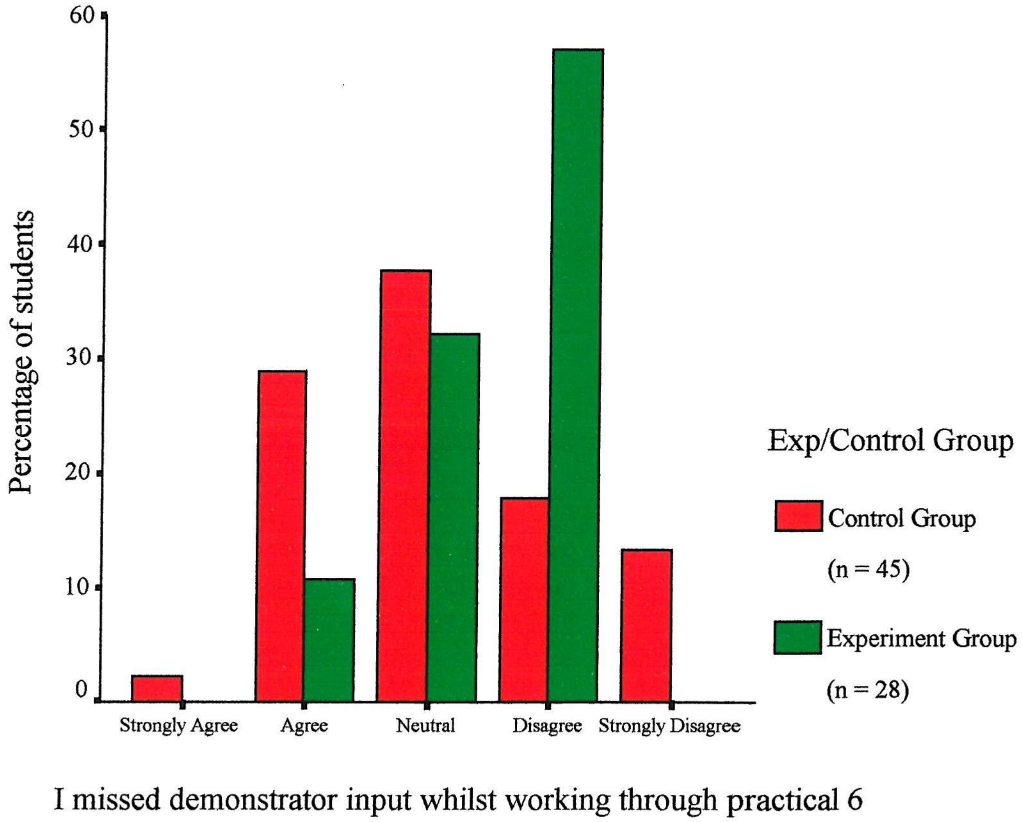


Figure 35 Bar chart on the distribution of students by the need for demonstrator input during practical 6. The data represents responses to indicator 3 from the multiple-item measure designed to stand for the concept of 'attitudes towards demonstrators' (see section 11.3.2).

11.10 Acceptance of SCALPEL and Demonstrator Assistance

This section addresses the effect of the need for demonstrator assistance on students' acceptance of SCALPEL.

11.10.1Operationalization of Sub Hypotheses

Research sub hypothesis 3 (see section 11.4) was operationalized to test for associations between the need for demonstrator assistance and acceptance of SCALPEL for the control and experiment groups.

H_0: expressed need for demonstrator assistance is independent of a student's acceptance of SCALPEL.

H_1: expressed need for demonstrator assistance is dependent on a student's acceptance of SCALPEL.

11.10.2Results

Because acceptance of SCALPEL was measured at the same time as expressed need for demonstrators, the independent (cause) and dependent (effect) variables could not be established by the time order of the two concepts. Theoretically the causal relationship between the two concepts can be reasoned either way, inferring that each concept is just as likely to have an effect on the other.

Table 3 presents Spearman's rho coefficients for acceptance of SCALPEL and the need for demonstrator assistance for the control and experiment groups. No significant relationship was observed in the control group. However, consistent with the relationships observed in the previous studies, a significant moderate negative relationship ($\rho = -0.43$, $p = 0.02$) was found to exist in the experiment group.

Table 3 Measure of association between students' expressed need for demonstrator assistance and acceptance of SCALPEL

Variable	Group	N	rho coefficient (ρ)	2-tailed Significance (p)
acceptmi	Control	43	-0.26	0.09
	Experiment	28	-0.43*	0.02

* Correlation is significant at the 0.05 level (2-tailed)

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered after exposure to practical 6. Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on the same questionnaire. The control and experiment group samples (N) were made up of students who expressed an attitude towards the acceptance of SCALPEL and the need for demonstrator assistance. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

11.11 Modelling the Role of Demonstrators

This section looks at the role of demonstrators, as perceived by students, during practical 6 with a focus on examining any underlying differences between the experiment and control groups. Of the 74 students in the sample frame, 68 (41 from the control and 27 from the experiment group) indicated on returned questionnaires the frequency of reasons for requiring demonstrator assistance (see Appendix 4, Table A4 - 4). Although given the opportunity to do so, none of the students indicated reasons not covered by the operationalized categories. Reasons for this are discussed in section 9.12.

If available, the sample frame of 68 students would have required 416 demonstrator consults, an average of 6 per student. A notable difference in the average number of consults per student was found between the two groups of students: the experiment group required on average 5 consults per student (135 consults / 27 students) compared with 7 in the control group (281 consults / 41 students).

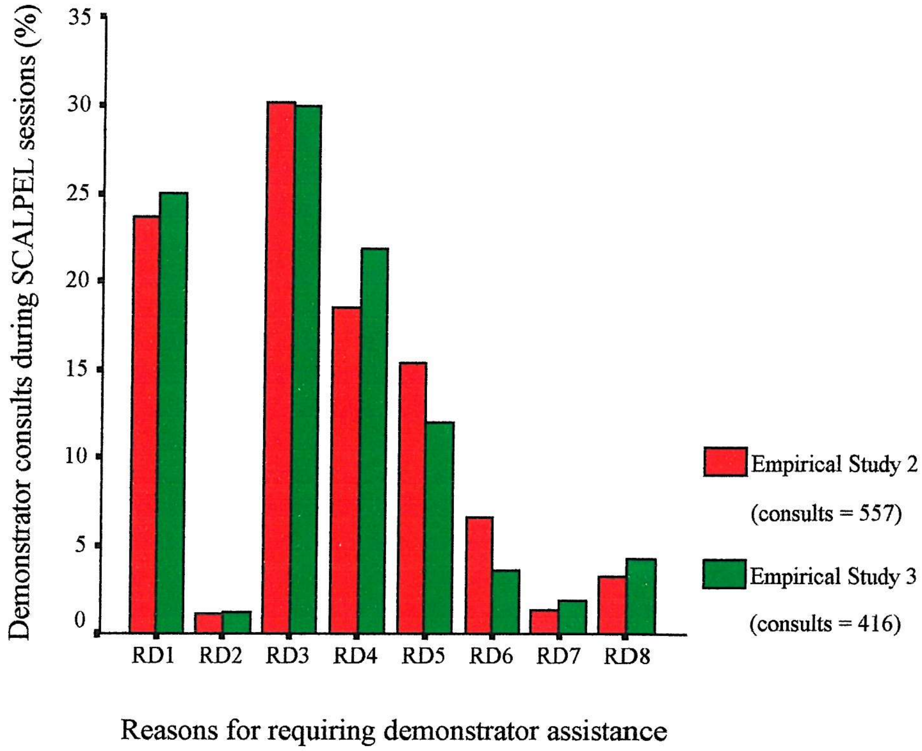


Figure 36 Bar chart on the distribution of required demonstrator consults during SCALPEL sessions. Empirical study 2 is based on the mean number of consults over 6 practicals. Empirical study 3 is based on practical 6 only.

Presented graphically in Figure 36, the distribution of the entire sample was very similar to the pattern highlighted in the previous study with the majority (92.5%) of reasons for demonstrator assistance being to discuss the pathology material. This was also the case for the distributions when the sample was presented as the control and experiment groups (see Figure 37). A notable exception was that the 5 required consults about the multiple choice window (2.1% of the entire sample, 3.6% of the experiment group consults) were confined to the experiment group. Problems in this category for the experiment group were expected since the new MCQ/Dialogue agent available to these students was integral to, and changed the functionality of the multiple choice engine (see chapter 10).

The importance placed by students on the necessity for demonstrators to be available to address problems was similar to the ranking observed in the previous study (see section 9.12). The rounded mean score of the sum of all ranks was taken to reflect the relative position of a category (see Appendix 4, Table A4 - 5) on a scale of 1 (most important) to 10 (least important). Ordering the categories by their mean importance for the entire sample ranked those representing domain material (RD3, RD1, RD4 and RD5) on the lower portion of the scale (most important), and those representing issues regarding application functionality (RD8, RD2 and RD7) at the higher end (least important). As expected, this observation conforms to the data on the number of required demonstrator consults, since it is reasonable to expect students to place more importance on demonstrators to be available to address their more frequent misconceptions.

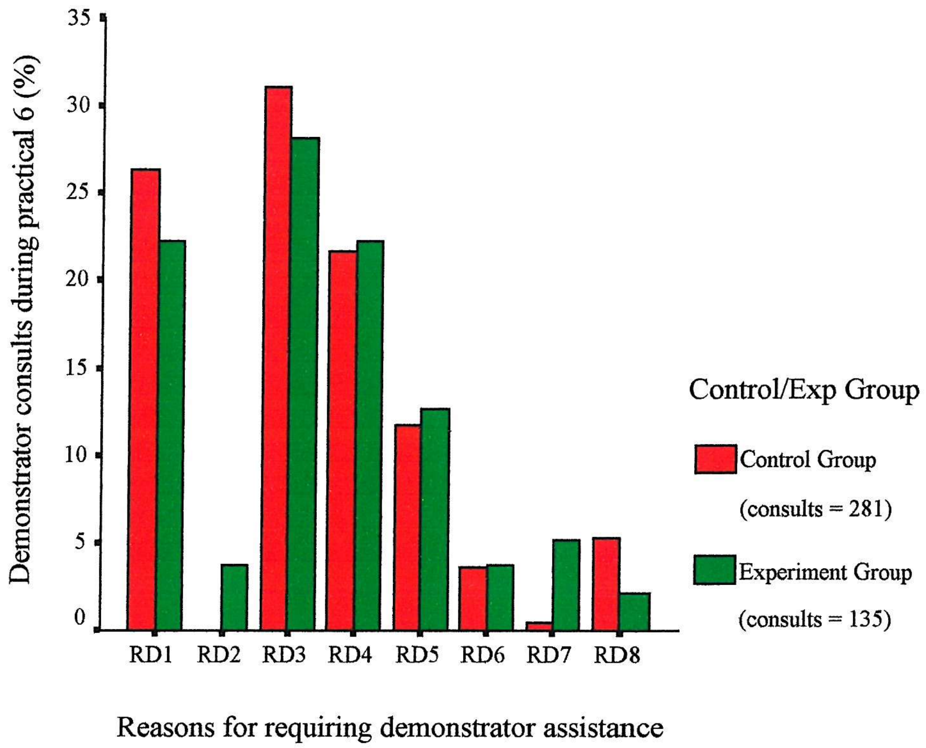


Figure 37 Bar chart on the distribution of required demonstrator consults during practical 6 for the control and experiment groups.

11.12 Solving Problems in the Absence of Demonstrators

With the population of 68 students requiring an estimated 416 demonstrator consults during practical 6, and with 92.5% of these consults concerning the pathology material, this section looks at the methods used by students in the control and experiment groups to solve problems in the absence of demonstrators. None of the students offered additional methods to the alternatives categorised in section 9.3.3.2.

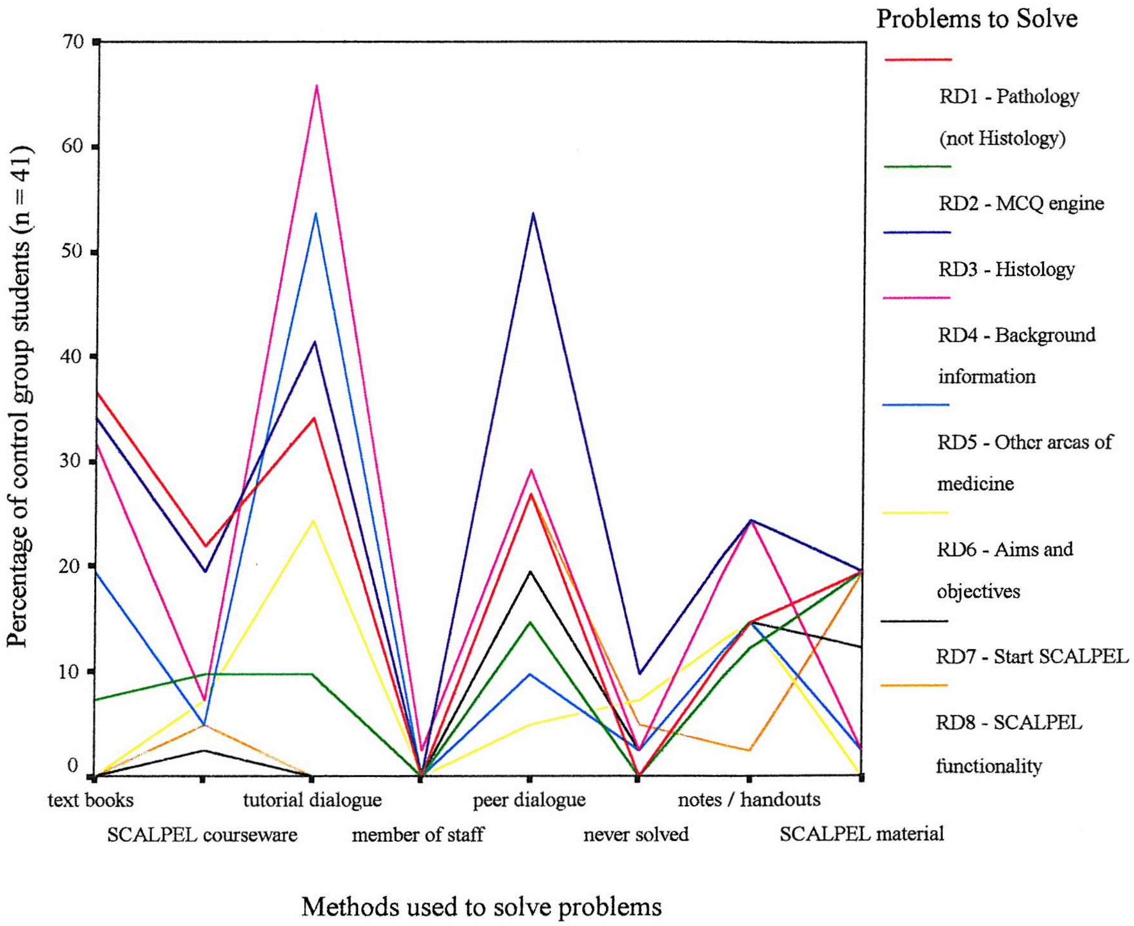


Figure 38 Percentage of control group students who used methods to solve problems encountered during practical 6 (for data see Appendix 4, Table A4 - 6).

Figure 38 shows the percentage of students in the control group (N = 41) who, in the absence of demonstrators, solved problems during SCALPEL sessions by using the identified methods. Spikes in the graph indicate that the use of text books, tutorial discussions and peer dialogue were the more favourable methods employed by students. It is also apparent that pathology problems (RD1, RD3, RD4 and RD5) were principally solved by these methods. Students in the experiment group, who had functionality available to address problems with the pathology material, exhibited a different pattern. Presented in Figure 39, spikes in the graph highlight a notable effect of the improved SCALPEL functionality with a higher percentage of the sample opting to use SCALPEL to solve problems with the pathology material (RD1, RD3, RD4).

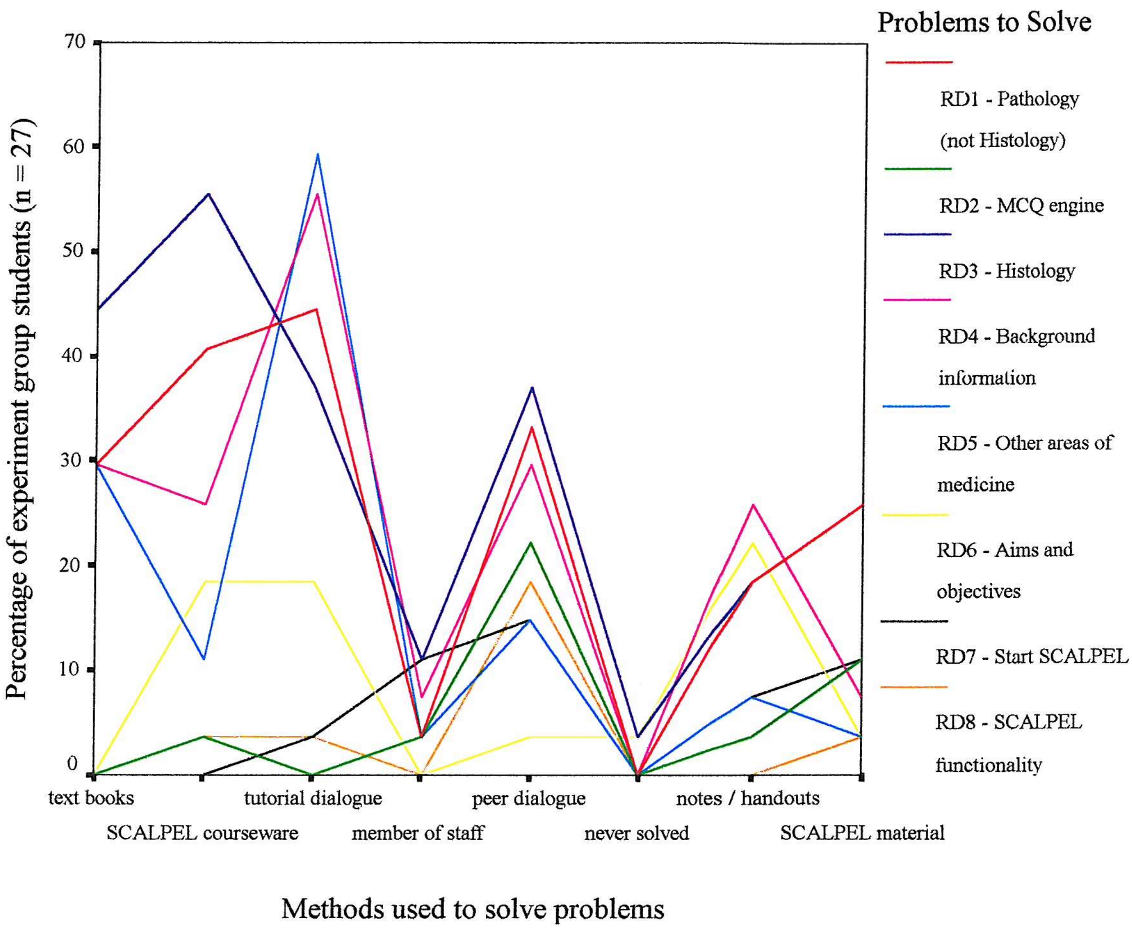


Figure 39 Percentage of experiment group students who used methods to solve problems encountered during practical 6 (for data see Appendix 4, Table A4 - 7).

Looking at the percentage of students who at least on one occasion experienced a problem for which a solution was never found, 9.8% of students in the control group (N = 41) expressed trouble in finding help to discuss Histology material (RD3). This was followed by 7.3% of the sample experiencing difficulty in clarifying the aims and objectives of the practical. In comparison, a single student (3.7%, N = 27) in the experiment group experienced a similar problem for each of these problem categories.

11.13 Possible Application Improvements - what the students want

This section looks at what students believed to be the most important of the proposed improvements to the SCALPEL courseware.

Although given the opportunity to do so, none of the students indicated improvements to the courseware not covered by the list of alternatives operationalized in section 9.3.3.3. The rounded mean score of the sum of all ranks was taken to reflect on a scale of 1 (most important) to 10 (least important) the importance of an improvement relative to the other proposals (see Appendix 4, Table A4 - 8). Functionality offering on-line text books, a glossary of terms, the ability to ask questions and receive answers about pathology material, and more links to background information about the pathology material, was considered by students to be the most important improvements to SCALPEL. Corresponding with the patterns observed in the data describing the role of demonstrators (section 11.11) and the methods used to solve problems given demonstrator absence (section 11.12), students placed more importance on including functionality to address the pathology material and not SCALPEL functionality.

11.14 Discussion

This study was informed by the empirical evidence presented by the previous evaluations of SCALPEL (see chapters 7, 8 and 9). A common theme held constant during each of the studies was students' acceptance of SCALPEL. The influence of students' attitudes towards demonstrator assistance during exposure to SCALPEL on these attitudes was first highlighted during the first study, and was reinforced by the second. Observing that students who expressed less of a need for demonstrator assistance during SCALPEL sessions asserted a higher degree of acceptance of SCALPEL, and that attitudes towards demonstrators were influenced by students being able to solve problems in the absence of demonstrators, it was reasoned that improving SCALPEL to help students resolve misconceptions would have a positive effect on student attitudes. This phase of the research was designed to compare the attitudes of students who had exposure to improved SCALPEL functionality (experiment group) with the attitudes of students for whom the improved functionality was unavailable (control group).

Beginning with a look at students' attitudes towards SCALPEL, taken as a complete sample, the distribution of attitudes highlighted a positive difference relative to the previous evaluations, with none of the sample expressing negative attitudes. However, breaking the sample into the control and experiment groups highlighted a significant difference in distributions ($U = 492.5$, $Z = -1.90$, $p = 0.05$). A higher percentage of the experiment group (89.3%) expressed positive attitudes (agree and strongly agree categories) than the control group (60.9%), with the difference explained by the percentage of students in the neutral category (10.7% for the experiment group and 39.1% for the control group). The overall distribution of student attitudes towards demonstrators also improved relative to the previous studies, with students expressing less of a need for demonstrator assistance in this study. Comparing the distribution of this attitude for the control and experiment groups found no significant difference ($U = 591.0$, $Z = -0.14$, $p = 0.89$). However, a higher percentage of the control group (31.1%) missed demonstrator assistance compared to the experiment group (10.7%). An almost identical percentage in the neutral category left the difference to be explained by a higher percentage of the experiment group (57.1% compared to 31.1% in control group) expressing that they did not miss demonstrators. The difference in the mean ranking of these attitudes ($U = 475.0$, $Z = -1.845$, $p = 0.65$) was just short of the 0.05 alpha level (α), leading to a possible explanation that although students in both groups expressed an equal need for demonstrator assistance, those in the experiment group didn't miss demonstrators as much as a result of their problems being addressed by the improved SCALPEL functionality available to them.

Turning to the factual data for evidence to support this supposition, the average number of required demonstrator consults per student was greater in the control group (7 per student) than in the experiment group (5 per student). Taking into consideration that higher percentages of students in the control group expressed that at least on one occasion problems were experienced for which a solution was never found, it can be reasoned that students in the experiment group faired better at solving problems. In addition, the observation that higher percentages of students in the experiment group were opting to use SCALPEL to solve problems is testimony that the improved functionality had a positive impact on students' ability to resolve misconceptions during pathology practicals. With all other variables held constant between the two groups, it is reasonable to hypothesise that the differences in attitudes towards demonstrators and SCALPEL are attributed to the implementation of improved functionality.

11.15 Conclusions

The evidence presented in this chapter supports the primary hypothesis that students who use SCALPEL's MCQ/Dialogue agent express agreeable attitudes towards SCALPEL compared with those students to whom the agent functionality is unavailable. A notable positive difference was observed in the distribution of student attitudes towards SCALPEL between the experiment and control groups. An additional difference was observed between the distribution of attitudes towards SCALPEL and the distribution of students who missed demonstrators during pathology practical 6. Following on from this, the average number of required demonstrator consults was higher in the group of students to whom the MCQ/Dialogue agent was unavailable, with more problems remaining unsolved. High percentages of students using SCALPEL to resolve misconceptions in the group that had the MCQ/Dialogue agent available illustrates the positive impact of this functionality on the pathology learning experience. Taking into consideration that all other variables remained constant between the two groups, the positive differences in student attitudes can be attributed to implementation of the MCQ/Dialogue agent.

In conclusion, an improvement in student attitudes towards SCALPEL and the need for demonstrator assistance relative to the previous evaluations was observed in the entire sample. A closer look at the distribution of attitudes between those students who worked through practical 6 using the MCQ/Dialogue agent and those students to whom the agent was unavailable highlighted positive differences in attitudes. Reasons for this were attributed to a reduction in required demonstrator consults, and an increase in problems solved as a consequence of the MCQ/Dialogue agent helping students to resolve misconceptions with the pathology material.

11.16 Summary

This chapter presented the third evaluation of SCALPEL. The entire research process was discussed relative to this phase of the research before presenting the results of the analytical and descriptive analysis. Finally a discussion was offered on the interpretation of the findings and their impact on the research presented in this thesis.

Chapter 12: Conclusion and Future Work

12.1 Research Synopsis

The research presented in this thesis constitutes work carried out as part of the Southampton Computer Assisted Learning Pathology Education Laboratory (SCALPEL) project. The project was charged with researching the impact of replacing traditional pathology laboratory teaching with pedagogic courseware.

12.1.1 Theoretical Influences on Education

The first stage of the research centred on investigating the use of technology in education. This began with a look at the theories that have underpinned contemporary moves to introduce learning systems. From the literature it was observed that educational technology models are broadly categorised as being influenced by either behaviourism or constructivism; the former is generally considered as subscribing to outdated theories of knowing and informs the majority of learning systems in higher education, while the latter has had more of an impact in schools (Hawkridge 1996). Although constructivist theories of knowing are now generally considered to inform the best approaches to learning, some (Carroll 1990) believe that such views are not evolutionary but cyclic, leading to instantiations of individualistic learning environments empirically driven by observations of 'what people want to do, how they want to do it, and the problems they are observed to experience' (Carroll 1990).

Carroll's comments make good advice. Although an understanding of theories of knowing is important, surely what is more important is to develop a learning environment based on an understanding of the learning goals for that individual learning task or course. This holds true for both conventional and computer-based learning. In the context of SCALPEL, the case-based pathology practicals had already proven to be successful; not only the learning goals but the entire learning experience had been integral to students' pathology learning in a laboratory environment. For computer-based learning environments an added consideration is the availability of appropriate technologies. For the purposes of this research an understanding was required of the technologies that could afford the transition from the laboratory to a computer-based pathology practical learning environment.

12.1.2 Choosing an Enabling Technology

The decision about the choice of technology was made after considering the format of the existing laboratory pathology practicals. Different media formats employed throughout and the implicit structuring of a practical pointed to the use of a hypermedia learning environment. This decision was further informed by the fact that hypermedia, as a consequence of supporting a variety of learning settings, is now seen as the dominant force in educational technology. From the myriad of available hypermedia systems Microcosm was chosen on the basis of the functionality offered, the extensive literature to support its use in education, and the support offered to developers and users of Microcosm-based learning environments at the University of Southampton.

12.1.3 Evaluating SCALPEL for the first time - a focus on pedagogic value

Having authored the SCALPEL courseware, an understanding was required of its impact on undergraduate medical students' education. The first empirical study was primarily charged with assessing the pedagogic value of SCALPEL in terms of education attainment and students' attitudes towards this method of delivering pathology practicals. A curiosity driven research agenda was also designed to investigate 'what students bring to learning': student-specific characteristics that may influence the success or otherwise of SCALPEL.

Focusing on the pedagogic value of SCALPEL relative to the traditional pathology laboratory, no statistical difference was found between the education attainment of two groups of students. Further analyses showed that on the whole individual students performed better in both methods of assessment at the end of study, therefore suggesting no bias towards a particular orientation of pathology practicals. These results gave confidence that pathology practicals are equally educationally effective, in terms of domain material, when delivered in the laboratory or by SCALPEL. Students' attitudes towards SCALPEL were not so encouraging with the majority of students expressing a neutral or dissatisfactory attitude. However, no evidence of a relationship between students' education attainment and attitudes suggested that SCALPEL was an effective pedagogic environment even for students who didn't express an agreeable attitude towards working with SCALPEL.

Turning to the curiosity driven research, three concepts were identified as being potentially influencing on the pedagogic value of SCALPEL. These were students' expressed need for demonstrator assistance, attitudes towards computers, and behavioural learning styles. No relationships were observed between the constituent concepts of pedagogic value and, learning styles and attitudes towards computers. In addition no relationship was found between students' expressed need for demonstrators during SCALPEL sessions and education attainment. However, a significant moderately negative relationship with acceptance of SCALPEL suggested that: students who expressed a need for demonstrators during exposure to SCALPEL indicated a lower acceptance of SCALPEL than those who suggested a lesser need. The distribution of student attitudes towards demonstrators and the findings of a focus group highlighted a student dependence on demonstrators whilst working through pathology practicals. This finding later proved to be significant in deciding upon functionality to improve SCALPEL.

With an overwhelming negative attitude towards SCALPEL, further relationships were investigated involving identified student-specific characteristics. These characteristics were categorised as: student demographics - age, gender and demographic area of post-primary education; socio-economic characteristics - previous higher education and access to computers; and behavioural characteristics - previous computer use. No relationships were found between students' attitudes towards SCALPEL and the conceptualised dimensions of each of these characteristics. This inferred that individual characteristics are less of a factor in the successful use of courseware applications, with more heterogeneous user populations being observed.

12.1.4 A focus on what students bring to learning

The following concepts were identified as being of particular interest to this research: students' expressed need for demonstrator assistance during SCALPEL sessions, attitudes towards computers, and learning styles. Analyses were carried out to look for associations between these concepts and other student-specific characteristics.

No relationships were observed between students' attitudes towards demonstrator assistance and demographic, socio-economic and behavioural characteristics. This was the same for the concept of attitudes towards computers. These findings infer that attitudes towards demonstrator assistance are formed during and/or after working through the practicals, and are not biased by characteristics brought by students. A significant relationship was observed with the socio-economic dimension of access to computers, inferring that accessibility to computers improves attitudes towards computers, which in turn inspires access to computers. Turning to students' behavioural characteristics, no relationships were found with two dimensions of previous computer use: the Internet and computers at work. However, weak significant relationships were found with formal learning courseware and use of computer games, and a moderate relationship was found with previous use of word processing packages. Measuring student attitudes pre and post exposure to SCALPEL found no significant changes in attitudes, leading to the conclusion that exposure to SCALPEL had no systematic effect on attitudes towards computers.

Turning to students' learning styles, no relationships were observed between this concept and the student demographic socio-economic and behavioural characteristics. However a look at changes in student learning styles pre and post exposure to SCALPEL indicated a significant shift in favour of exploratory learning.

12.1.5 Evaluating SCALPEL a second time - a focus on demonstrators

A second empirical study of SCALPEL was performed with two significant differences informing its design. First, the series of pathology practicals under evaluation (the same six practicals used in the first study) were taught entirely using SCALPEL. Second, demonstrators were not available during the SCALPEL sessions.

This study's focus was again on students' acceptance of SCALPEL, but also on the role of demonstrators during pathology practicals. Relative to the first study, students expressed improved attitudes towards SCALPEL and less of a need for demonstrator assistance. A significant relationship was again observed between these two attitudinal concepts. In the case of attitudes towards demonstrators, students who managed to solve all problems encountered during the pathology practicals expressed less of a need for demonstrators than those students who experienced unsolved misconceptions. For attitudes towards computers, a statistical difference was highlighted between these categories of students, with the distribution suggesting that students who solved all problems expressed a higher degree of acceptance of SCALPEL. In addition, the overall low percentage of unsolved problems inferred that the improved student attitudes were as a consequence of students being pro active in solving problems encountered during the practicals. The data also showed that the SCALPEL courseware was a relatively ineffective method of solving problems with students opting for text books and tutorial/peer dialogue. Furthermore, students placed more importance on improvements to SCALPEL that would replicate the methods they most frequently used during this evaluation, namely, on-line text books and the ability to ask questions and receive answers about the pathology material. This empirical evidence underlined the need for improved SCALPEL functionality that would take away from students the onus of having to look outside of SCALPEL for solutions to pathology problems. If was inferred that addressing student needs for demonstrator assistance by offering effective SCALPEL functionality could result in a marked improvement in students' acceptance of this method of delivering pathology practicals.

12.1.6 Improving SCALPEL with a dialogue model

The first implementation of SCALPEL was limited in the way it offered help to students when faced with misconceptions, or offered help to reinforce students' correct understanding of concepts. Within SCALPEL the only means of offering students help in resolving misconceptions was through SToMP's MCQ engine offering relevant supplementary domain material. Observations of demonstrator assistance in the traditional laboratory, empirical data from the two empirical evaluations of SCALPEL, and the education literature, informed the decision to look into ways of incorporating into SCALPEL a means of mediating student-demonstrator and peer dialogue.

Improvements to SCALPEL were seen as including functionality that would mediate student-demonstrator and peer dialogue about context sensitive concepts of any knowledge domain authored into the learning environment. Reasoning that previous discussion threads and/or domain material could already be available to satisfy any given query, a secondary consideration was that raising a new line of enquiry should be the last resort in resolving problems during a pathology practical.

After looking at existing instantiations of learning environments that facilitate dialogue, a MCQ/Dialogue agent was developed that allows students to tap into a resource of previously asked questions relevant to the concepts being addressed, and in the correct context. A concept hierarchy was determined to be the ideal structure for presenting contextualised concepts within the domain. By taking responsibility for attaching a relevant concept and context to an enquiry, students benefit by being presented with relevant past discussions about the concept they are currently addressing, while demonstrators benefit by being informed of the concept and context that underpins any new line of enquiry.

12.1.7 Evaluating the MCQ/Dialogue agent

After designing the MCQ/Dialogue agent, a proof-of-concept application was implemented for evaluation relative to the first implementation of SCALPEL. The design of the evaluation was informed by the second empirical study but differed in two significant aspects. First, the evaluation techniques were designed to focus on a single practical, practical 6, and not the entire series of practicals used in the previous evaluations. Second, SCALPEL was amended to include the MCQ/Dialogue agent functionality for use with an experimental group of students.

The primary objective of this phase of the research was to assess the effectiveness of the MCQ/Dialogue agent. The attitudinal concepts 'acceptance of SCALPEL' and 'attitudes towards the need for demonstrators', and their relationship, continued to underpin this research. Curiosity driven research was charged with highlighting possible reasons for the distribution of these attitudinal concepts and their observed relationship. A secondary objective was the continuing need to improve SCALPEL.

The analyses showed that students who used the MCQ/Dialogue agent expressed agreeable attitudes towards SCALPEL compared with those students to whom the agent was unavailable. Notable positive differences were observed between the two student samples in the distribution of student attitudes towards SCALPEL and demonstrator unavailability.

The average number of required demonstrator consults was higher in the group of students to whom the MCQ/Dialogue agent was unavailable, with more problems remaining unsolved. High percentages of students using SCALPEL to resolve misconceptions in the group that used the MCQ/Dialogue agent illustrates the positive impact of this functionality on the pathology learning experience. Taking into consideration that all other variables remained constant between the two groups, the positive differences in student attitudes can be attributed to the implementation of functionality designed to resolve student misconceptions within the courseware domain.

In conclusion, an improvement in student attitudes towards SCALPEL and the need for demonstrator assistance relative to the previous evaluations was observed in the entire sample. A closer look at the distribution of attitudes between those students who worked through practical 6 using the MCQ/Dialogue agent and those students to whom the functionality was unavailable highlighted positive differences in attitudes. Reasons for this were attributed to a reduction in required demonstrator consults, and an increase in problems solved as a consequence of the MCQ/Dialogue agent helping students to resolve misconceptions with the pathology material.

12.2 Future Work

Future work on SCALPEL can be categorised into three areas: improvements to SCALPEL functionality relative to its use as a Microcosm application; evaluation of the MCQ/Dialogue agent in different domains; and research into more fundamental changes to SCALPEL's functionality in terms of breakthroughs in technology.

12.2.1 Improving SCALPEL as a Microcosm application

The MCQ/Dialogue agent was implemented as a proof-of-concepts application. Although the primary dialogue functionality was realised additional functionality is required to facilitate the authoring and maintenance of pathology practicals.

To take advantage of the MCQ/Dialogue functionality within SCALPEL requires the co-ordination of concept ID allocation and context ID determination. When a new practical is authored, two components of the MCQ/Dialogue agent need to be configured. A new configuration script needs to be authored and the concept hierarchy needs to be amended to reflect the addition of new concepts. For the purposes of this research this process (described in chapter 10) was performed manually. It is not acceptable for the manual approach to be used by pathology practical authors on a continued basis as it leaves open the opportunity for inconsistencies with concept and context IDs. Such inconsistencies would render the dialogue functionality useless. A toolkit is therefore required to semi-automate this process: a user interface that allows authors to include new concepts (terms, practicals and questions) on the concept hierarchy and facilitates the authoring of configuration scripts. The best approach would be to use a visual representation or map of the concept hierarchy. The user could then simply indicate with a few mouse clicks the correct position of the new concepts. This process would automatically generate unique concept IDs for each concept, determine the context IDs, and author these into the relevant configuration script.

A second interface is required for use by demonstrators. Again within the implementation used for the purposes of this research, all new student queries were dealt with manually. A demonstrator-answering interface would bring to a demonstrator's attention all queries that need to be answered. If different demonstrators are available for different terms, practicals or even questions, queries specific to these concepts can be routed accordingly. This has the advantage of distributing the load across a number of demonstrators or perhaps more importantly, routing queries to experts who are recognised as having specific skills.

When implementing a toolkit the potential of 'domain portability' needs to be considered. In the context of SCALPEL, a toolkit is required to assist with the authoring and maintainability of pathology practicals and their constituent MCQ questions and discussion threads. However, as soon as a tool kit is developed the potential of porting the functionality to other domains is compromised. This is because the functionality, by definition of servicing user requirements, becomes specific to the pathology environment. Consequently, when considering a toolkit care needs to be taken that the underlying

MCQ/Dialogue agent functionality is not tied to a specific implementation, e.g. pathology. Developing courseware toolkits that adhere to domain portability is a logical progression to the work presented in this thesis, and remains an area of research that could have a significant impact on the use of computer-based learning environments.

12.2.2 Using SCALPEL in different domains

The implementation and evaluation of the SCALPEL approach is specific to pathology practicals at the University of Southampton. Caution is accordingly advised when generalising the work in this thesis to other domains. However, conceptually, it is reasoned that the approach adopted by SCALPEL is portable to other domains where the structure of the material is case-based. To test this assumption and to improve on the empirical evidence in support of SCALPEL, the implementation and evaluation of SCALPEL in other domains is desirable and forms a valuable exercise for future research.

12.2.3 SCALPEL and breakthroughs in technology

Educational technology is strongly influenced by hypermedia. The World Wide Web is the predominant hypermedia medium in widespread use, with higher education practitioners in particular rushing to produce online versions of their courses. It is difficult amongst such euphoria not to view the Web as the Holy Grail of educational technology. However, without large amounts of resources being used to develop each instantiation of a learning environment, the educational benefits of the Web are often very limited. From a hypermedia perspective such limitations are grounded in the Web's inability to support multiple linking strategies (i.e. local, generic, dynamic). Webcosm goes some way in tackling this problem by removing embedded links from documents and storing them in separate link databases. In this form functionality can be authored to provide the linking benefits that are found in Microcosm, e.g. show and computed links. From a multimedia perspective, breakthroughs in the conceptual indexing of audio/visual media are required to improve the retrieval accuracy of relevant material for a given knowledge domain.

Coupled with artificial intelligence, which is once again making progress under the guise of software agents, it is not unrealistic to foresee Web-based learning environments in which students have personalised agents that accurately model their individual knowledge state, and interact with virtual demonstrators and other students to dynamically manipulate links to domain material that is extremely sensitive to their individual needs. Embracing the ever increasing user base of the Web, intelligent software agents that are capable of autonomous communication offer the potential of bringing together learning resources worldwide. Realising such a vision, pathology students at Southampton would not be restricted to the viewpoints offered to them by their local domain experts, but would be able to take advantage of resources that have always been available, but out of reach.

It is inevitable that SCALPEL will be ported to the Web. It is important that when this happens, functionality is not lost but improved.

12.3 Project SCALPEL - A Personnel Perspective

The research presented in this thesis was not motivated by the need to replace or improve an existing learning environment of poor pedagogic value. Indeed, the traditional laboratory environment was an already proven and successful way of delivering pathology practicals. Rather, this research was driven by the continued need to deal with rising staff-student ratios, exacerbated by the logistical and financial problems of accommodating a growing student base in a traditional laboratory.

The initial primary research agenda focused on satisfying the medical faculty that replacing the traditional laboratory with computer-based pathology practicals would not be detrimental to student learning. This prompted the inclusion of research measures to assess students' education attainment. Having presented data highlighting no significant difference in the education attainment of students who worked in the traditional laboratory, and those who used SCALPEL, the research focus shifted away from education attainment and concentrated more on students' attitudes towards SCALPEL. Most empirical studies of educational technology focus on education attainment as the prominent 'success' indicator. During the evaluation of SCALPEL students' attitudes towards SCALPEL were considered to be more advantageous to determining the success of SCALPEL as a vehicle for delivering pathology practicals. This view was taken as a consequence of past empirical studies continually reporting no significant improvement in the education of students who have exposure to computer-based learning environments.

From a personal perspective, I did not expect SCALPEL to radically improve test results, I saw the challenge as one of implementing a learning environment that engaged students' interest, with the consequence of realising an educational tool that would be repetitively used by students on a voluntary basis. It was not the pathology domain material that was changing during the transition from the laboratory to the computer, it was the enabling environment. Given that there would be no way of enforcing student participation in computer-based practicals, the environment needed to be 'accepted' by students as being worthy of their time. A key consideration here was that any instantiation was going to 'replace' and not 'augment', as is commonly the case, an existing teaching practice. This meant that the courseware was not going to be deposited as a supplementary learning resource in a library, available to students if deemed necessary. Rather it would constitute the only way for students to strengthen their understanding of pathology concepts introduced in earlier lectures. To this end the variables considered to be of interest to this research centred on individual characteristics and identified aspects of the traditional laboratory, and their relationship with students' acceptance of SCALPEL. The reasoning being that if individual or groups of students could be highlighted as sharing positive or negative attitudes towards SCALPEL, targeted preparation and support could be offered to support their working through the computer-based practicals. In addition, if students expressed attitudes towards identified facets of the traditional laboratory, as was the case with students expressing a need for demonstrator availability, research could be carried out into ways of offering effective functionality to satisfy such requirements.

Listening to student opinion and acting upon what they expressed as being important to them during pathology practicals proved to be rewarding in improving attitudes towards SCALPEL. Highlighted in the first empirical study, attitudes towards the need for demonstrator availability was overwhelming, with the added interest that the distribution of these attitudes was significantly related to the distribution of attitudes towards the acceptance of SCALPEL. The subsequent research and implementation of functionality to facilitate student-demonstrator dialogue resulted in the collection of data showing that such functionality reduces students' need for demonstrator assistance, with the subsequent effect of improving attitudes towards SCALPEL.

Established research techniques and the experiences of past empirical studies informed the research presented in this thesis. Time and money restraints invariably contribute to the choice of research techniques for any given project. This means that best practice is defined by the variables specific to each research initiative. To a large extent the evaluation of SCALPEL was dictated by the availability of students. Given that this project was a 'live' project, i.e. the evaluation was integral to and not supplementary to students' medical education, a primary consideration was that evaluations had to coincide with the timetable of academic terms and more specifically, pathology practicals. This meant that the research, design and implementation of both the SCALPEL courseware and the appropriate evaluation instruments had strict timetable constraints. A compromise is often required, as was the case during this research, between employing satisfactory empirical measures that stand up to critical scrutiny, whilst working within strict financial and logistical constraints.

12.4 Conclusion

From a holistic viewpoint, this thesis communicates a case study of replacing a traditional case-based teaching environment with a hypermedia learning environment called SCALPEL. An initial empirical study showed that the education attainment of students who used SCALPEL was no worse than the education attainment of those students who worked in the traditional laboratory environment. Concerns were raised about SCALPEL's ability to help students resolve problems with the pathology material. Further investigations led to the implementation of dialogue functionality through a MCQ/Dialogue agent. A final empirical study highlighted a reduction in required demonstrator consults, and a notable improvement in students' acceptance of SCALPEL as an environment for working through pathology practicals.

Appendix 1: Quantitative Statistics of the First Empirical Study: Assessing the Pedagogic Value of SCALPEL

This appendix presents statistical data from the evaluation presented in chapter 7 - The First Empirical Study: Assessing the Pedagogic Value of SCALPEL.

Table A1 - 1 Distribution of students by age group

Value	Frequency	Percent	Valid Percent	Cumulative Percent
under 20	123	78.3	78.3	78.3
20 - 25	27	17.2	17.2	95.5
26 - 35	7	4.5	4.5	100.0
over 35	0	0	0	100.0
Sample Total	157	100.0	100.0	

Age group was operationalized by asking students to indicate their age group on a questionnaire administered before exposure to SCALPEL. The sample population (sample total) constituted all first year medical students who indicated their age on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A1 - 2 Mean difference in MCQ test results between students who worked in the traditional laboratory and those who had exposure to SCALPEL during practicals 1-3

Variable	Exposure	N	Mean Mark	SD	SE of Mean
mcqtest1	LAB	77	7.68	5.10	0.58
	SCALPEL	77	6.97	4.78	0.54

Mean Difference = 0.70

Variances	t-value	df	2-tailed Sig	SE of Diff	95% CI for Diff
Equal	0.88	152	0.38	0.80	(-0.87, 2.27)

The MCQ test (mcqtest1) was administered to students at crossover (after practical 3). Half the students constituted the control group by working through the practicals in the traditional laboratory (exposure = LAB) while the experimental group had exposure to the SCALPEL courseware (exposure = SCALPEL). The mean marks out of 20, standard deviation (SD) and standard error (SE) of the test results are expressed for both groups. The data were analysed by independent samples t-test for the equality of means expressing: the mean, t value, degrees of freedom (df), two-tailed significance, standard error of the mean difference (SE of Diff) and 95% confidence intervals for the difference (95% CI for Diff).

Table A1 - 3 Mean difference in short answer test results between students who worked in the traditional laboratory and those who had exposure to SCALPEL during practicals 1-3

Variable	Exposure	N	Mean Mark	SD	SE of Mean
satest1	LAB	61	3.13	1.98	0.25
	SCALPEL	67	3.55	2.38	0.29

Mean Difference = -0.42

Levene's Test for Equality of Variances: F = 3.96, P = 0.49

Variances	t-value	df	2-tailed Sig	SE of Diff	95% CI for Diff
Unequal	-1.09	124.98	0.28	0.39	(-1.18, 0.34)

The short answer test (satest1) was administered to students at crossover (after practical 3). Half the students constituted the control group by working through the practicals in the traditional laboratory (exposure = LAB) while the experimental group had exposure to the SCALPEL courseware (exposure = SCALPEL). The mean marks out of 10, standard deviation (SD) and standard error (SE) of the test results are expressed for both groups. Levene's test for equality of variances indicates the ratio of the two sample variances (F) and the probability (P) of the difference occurring in the underlying population. The data were analysed by independent samples t-test for the equality of means with unequal variances expressing: the mean difference, t value, degrees of freedom (df), two-tailed significance, standard error of the mean difference (SE of Diff) and 95% confidence intervals for the difference (95% CI for Diff).

Table A1 - 4 Mean difference in MCQ test results between students who worked in the traditional laboratory and those who had exposure to SCALPEL during practicals 4-6

Variable	Exposure	N	Mean Mark	SD	SE of Mean
mcqtest2	SCALPEL	75	11.71	3.86	0.45
	LAB	76	11.79	4.35	0.50

Mean Difference = -0.08

Variances	t-value	df	2-tailed Sig	SE of Diff	95% CI for Diff
Equal	-0.12	149	0.90	0.67	(-1.41, 1.24)

The MCQ test (mcqtest2) was administered to students at the end of the study (after practical 6). Half the students constituted the control group by working through the practicals in the traditional laboratory (exposure = LAB) while the experimental group had exposure to the SCALPEL courseware (exposure = SCALPEL). The mean marks out of 20, standard deviation (SD) and standard error (SE) of the test results are expressed for both groups. The data were analysed by independent samples t-test for the equality of means expressing: the mean difference, t value, degrees of freedom (df), two-tailed significance, standard error of the mean difference (SE of Diff) and 95% confidence intervals for the difference (95% CI for Diff).

Table A1 - 5 Mean difference in short answer test results between students who worked in the traditional laboratory and those who had exposure to SCALPEL during practicals 4-6

Variable	Exposure	N	Mean Mark	SD	SE of Mean
satest2	SCALPEL	78	4.26	1.92	0.22
	LAB	77	4.61	1.89	0.21

Mean Difference = -0.35

Variances	t-value	df	2-tailed Sig	SE of Diff	95% CI for Diff
Equal	-1.16	153	0.25	0.31	(-0.96, 0.25)

The short answer test (satest2) was administered to students at the end of the study (after practical 6). Half the students constituted the control group by working through the practicals in the traditional laboratory (exposure = LAB) while the experimental group had exposure to the SCALPEL courseware (exposure = SCALPEL). The mean marks out of 10, standard deviation (SD) and standard error (SE) of the test results are expressed for both groups. The data were analysed by independent samples t-test for the equality of means expressing: the mean difference, t value, degrees of freedom (df), two-tailed significance, standard error of the mean difference (SE of Diff) and 95% confidence intervals for the difference (95% CI for Diff).

Table A1 - 6 Mean difference in MCQ test results of students who worked in the traditional laboratory during practicals 1-3 and had exposure to SCALPEL during practicals 4-6 (exposure order LAB-SCALPEL)

Variable	N	Mean Mark	SD	SE of Mean	Corr	2-tailed Sig
mcqtest1	74	7.61	5.06	0.59	0.47	0.000
mcqtest2	74	11.73	3.88	0.45		

Paired Differences

Mean	SD	SE of Mean	95% CI for Diff	t-value	df	2-tailed Sig
-4.12	4.70	0.55	(-5.21, -3.03)	-7.55	73	0.000

The first MCQ test (mcqtest1) was administered at crossover, the second (mcqtest2) at the end of the study. The sample pairs (N) consisted of students who worked in the traditional laboratory before crossover and completed both MCQ tests. The mean, standard deviation (SD) and standard error (SE) of both test results are expressed together with their correlation (Corr) and 2-tailed significance. The data were analysed by paired-samples t-test for the mean difference expressing: the mean difference, standard deviation, standard error, 95% confidence intervals for the difference (95% CI for Diff), t value, degrees of freedom (df) and two-tailed significance.

Table A1 - 7 Mean difference in short answer test results for students who worked in the traditional laboratory during practicals 1-3 and had exposure to SCALPEL during practicals 4-6 (exposure order LAB-SCALPEL)

Variable	N	Mean Mark	SD	SE of Mean	Corr	2-tailed Sig
satest1	61	3.13	1.98	0.25	0.15	0.25
satest2	61	4.30	1.80	0.23		

Paired Differences

Mean	SD	SE of Mean	95% CI for Diff	t-value	df	2-tailed Sig
-1.16	2.47	0.32	(-1.80,-0.53)	-3.68	60	0.001

The first short answer test (satest1) was administered at crossover, the second (satest2) at the end of the study. The sample pairs (N) consisted of students who worked in the traditional laboratory before crossover and completed both short answer tests. The mean, standard deviation (SD) and standard error (SE) of both test results are expressed together with their correlation (Corr) and 2-tailed significance. The data were analysed by paired-samples t-test for the mean difference expressing: the mean difference, standard deviation, standard error, 95% confidence intervals for the difference (95% CI for Diff), t value, degrees of freedom (df) and two-tailed significance.

Table A1 - 8 Mean difference in MCQ test results for students who had exposure to SCALPEL during practicals 1-3 and worked in the traditional laboratory during practicals 4-6 (exposure order SCALPEL-LAB)

Variable	N	Mean Mark	SD	SE of Mean	Corr	2-tailed Sig
mcqtest1	75	6.96	4.72	0.55	0.49	0.000
mcqtest2	75	11.81	4.37	050		

Paired Differences

Mean	SD	SE of Mean	95% CI for Diff	t-value	df	2-tailed Sig
-4.85	4.60	0.53	(-5.91,-3.80)	-9.14	74	0.000

The first MCQ test (mcqtest1) was administered at crossover, the second (mcqtest2) at the end of the study. The sample pairs (N) consisted of students who had exposure to SCALPEL before crossover and completed both MCQ tests. The mean, standard deviation (SD) and standard error (SE) of both tests are expressed together with their correlation (Corr) and its 2-tailed significance. The data were analysed by paired-samples t-test for the mean difference expressing: the mean difference, standard deviation, standard error, 95% confidence intervals for the difference (95% CI for Diff), t value, degrees of freedom (df) and two-tailed significance.

Table A1 - 9 Mean difference in short answer test results for students who had exposure to SCALPEL during practicals 1-3 and worked in the traditional laboratory during practicals 4-6 (exposure order SCALPEL-LAB)

Variable	N	Mean Mark	SD	SE of Mean	Corr	2-tailed Sig
satest1	67	3.55	2.38	0.29	0.41	0.74
satest2	67	4.73	1.85	0.23		

Paired Differences

Mean	SD	SE of Mean	95% CI for Diff	t-value	df	2-tailed Sig
-1.18	2.95	0.36	(-1.90,-0.46)	-3.27	66	0.002

The first short answer test (satest1) was administered at crossover, the second (satest2) at the end of the study. The sample pairs (N) consisted of students who had exposure to SCALPEL before crossover and completed both short answer tests. The mean, standard deviation (SD) and standard error (SE) of both tests are expressed together with their correlation (Corr) and its 2-tailed significance. The data were analysed by paired-samples t-test for the mean difference expressing: the mean difference, standard deviation, standard error, 95% confidence intervals for the difference (95% CI for Diff), t value, degrees of freedom (df) and two-tailed significance.

Table A1 - 10 Distribution of students' acceptance of SCALPEL

Value	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Agree	4	2.5	2.8	2.8
Agree	29	18.5	20.6	23.4
Neutral	71	45.2	50.4	73.8
Disagree	35	22.3	24.8	98.6
Strongly Disagree	2	1.3	1.4	100.0
Valid Total	141	89.8	100.0	
Missing	16	10.2		
Sample Total	157	100.0		

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. The sample population (sample total) constituted all first year medical students who expressed an attitude on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A1 - 11 Measures of association between students' acceptance of SCALPEL and the results of the MCQ and short answer tests administered during the study

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
mcqtest1	69	0.09	0.94
satest1	62	-0.03	0.81
mcqtest2	70	0.11	0.37
satest2	71	0.09	0.44

An MCQ (mcqtest1) and short answer test (satest1) was administered at crossover and at the end of the study (mcqtest2, satest2). Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Each sample (N) was made up of students who had exposure to SCALPEL immediately prior to completing the relevant test, and expressed an attitude towards the acceptance of SCALPEL. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A1 - 12 Distribution of students' expressed need for demonstrators during SCALPEL sessions

Value	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Agree	15	9.6	10.9	10.9
Agree	64	40.8	46.4	57.2
Neutral	52	33.1	37.7	94.9
Disagree	7	4.5	5.1	100.0
Strongly Disagree	0	0	0	100.0
Valid Total	138	87.9	100.0	
Missing	19	12.1		
Sample Total	157	100.0		

Students' need for demonstrator assistance was operationalized by a multiple-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. The sample population (sample total) constituted first year medical students who expressed an attitude on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A1 - 13 Measures of association between students' expressed need for demonstrator assistance and the results of the MCQ and short answer tests administered during the study

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
mcqtest1	136	0.02	0.85
satest1	116	-0.03	0.74
mcqtest2	137	-0.14	0.09
satest2	138	-0.04	0.66

An MCQ (mcqtest1) and short answer test (satest1) was administered at crossover and at the end of the study (mcqtest2, satest2). Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. Each sample (N) was made up those students who completed the relevant test, and expressed an opinion towards the need for demonstrator assistance. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A1 - 14 Measures of association between students' pre-SCALPEL attitudes towards computers and the results of the MCQ and short answer tests administered during the study

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
mcqtest1	72	0.00	1.00
satest1	64	-0.01	0.95
mcqtest2	66	-0.14	0.26
satest2	69	0.01	0.90

An MCQ (mcqtest1) and short answer test (satest1) was administered at crossover and at the end of the study (mcqtest2, satest2). Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up of students who had exposure to SCALPEL immediately prior to completing the relevant test, and expressed an attitude towards computers. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A1 - 15 Measure of association between student attitudes towards computers and acceptance of SCALPEL

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
preattmi	127	0.05	0.59

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Student attitudes towards computers (preattmi) were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. The sample (N) was made up of students who expressed an attitude towards the acceptance of SCALPEL and computers. Spearman's rho correlation coefficient (ρ) and 2-tailed significance are expressed for the measure of association.

Table A1 - 16 Measures of association between students' pre-SCALPEL learning styles and the results of the MCQ and short answer tests administered during the study

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
mcqtest1	71	0.02	0.84
satest1	64	-0.11	0.40
mcqtest2	67	-0.14	0.27
satest2	70	0.08	0.51

An MCQ (mcqtest1) and short answer test (satest1) was administered at crossover and at the end of the study (mcqtest2, satest2). Student learning styles were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up those students who had exposure to SCALPEL immediately prior to completing the relevant test, and expressed a behavioural learning style. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A1 - 17 Measure of association between student learning styles and acceptance of SCALPEL

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
stylesmi	127	-0.10	0.26

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Student learning styles (stylesmi) were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. The sample (N) was made up of students who expressed an attitude towards the acceptance of SCALPEL and a behavioural learning style. Spearman's rho correlation coefficient (ρ) and 2-tailed significance are expressed for the measure of association.

Table A1 - 18 Measure of association between students' acceptance of SCALPEL and age

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
agegroup	141	-0.18	0.30

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. An indicator on a questionnaire administered before exposure to SCALPEL operationalized age (agegroup). The sample (N) was made up of students who expressed an attitude towards the acceptance of SCALPEL and indicated their age. Spearman's rho correlation coefficient (ρ) and 2-tailed significance are expressed for the measure of association.

Table A1 - 19 Measures of association between students' acceptance of SCALPEL and student demographics measured at the nominal level

Variable	N	Goodman-Kruskal tau	2-tailed Significance (p)
gender	141	0.06	0.68
schoolda	141	0.02	0.72

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Students' gender and demographic area of post-primary education (schoolda) were operationalized through indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up of students who expressed an attitude towards acceptance of SCALPEL and indicated their gender and post-primary education details. For each measure of association the Goodman-Kruskal tau and 2-tailed significance are expressed.

Table A1 - 20 Measures of association between students' acceptance of SCALPEL and social economic characteristics

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
hiedcode	141	-0.10	0.20
accscale	128	0.09	0.25

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Students' previous higher education (hiedcode) and access to computers (accscale) were operationalized through indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up of students who expressed an attitude towards acceptance of SCALPEL and indicated their previous higher education and access to computers. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A1 - 21 Measures of association between students' acceptance of SCALPEL and previous computer use

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
usagefl	136	0.06	0.43
usagegam	139	0.10	0.20
usageint	137	0.05	0.47
usagewp	140	0.08	0.31
usagewrk	133	0.15	0.07

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 17 indicators on a questionnaire administered at the end of the study. Operationalized by indicators on a questionnaire administered before exposure to SCALPEL, the previous use of computers for formal learning (usagefl), games (usagegam), internet (usageint), word processing (usagewp) and work (usagewrk) were measured. Each sample (N) was made up of students who expressed an attitude towards acceptance of SCALPEL and indicated their previous computer use for the relevant variable. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Appendix 2: Quantitative Statistics of the First Empirical Study: What Students Bring to Pathology Practicals

This appendix presents statistical data from the evaluation presented in chapter 8 - The First Empirical Study: What Students Bring to Pathology Practicals.

Table A2 - 1 Measure of association between students' expressed need for demonstrator assistance and age

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
agegroup	138	-0.13	0.12

Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. An indicator on a questionnaire administered before exposure to SCALPEL operationalized age (agegroup). The sample (N) was made up of students who expressed an opinion towards the need for demonstrator assistance and indicated their age. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

Table A2 - 2 Measures of association between students' expressed need for demonstrator assistance and student demographics measured at the nominal level

Variable	N	Goodman-Kruskal tau	2-tailed Significance (p)
gender	138	0.02	0.09
schoolda	138	0.08**	0.002

** Correlation is significant at the 0.01 level (2-tailed)

Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. Gender and demographic area of post-primary education (schoolda) were operationalized through indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up of students who expressed an opinion towards the need for demonstrator assistance and indicated their gender and post-primary education. For each measure of association the Goodman-Kruskal tau and 2-tailed significance (p) are expressed.

Table A2 - 3 Measures of association between students' expressed need for demonstrator assistance and social economic characteristics

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
hiedcode	138	-0.04	0.60
accscale	125	-0.08	0.40

Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. Students' previous higher education (hiedcode) and access to computers (accscale) were operationalized through indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up of students who expressed an opinion towards the need for demonstrator assistance and indicated their previous higher education and access to computers. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 4 Measures of association between students' expressed need for demonstrator assistance and previous computer use

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
usagefl	134	-0.02	0.78
usagegam	137	-0.15	0.07
usageint	134	-0.10	0.25
usagewp	137	-0.02	0.85
usagewrk	131	-0.03	0.74

Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. Indicators on a questionnaire administered before exposure to SCALPEL measured previous use of computers for formal learning (usagefl), games (usagegam), Internet (usageint), word processing (usagewp) and work (usagewrk). Each sample (N) was made up of students who expressed an opinion towards the need for demonstrator assistance and indicated their previous computer use for the relevant variable. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 5 Measure of association between students' expressed need for demonstrator assistance and learning styles

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
stylesmi	124	-0.00	0.97

Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. Student learning styles (stylesmi) were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. The sample (N) was made up of students who expressed an opinion towards the need for demonstrator assistance and a behavioural learning style. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

Table A2 - 6 Measures of association between students' expressed need for demonstrator assistance and attitudes towards computers

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
preattmi	124	0.05	0.54

Students' need for demonstrator assistance was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered at the end of the study. Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. Each sample (N) was made up of students who expressed an opinion towards computers and the need for demonstrator assistance. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 7 Measure of association between students' pre-SCALPEL attitudes towards computers and age

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
agegroup	142	-0.02	0.79

A student's attitude towards computers was operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. An indicator on the same questionnaire operationalized age (agegroup). The sample (N) was made up of students who expressed an attitude towards computers and indicated their age. Spearman's rho correlation coefficient (ρ) and 2-tailed significance are expressed for the measure of association.

Table A2 - 8 Measures of association between students' pre-SCALPEL attitudes towards computers and student demographics measured at the nominal level

Variable	N	Goodman-Kruskal tau	2-tailed Significance (p)
gender	142	0.02*	0.04
schoolda	142	0.02	0.80

* Correlation is significant at the 0.05 level (2-tailed)

Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. Gender and demographic area of post-primary education (schoolda) were operationalized through indicators on the same questionnaire. Each sample (N) was made up of students who expressed an attitude towards computers and indicated their gender and post-primary education. For each measure of association the Goodman-Kruskal tau and 2-tailed significance (p) are expressed.

Table A2 - 9 Measures of association between students' pre-SCALPEL attitudes towards computers and social economic characteristics

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
hiedcode	142	-0.07	0.40
accscale	130	0.32**	0.000

** Correlation is significant at the 0.01 level (2-tailed)

Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. Students' previous higher education (hiedcode) and access to computers (accscale) were operationalized through indicators on the same questionnaire. Each sample (N) was made up of students who expressed an attitude towards computers and indicated their previous higher education and access to computers. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 10 Measures of association between students' pre-SCALPEL attitudes towards computers and previous computer use

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
usagefl	138	0.21*	0.01
usagegam	140	0.24**	0.005
usageint	139	0.12	0.15
usagewp	141	0.44**	0.000
usagewrk	134	0.16	0.07

* Correlation is significant at the 0.05 level (2-tailed)
** Correlation is significant at the 0.01 level (2-tailed)

Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. Indicators on a questionnaire administered before exposure to SCALPEL measured previous use of computers for formal learning (usagefl), games (usagegam), Internet (usageint), word processing (usagewp) and work (usagewrk). Each sample (N) was made up of students who expressed an attitude towards computers and indicated their previous computer use for the relevant variable. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 11 Measure of association between students' pre-SCALPEL attitudes towards computers and learning styles

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
stylesmi	141	0.17*	0.04

* Correlation is significant at the 0.05 level (2-tailed)

Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered before exposure to SCALPEL. Student learning styles (stylesmi) were operationalized by a multi-item measure consisting of 14 indicators on the same questionnaire. The sample (N) was made up of students who expressed an attitude towards computers and a behavioural learning style. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

Table A2 - 12 Wilcoxon matched pairs test comparing students' attitudes towards computers pre and post exposure to SCALPEL

	N	Mean Rank	Sum of Ranks
Negative Ranks	13[a]	19.00	247.00
Positive Ranks	24[b]	19.00	456.00
Ties	94[c]		
Total	131		
Z	-1.81		
2-tailed significance (p)	0.07		

a. Post-SCALPEL attitudes towards computers < Pre-SCALPEL attitudes towards computers
b. Post-SCALPEL attitudes towards computers > Pre-SCALPEL attitudes towards computers
c. Pre-SCALPEL attitudes towards computers = Post-SCALPEL attitudes towards computers

Students' attitudes towards computers were operationalized by a multi-item measure consisting of 6 indicators on a questionnaire administered pre and post exposure to SCALPEL. The sample of matched pairs consisted of those students who expressed an attitude towards computers on both questionnaires. The number of cases (N), mean rank and sum of ranks are expressed for the negative and positive ranks, together with the number of tied ranks, total number of cases and test statistic (*Z*) with its 2-tailed significance (p).

Table A2 - 13 Measure of association between students' pre-SCALPEL learning styles and age

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
agegroup	142	-0.02	0.77

Student learning styles were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. An indicator on the same questionnaire operationalized age (agegroup). The sample (N) was made up of students who expressed a behavioural learning style and indicated their age. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for the measure of association.

Table A2 - 14 Measures of association between students' pre-SCALPEL learning styles and student demographics measured at the nominal level

Variable	N	Goodman-Kruskal tau	2-tailed Significance (p)
gender	142	0.03	0.06
schoolda	142	0.01	0.80

Student learning styles were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. Gender and demographic area of post-primary education (schoolda) were operationalized through indicators on the same questionnaire. Each sample (N) was made up of students who expressed a behavioural learning style and indicated their gender and post-primary education. For each measure of association the Goodman-Kruskal tau and 2-tailed significance (p) are expressed.

Table A2 - 15 Measures of association between students' pre-SCALPEL learning styles and social economic characteristics

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
hiedcode	142	-0.02	0.79
accscale	130	0.09	0.30

Student learning styles were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. Students' previous higher education (hiedcode) and access to computers (accscale) were operationalized through indicators on the same questionnaire. Each sample (N) was made up of students who expressed a behavioural learning style and indicated their previous higher education and access to computers. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 16 Measures of association between students' pre-SCALPEL learning styles and previous computer use

Variable	N	rho coefficient (ρ)	2-tailed Significance (p)
usagefl	138	0.12	0.15
usagegam	140	0.14	0.10
usageint	139	-0.17*	0.04
usagewp	141	0.12	0.17
usagewrk	134	0.03	0.74

* Correlation is significant at the 0.05 level (2-tailed)

Student learning styles were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered before exposure to SCALPEL. Indicators on a questionnaire administered before exposure to SCALPEL measured previous use of computers for formal learning (usagefl), games (usagegam), Internet (usageint), word processing (usagewp) and work (usagewrk). Each sample (N) was made up of students who expressed a behavioural learning style and indicated their previous computer use for the relevant variable. Spearman's rho correlation coefficient (ρ) and 2-tailed significance (p) are expressed for each measure of association.

Table A2 - 17 Wilcoxon matched pairs test comparing students' learning styles pre and post exposure to SCALPEL

	N	Mean Rank	Sum of Ranks
Negative Ranks	24[a]	17.00	408.00
Positive Ranks	9[b]	17.00	153.00
Ties	99[c]		
Total	132		
Z	-261		
2-tailed significance (p)	0.01		

a. Post-SCALPEL learning styles < Pre-SCALPEL learning styles (passive to explorative)
b. Post-SCALPEL learning styles > Pre-SCALPEL learning styles (explorative to passive)
c. Pre-SCALPEL learning styles = Post-SCALPEL learning styles

Student learning styles were operationalized by a multi-item measure consisting of 14 indicators on a questionnaire administered pre and post exposure to SCALPEL. The sample of matched pairs consisted of those students who expressed a behavioural learning style on both questionnaires. The number of cases (N), mean rank and sum of ranks are expressed for the negative and positive ranks, together with the number of tied ranks, total number of cases and test statistic (*Z*) with its 2-tailed significance (p).

Appendix 3: Quantitative Statistics of the Second Empirical Study

This appendix presents statistical data from the evaluation presented in chapter 9 - The Second Empirical Study.

Table A3 - 1 Distribution of students by age group

Value	Frequency	Percent	Valid Percent	Cumulative Percent
under 20	68	70.1	70.1	70.1
20 - 25	25	25.8	25.8	95.9
26 - 35	4	4.1	4.1	100.0
over 35	0	0	0	100.0
Sample Total	97	100.0	100.0	

Age group was operationalized by asking students to indicate their age group on a questionnaire administered during the last lecture of the pathology course. The sample frame (sample total) constituted all first year medical students who indicated their age on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A3 - 2 Distribution of students' acceptance of SCALPEL

Value	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Agree	6	6.2	6.3	6.3
Agree	39	40.2	41.1	47.4
Neutral	42	43.3	44.2	91.6
Disagree	8	8.2	8.4	100.0
Strongly Disagree	0	0	0	100.0
Valid Total	95	97.9	100.0	
Missing	2	2.1		
Sample Total	97	100.0		

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered during the last lecture of the pathology course. The sample frame (sample total) constituted all first year medical students who expressed an attitude on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A3 - 3 Distribution of students' expressed need for demonstrators during SCALPEL sessions

Value	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Agree	7	7.2	7.5	7.5
Agree	34	35.1	36.6	44.1
Neutral	34	35.1	36.6	80.6
Disagree	16	16.5	17.2	97.8
Strongly Disagree	2	2.1	2.2	100.0
Valid Total	93	95.9	100.0	
Missing	4	4.1		
Sample Total	97	100.0		

Students' need for demonstrator assistance was operationalized by a multiple-item measure consisting of 4 indicators on a questionnaire administered during the last lecture of the pathology course. The sample frame (sample total) constituted first year medical students who expressed an attitude on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A3 - 4 Distribution of reasons for students expressing a need for demonstrator assistance during a SCALPEL session

SPSS Key	Reason for requiring demonstrator assistance	Frequency of consults	Percentage of total consults
RD1	Discuss pathology material being taught other than Histology	132	23.7
RD2	Explain how to use the multiple choice window	6	1.1
RD3	Discuss Histology	168	30.2
RD4	Discuss background information about any aspect of the material being taught	103	18.5
RD5	Discuss how the material being taught relates to other areas of medicine	86	15.4
RD6	Clarify aims and objectives of the practical session	37	6.6
RD7	Explain how to start the application	7	1.3
RD8	Explain how to use the application features (i.e. video, image zoom)	18	3.2
	Total	557	100.0

Students were asked to indicate how many times on average in each practical they would like to have consulted a demonstrator for each of the reasons. The sample (N = 66) was made up of students who indicated on a returned questionnaire the frequency of reasons for requiring demonstrators. The data given represents an average SCALPEL session. For each reason the average total number of demonstrator consults and the percentage of the total number of consults is given.

Table A3 - 5 Importance placed by students on demonstrators to be available to deal with specific problems. Expressed as the rounded mean score of the sum of all ranks in ascending order of importance (most important first)

SPSS Key	Reason for requiring demonstrator assistance	N	Rank Importance Mean	Mode
RD3	Discuss Histology	66	2.23	1
RD1	Discuss pathology material being taught other than Histology	66	3.03	1
RD4	Discuss background information about any aspect of the material being taught	63	3.30	3
RD5	Discuss how the material being taught relates to other areas of medicine	55	4.04	3
RD6	Clarify aims and objectives of the practical session	44	5.70	5
RD8	Explain how to use the application features (i.e. video, image zoom)	32	5.91	6
RD2	Explain how to use the multiple choice window	32	7.03	10
RD7	Explain how to start the application	32	7.09	10

Students were asked to indicate using a scale of 1 (most important) to 10 (least important) the importance they placed on demonstrators to be available for each of the categories. Each sample (N) consisted of students who indicated on a returned questionnaire a rank for the relevant category. The data given represents the mode rank and the rounded mean score of the sum of all ranks for that category.

Table A3 - 6 Number of students who used identified methods to solve problems in the absence of demonstrators

Methods use to solve problems		Reasons for requiring demonstrator assistance (see Table A3 - 4 for description of categories RD1 - RD8)															
		RD1		RD2		RD3		RD4		RD5		RD6		RD7		RD8	
		N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
a	text books	33	(50.0)	6	(9.1)	51	(77.3)	37	(56.1)	22	(33.3)	5	(7.6)	1	(1.5)	1	(1.5)
b	help provided by the pathology application	11	(16.7)	11	(16.7)	13	(19.7)	3	(4.5)	2	(3.0)	9	(13.6)	16	(24.2)	14	(21.2)
c	asked in tutorial	35	(53.0)	4	(6.1)	39	(59.1)	50	(75.8)	44	(66.7)	19	(28.8)	1	(1.5)	3	(4.5)
d	asked a member of staff outside of tutorial	2	(3.0)	1	(1.5)	2	(3.0)	6	(9.1)	5	(7.6)	5	(7.6)	1	(1.5)	2	(3.0)
e	asked another student	21	(31.8)	21	(31.8)	34	(51.5	23	(34.8)	15	(22.7)	6	(9.1)	23	(34.8)	29	(43.9)
f	never found the answer	2	(3.0)	2	(3.0)	5	(7.6)	4	(6.1)	7	(10.6)	10	(15.2)	1	(1.5)	7	(10.6)
g	read lecture notes / handouts	9	(13.6)	6	(9.1)	9	(13.6)	9	(13.6)	10	(15.2)	13	(19.7)	10	(15.2)	2	(3.0)
h	found the answer in the application	6	(9.1)	10	(15.2)	12	(18.2)	3	(4.5)	3	(4.5)	4	(6.1)	6	(9.1)	3	(4.5)
	Total of all methods	119		61		165		135		108		71		59		61	

Students were asked to indicate the methods used to solve problems encountered during SCALPEL sessions. For each reason for requiring demonstrator assistance (RD_n), the number of students (N) who indicated using the relevant method is expressed together with the percentage (%) of the sample population (N = 66). Note that using a method was not mutually exclusive for any given reason for requiring demonstrator assistance, i.e. a student may have indicated using more than one method for any given reason.

Table A3 - 7 Importance placed by students on implementing each of the proposed SCALPEL improvements. Expressed as the mean score of the sum of all ranks in ascending order of importance (most important first)

SPSS Key	Proposed Application Improvements	N	Rank Importance Mean	Rank Importance Mode
IMP3	Ability to ask questions and receive answers about pathology material	74	2.61	1
IMP1	Glossary of terms	73	2.99	2
IMP2	On-line text books	73	3.36	1
IMP5	More links to background information about pathology material	74	3.45	3
IMP7	More animation like the ECG	72	5.18	7
IMP6	Optional practicals	70	5.47	6
IMP8	Introduction of sound to listen to patient history etc	72	5.69	6
IMP4	On-line notepad	68	6.35	8

Students were asked to indicate using a scale of 1 (most important) to 10 (least important) the importance they placed on each of the proposed SCALPEL improvements. Each sample (N) consisted of students who indicated on a returned questionnaire a rank for the relevant application improvement. The data given represents the mode rank and the rounded mean score of the sum of all ranks for that proposed improvement.

Appendix 4: Quantitative Statistics of the Third Empirical Study

This appendix presents statistical data from the evaluation presented in chapter 11 - The Third Empirical Study.

Table A4 - 1 Distribution of students by age group

Value	Frequency	Percent	Valid Percent	Cumulative Percent
under 20	59	79.7	79.7	79.7
20 - 25	12	16.2	16.2	95.9
26 - 35	3	4.1	4.1	100.0
over 35	0	0	0	100.0
Sample Total	74	100.0	100.0	

Age group was operationalized by asking students to indicate their age group on a questionnaire administered after working through practical 6. The sample frame (sample total) constituted all first year medical students who indicated their age on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A4 - 2 Distribution of students' acceptance of SCALPEL

Value	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Agree	11	14.9	14.9	14.9
Agree	42	56.8	56.8	71.6
Neutral	21	28.4	28.4	100.0
Disagree	0	0	0	100.0
Strongly Disagree	0	0	0	100.0
Valid Total	74	100.0	100.0	
Missing	0	0		
Sample Total	74	100.0		

Acceptance of SCALPEL was operationalized by a multi-item measure consisting of 4 indicators on a questionnaire administered after exposure to practical 6. The sample frame (sample total) constituted all first year medical students who expressed an attitude on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A4 - 3 Distribution of students' expressed need for demonstrators during SCALPEL sessions

Value	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Agree	0	0	0	0
Agree	6	8.1	8.5	8.5
Neutral	31	41.9	43.7	52.1
Disagree	26	35.1	36.6	88.7
Strongly Disagree	8	10.8	11.3	100.0
Valid Total	71	95.9	100.0	
Missing	3	4.1		
Sample Total	74	100.0		

Students' need for demonstrator assistance was operationalized by a multiple-item measure consisting of 4 indicators on a questionnaire administered after exposure to practical 6. The sample frame (sample total) constituted first year medical students who expressed an attitude on the returned questionnaire. The frequency, percentage of sample population (Percent), percentage of valid responses (Valid Percent) and cumulative percentage are expressed.

Table A4 - 4 Distribution of reasons for students expressing a need for demonstrator assistance during practical 6

SPSS Key	Reason for requiring demonstrator assistance	Control Group (N = 41)		Experiment Group (N = 27)		Entire Sample (N = 68)	
		No of Consults	Group (%)	No of Consults	Group (%)	No of Consults	Sample (%)
RD1	Discuss pathology material being taught other than Histology	74	(26.3)	30	(22.2)	104	(25.0)
RD2	Explain how to use the multiple choice window	0	(0)	5	(3.7)	5	(1.2)
RD3	Discuss Histology	87	(31.0)	38	(28.1)	125	(30.0)
RD4	Discuss background information about any aspect of the material being taught	61	(21.7)	30	(22.2)	91	(21.9)
RD5	Discuss how the material being taught relates to other areas of medicine	33	(11.7)	17	(12.6)	50	(12.0)
RD6	Clarify aims and objectives of the practical session	10	(3.6)	5	(3.7)	15	(3.6)
RD7	Explain how to start the application	1	(0.4)	7	(5.2)	8	(1.9)
RD8	Explain how to use the application features (i.e. video, image zoom etc.)	15	(5.3)	3	(2.2)	18	(4.3)
	Total	281	(100.0)	135	(100.0)	416	(100.0)

Students were asked to indicate approximately how many times during practical 6 they would like to have consulted a demonstrator for each of the reasons. The samples were made up of students who indicated on a returned questionnaire the frequency of reasons for requiring demonstrators. For each reason the data given represents the total number of demonstrator consults and the percentage of the total number of consults for the control and experiment groups, and for the entire sample.

Table A4 - 5 Importance placed by students on demonstrators to be available to deal with specific problems. Expressed as the rounded mean score of the sum of all ranks in ascending order of importance (most important first)

SPSS Key	Reason for requiring demonstrator assistance	Control Group (N = 41) Rank Importance				Experiment Group (N = 27) Rank Importance				Entire Sample (N = 68) Rank Importance			
		n	R	Mean	Mode	n	R	Mean	Mode	n	R	Mean	Mode
RD1	Discuss pathology material being taught other than Histology	32	2	3.16	1	19	2	2.84	1	51	2	3.04	1
RD2	Explain how to use the multiple choice window	14	8	8.64	8	11	8	6.27	10	25	8	7.60	8
RD3	Discuss Histology	32	1	2.59	1	19	4	3.32	1	51	1	2.86	1
RD4	Discuss background information about any aspect of the material being taught	33	3	3.70	3	19	3	2.84	1	52	3	3.38	3
RD5	Discuss how the material being taught relates to other areas of medicine	24	4	4.38	3	12	5	3.50	2	36	4	4.08	3
RD6	Clarify aims and objectives of the practical session	18	6	5.83	5	9	1	2.78	2	27	5	4.81	5
RD7	Explain how to start the application	14	7	8.50	7	10	6	5.10	7	24	7	7.08	7
RD8	Explain how to use the application features (i.e. video, image zoom)	19	5	5.63	6	9	7	6.11	6	28	6	5.79	6

Students were asked to indicate using a scale of 1 (most important) to 10 (least important) the importance they placed on demonstrators to be available for each of the categories. Each sample (n) consisted of students who indicated on a returned questionnaire a rank for the relevant category. The data given represents the mode rank and the rounded mean score of the sum of all ranks for that category.

Table A4 - 6 Number of control group students who used identified methods to solve problems in the absence of demonstrators

	Methods use to solve problems	**Reasons for requiring demonstrator assistance (see Table A3 - 4 for description of categories RD1 - RD8)**															
		RD1		**RD2**		**RD3**		**RD4**		**RD5**		**RD6**		**RD7**		**RD8**	
		N	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**	**N**	**(%)**
a	text books	15	(36.6)	3	(7.3)	14	(34.1)	13	(31.7)	8	(19.5)	0	(0.0)	0	(0.0)	0	(0.0)
b	help provided by the pathology application	9	(22.0)	4	(9.8)	8	(19.5)	3	(7.3)	2	(4.9)	3	(7.3)	1	(2.4)	2	(4.9)
c	asked in tutorial	14	(34.1)	4	(9.8)	17	(41.5)	27	(65.9)	22	(53.7)	10	(24.4)	0	(0.0)	0	(0.0)
d	asked a member of staff outside of tutorial	0	(0.0)	0	(0.0)	0	(0.0)	1	(2.4)	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
e	asked another student	11	(26.8)	6	(14.6)	22	(53.7)	12	(29.3)	4	(9.8)	2	(4.9)	8	(19.5)	11	(26.8)
f	never found the answer	0	(0.0)	0	(0.0)	4	(9.8)	1	(2.4)	1	(2.4)	3	(7.3)	1	(2.4)	2	(4.9)
g	read lecture notes / handouts	6	(14.6)	5	(12.2)	10	(24.4)	10	(24.4)	6	(14.6)	6	(14.6)	6	(14.6)	1	(2.4)
h	found the answer in the application	8	(19.5)	8	(19.5)	8	(19.5)	1	(2.4)	1	(2.4)	0	(0.0)	5	(12.2)	8	(19.5)
	Total of all methods	63		30		83		68		44		24		21		24	

Students were asked to indicate the methods used to solve problems encountered during practical 6. For each reason for requiring demonstrator assistance (RD-$_n$), the number of students (N) who indicated using the relevant method is expressed together with the percentage (%) of the control group (N = 41). Note that using a method was not mutually exclusive for any given reason for requiring demonstrator assistance, i.e. a student may have indicated using more than one method for any given reason.

Table A4 - 7 Number of experiment group students who used identified methods to solve problems in the absence of demonstrators

	Methods use to solve problems	RD1 N	RD1 (%)	RD2 N	RD2 (%)	RD3 N	RD3 (%)	RD4 N	RD4 (%)	RD5 N	RD5 (%)	RD6 N	RD6 (%)	RD7 N	RD7 (%)	RD8 N	RD8 (%)
		Reasons for requiring demonstrator assistance (see Table A3 - 4 for description of categories RD1 - RD8)															
a	text books	8	(29.6)	0	(0.0)	12	(44.4)	8	(29.6)	8	(29.6)	0	(0.0)	0	(0.0)	0	(0.0)
b	help provided by the pathology application	11	(40.7)	1	(3.7)	15	(55.6)	7	(25.9)	3	(11.1)	5	(18.5)	0	(0.0)	1	(3.7)
c	asked in tutorial	12	(44.4)	0	(0.0)	10	(37.0)	15	(55.6)	16	(59.3)	5	(18.5)	1	(3.7)	1	(3.7)
d	asked a member of staff outside of tutorial	1	(3.7)	1	(3.7)	3	(11.1)	2	(7.4)	1	(3.7)	0	(0.0)	3	(11.1)	0	(0.0)
e	asked another student	9	(33.3)	6	(22.2)	10	(37.0)	8	(29.6)	4	(14.8)	1	(3.7)	4	(14.8)	5	(18.5)
f	never found the answer	0	(0.0)	0	(0.0)	1	(3.7)	0	(0.0)	0	(0.0)	1	(3.7)	0	(0.0)	0	(0.0)
g	read lecture notes / handouts	5	(18.5)	1	(3.7)	5	(18.5)	7	(25.9)	2	(7.4)	6	(22.2)	2	(7.4)	0	(0.0)
h	found the answer in the application	7	(25.9)	3	(11.1)	7	(25.9)	2	(7.4)	1	(3.7)	1	(3.7)	3	(11.1)	1	(3.7)
	Total of all methods	53		12		63		79		35		19		13		8	

Students were asked to indicate the methods used to solve problems encountered during practical 6. For each reason for requiring demonstrator assistance ($RD-_n$), the number of students (N) who indicated using the relevant method is expressed together with the percentage (%) of the experiment group (N = 27). Note that using a method was not mutually exclusive for any given reason for requiring demonstrator assistance, i.e. a student may have indicated using more than one method for any given reason.

Table A4 - 8 Importance placed by students on implementing each of the proposed SCALPEL improvements. Expressed as the mean score of the sum of all ranks in ascending order of importance (most important first)

SPSS Key	Proposed Application Improvements	Control Group Rank Importance				Experiment Group Rank Importance				Entire Sample Rank Importance			
		n	R	Mean	Mode	n	R	Mean	Mode	n	R	Mean	Mode
IMP1	Glossary of terms	44	2	2.93	2	28	1	2.96	2	72	1	2.94	2
IMP2	On-line text books	44	1	2.64	2	27	4	4.04	1	71	2	3.17	2
IMP3	Ability to ask questions and receive answers about pathology material	44	3	3.14	1	28	2	3.32	1	72	3	3.21	1
IMP4	On-line notepad	42	8	7.29	8	25	8	5.96	8	67	8	6.79	8
IMP5	More links to background information about pathology material	44	4	3.32	3	28	3	3.43	3	72	4	3.36	3
IMP6	Optional practicals	43	7	5.67	5	26	7	5.31	5	69	7	5.54	5
IMP7	More animation like the ECG	43	5	4.77	5	28	5	4.86	3	71	5	4.80	5
IMP8	Introduction of sound to listen to patient history etc	43	6	5.44	5	26	6	5.00	2	69	6	5.28	5

Students were asked to indicate using a scale of 1 (most important) to 10 (least important) the importance they placed on each of the proposed SCALPEL improvements. Each sample (N) consisted of students who indicated on a returned questionnaire a rank for the relevant application improvement. The data given represents the mode rank and the rounded mean score of the sum of all ranks for that proposed improvement.

Appendix 5: Example of an Original Pathology Practical Case

FIRST YEAR MEDICAL CURRICULUM

TERM 2 - CARDIOPULMONARY

Pathology Practical 1 - 18 January 1996

Classroom Level E

Group A - 11.00 am to 12.30 pm

Group B - 2.00 pm to 3.30 pm

CASE HISTORY

Mr Coady, a 70 year old man, was admitted to hospital with a history of worsening breathlessness. He gave a history of similar episodes which had resolved on treatment from his general practitioner. Overall, his condition was gradually deteriorating. Mr Coady had smoked and inhaled 30 cigarettes per day for the previous 45 years.

On examination, Mr Coady was distressed and had swollen ankles. He was noted to be using the accessory muscles of respiration.

Question

What are the accessory muscles of respiration?

On listening to his chest, wheezes were heard and on examination of his abdomen, his liver was enlarged.

An arterial blood sample was taken and the following results were obtained.

pH	7.24	Normal range 7.34 - 7.44
p02	9.5 kPa	Normal range 12 - 15 kPa
pC02	8.5 kPa	Normal range 4.5 - 6.1 kPa

What is the significance of the above figures?

A venous blood sample was taken and centrifuged in order to separate the red cells from the plasma. This allows for the estimation of the packed cell volume. You are supplied with photographs of a normal sample (N) and samples from three patients. Using the ruler at the bottom of the page, measure the length of the whole sample, not including the white plug in the bottom of the tube, and the length of the column of red cells. Calculate the proportion which is occupied by the packed red cells.

Which sample do you think came from Mr Coady? What do you think was wrong with the two other patients?

Mr Coady's haemoglobin was measured at 18.0g/dl, the upper limit of normal being 17g/dl. The normal packed cell volume is 0.4 ±0.07. How does the haemoglobin measurement relate to your findings? What is the mechanism behind this abnormality?

Despite medical and nursing care, Mr Coady deteriorated further and died. Permission for autopsy was obtained and an autopsy performed. You are provided with photographs of the heart and of the liver and histology of the liver.

What abnormality do you see in the heart? (NB the normal ratio of thickness of the left to right ventricle is 5:1).

What was the cause of this abnormality?

Describe the appearances of the cut surface of the liver.

You are provided with two histology slides (Nos 97 and 98). Which of these slides do you think came from Mr Coady's liver? What are the causes and mechanisms of these abnormalities?

Summarise the findings of Mr Coady's last illness and describe the relationship between the abnormalities which you have found.

Appendix 6: Example of a SToMP MCQ Configuration Script

[header]
title=Term 2 - Pathology Practical 1
environment=stomp
mode=linked
test=formative
questions=8
rubric=Answer all the questions you can and then press the 'mark' button.
AlwaysOnTop=True

[Question 1]
title=1
style=1
choice=5
KeyWords=
Rubric=Select one button
right=c
IfRight=Your answer to question one is correct. Sternocleidomastoid, pectoralis major and minor, and latissimus muscles are the accessory muscles of respiration.
Rightkey=
Others=0
IfWrong=Your answer to question one is wrong, try again.
Wrongkey=

[Question 2]
title=2
style=1
choice=5
KeyWords=
Rubric=Select one button
right=b
IfRight=Your answer to question two is correct. The data inidicates hypoxia and respiratory acidosis because the oxygen saturation is low, the pH is also low, i.e. [H+] increased, and the partial pressure of CO2 is increased, due to respiratory failure.
Rightkey=
Others=0
IfWrong=Your answer to question two is wrong, try again.
Wrongkey=

Appendix 7: Example of the MCQ/Demonstrator Agent Configuration Script

```
[header]
title=Term 2 - Pathology Practical 6
environment=stomp
mode=linked
test=formative
questions=20
rubric=Answer the question and press the 'Mark Question' button.
AlwaysOnTop=True
ContextID=1.2.5

[Question 1]
title=Question 1
style=1
choice=5
KeyWords=
Rubric=Select one button
right=c
IfRight=Your answer to question one is correct. Sternocleidomastoid, pectoralis major and minor, and latissimus muscles are the accessory muscles of respiration.
Rightkey=
Others=0
IfWrong=Your answer to question one is wrong, try again.
Wrongkey=
ContextID=1.2.5.6

[Question 2]
title=Question 2
style=1
choice=3
KeyWords=
Rubric=Select one button
right=c
IfRight=Your answer to question two is correct. Eosinophil leukocytosis occurs in a variety of disorders but allergic disease is one of the commonest causes in the UK.
Rightkey=
Others=1
OtherCondition1=a
OtherMessage1=Your answer to question two is wrong. Another type of granulocyte is increased in this blood film.
IfWrong=Your answer to question two is wrong. Eosinophil leukocytes are not involved in the response to acute bacterial infection.
Wrongkey=
ContextID=1.2.5.7
```

Appendix 8: Concept Hierarchy - Relational Database Schema

A8.1 Concept_Contexts table with example data

Concept_Contexts

ConceptID	ContextID	Description
1	1	Pathology
2	1.2	Term 2
3	1.3	Term 3
4	1.2.4	Practical 1
5	1.2.5	Practical 6
6	1.2.4.6	Q1
6	1.2.5.6	Q1
7	1.2.5.7	Q2

A8.2 Questions table with example data

Questions

QuestionNumber	Originator	DateRaised	Question	DomainExpert	DateAnswered	Answer	ContextID
1034	mwk	05/04/1998	...	William Roche	06/04/1998	...	1.2.5.6
1035	std1	05/04/1998	...	William Roche	06/04/1998	...	1.2.4.6
1097	mwk	07/04/1998	...	Wendy Hall	07/04/1998	...	1.2.5.7

Appendix 9: Questionnaires of the First Empirical Study

This appendix presents the questionnaires used during the first empirical study. Please note that for each questionnaire the left-hand margin has been reduced to accommodate the layout requirements of this thesis. The effect of this is most notable with the indicators based on Likert scaling where the text appears clustered relative to the actual questionnaires. In all other respects the questionnaires are presented here as they were to the respondents.

The first questionnaire (see section A9.1) was administered to the students before the start of the first practical (in the traditional laboratory for group A and in the computer suite for group B). Students were given the second questionnaire (see section A9.2) to complete at the end of their last practical using SCALPEL (after practical 3 for group B and practical 6 for group A). The final questionnaire (see section A9.3) was given to students to complete before starting the MCQ and short answer examinations administered at the end of the first study. To promote anonymity, each student was assigned a number at the beginning of the study. Each questionnaire was stamped with a valid student number and a piece of paper with the corresponding student name was stapled to the front page. The questionnaires were then administered to the correct student at the appropriate time. Before returning a completed questionnaire students were instructed to remove the piece of paper with their name on it. The student numbers stamped on each of the questionnaires were then used to ensure that students were matched to the correct three questionnaires during the data processing stage.

A9.1 The First Empirical Study - Questionnaire 1

Term 2
Cardio-Respiratory Systems Course
Pathology

This term, we will be introducing a computer based pathology teaching programme that we would like you to use. Half the year will use the computer programme for half the term and then the two halves will swap over. Everyone will have the same teaching overall by the end of the term.

In order to evaluate the usefulness of this teaching format, we need to know what you think about it. We would be very grateful if you could take some time to answer this questionnaire. Your answers will be treated confidentially. Feel free to say how you personally feel about the appropriateness and acceptability of the computer programme.

Thank you for your co-operation.

....

1. Student number ....

(this has been allocated to you for the study in order to protect your confidentiality)

2 Age *please tick one box*

under 20 ☐
20-25 ☐
26-35 ☐
over 35 ☐

3 Sex *please tick one box*

male ☐
female ☐

4. Country in which you went to school after the age of 11

5. Previous degree (if any)

5.1 Country in which you took this degree

6. Previous computer usage (before starting medical school)

please tick as many boxes as appropriate

	frequent (more than 1x per week)	**sometimes** (less than 1x per week)	**never before**
games	☐	☐	☐
formal learning programme	☐	☐	☐
Internet	☐	☐	☐
word processing	☐	☐	☐
at work	☐	☐	☐

7. Access to computers

please tick boxes as appropriate

	yes	**no**
I own a portable (laptop) computer	☐	☐
I have a computer in my room at University	☐	☐
I have a desk top computer at home	☐	☐
I have access to a computer at home	☐	☐
I do not have access to a computer other than the University workstations	☐	☐

8. The following set of statements relate to your attitudes to computers and other learning methods

*please circle **one** number only for each statement*

	strongly agree	agree	neutral	disagree	strongly disagree
Computers often make more problems than they solve	1	2	3	4	5
Computers can be useful, but I have never found them so	1	2	3	4	5
I sometimes find computers useful	1	2	3	4	5
I find computers useful and efficient tools	1	2	3	4	5
I could not function without computers	1	2	3	4	5
I prefer to use handouts instead of textbooks	1	2	3	4	5
We should be told what sections of textbooks to read	1	2	3	4	5
Textbooks are helpful aids to fill in areas of learning	1	2	3	4	5
We should be left to find the information about subjects from textbooks	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
We should be allowed to explore subjects according to our interests	1	2	3	4	5
Lectures should tell us all we need to know	1	2	3	4	5
Lectures are unnecessary	1	2	3	4	5
Lectures should cover all I need to know about a subject	1	2	3	4	5
Lectures should provide me with a guide to additional information that I should know about a subject area	1	2	3	4	5
Lectures can be a helpful way of finding out about a subject area	1	2	3	4	5
Practicals should be demonstrations of what we need to know	1	2	3	4	5
Practicals should require that we recognise things that we learnt in the course	1	2	3	4	5
Practicals should include some problem solving work	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
Practicals should be about solving problems relating to areas that we have not covered	1	2	3	4	5
Computers make me uneasy	1	2	3	4	5

9. If you have any comments to make about use of computers for learning, please write them below

Thank you for your co-operation

A9.2 The First Empirical Study - Questionnaire 2

Term 2
Cardio-Respiratory Systems Course
Pathology

Thank you for filling out the previous questionnaire. Now that you have used the Pathology Computer Application we would like to know if your views on computer aided learning within the Pathology course have changed, therefore please could you fill out the following questionnaire.

Again your answers will be treated confidentially. Feel free to say how you personally feel about the appropriateness and acceptability of the computer programme.

Thank you for your co-operation.

....

1. Student number ....

(this has been allocated to you for the study in order to protect your confidentiality)

Attitudes to Computer Aided Learning

2. The following set of statements relate to your attitudes to computers and other learning methods

*please circle **one** number only for each statement*

	strongly agree	agree	neutral	disagree	strongly disagree
Computers often make more problems than they solve	1	2	3	4	5
Computers can be useful, but I have never found them so	1	2	3	4	5
I sometimes find computers useful	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
I find computers useful and efficient tools	1	2	3	4	5
I could not function without computers	1	2	3	4	5
I prefer to use handouts instead of textbooks	1	2	3	4	5
We should be told what sections of textbooks to read	1	2	3	4	5
Textbooks are helpful aids to fill in areas of learning	1	2	3	4	5
We should be left to find the information about subjects from textbooks	1	2	3	4	5
We should be allowed to explore subjects according to our interests	1	2	3	4	5
Lectures should tell us all we need to know	1	2	3	4	5
Lectures are unnecessary	1	2	3	4	5
Lectures should cover all I need to know about a subject	1	2	3	4	5
Lectures should provide me with aguide to additional information that I should know about a subject area	1	2	3	4	5
Lectures can be a helpful way of finding out about a subject area	1	2	3	4	5
Practicals should be demonstrations of what we need to know	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
Practicals should require that we recognise things that we learnt in the course	1	2	3	4	5
Practicals should include some problem solving work	1	2	3	4	5
Practicals should be about solving problems relating to areas that we have not covered	1	2	3	4	5
Computers make me uneasy	1	2	3	4	5

3. If you have any comments to make about use of computers for learning, please write them below

Thank you for your co-operation

A9.3 **The First Empirical - Questionnaire 3**

Term 2
Cardio-Respiratory Systems Course
Pathology

Thank you for filling out the previous questionnaires. Now that you have completed the series of term 2 pathology practicals we would like to know your views on the two teaching methods used, computer aided learning and laboratory teaching. Therefore please could you fill out the following questionnaire.

Again your answers will be treated confidentially. Feel free to say how you personally feel about the two teaching methods.

Thank you for your co-operation.

....

1. Student number ....

(this has been allocated to you for the study in order to protect your confidentiality)

2. The following set of statements relate to your attitudes to computers and other learning methods

*please circle **one** number only for each statement*

	strongly agree	agree	neutral	disagree	strongly disagree
I find the computer application a satisfying learning activity	1	2	3	4	5
I find more teaching material is provided in the laboratory	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
I prefer to use the computer application for pathology practicals	1	2	3	4	5
I learn more pathology using the computer application	1	2	3	4	5
I do not need the help of demonstrators when I use the computer application	1	2	3	4	5
I do not feel that laboratory teaching is a satisfying learning activity	1	2	3	4	5
I participate more in the practical using the computer application	1	2	3	4	5
I understand the teaching material better when it is presented by the computer application	1	2	3	4	5
Not enough teaching material is provided in the laboratory	1	2	3	4	5
I felt lost and confused using the computer application	1	2	3	4	5
I believe demonstrators should be available at set times in the future	1	2	3	4	5
I like the consistency of explanations provided by the computer application	1	2	3	4	5
I prefer Laboratory teaching for pathology practicals	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
I like the environment the computer application provides	1	2	3	4	5
I miss demonstrator input whilst using the computer application	1	2	3	4	5
I find practicals in the laboratory a satisfying learning activity	1	2	3	4	5
I participate more in the laboratory	1	2	3	4	5
I understand the teaching material better when it is presented in the laboratory	1	2	3	4	5
I learn more pathology in the laboratory	1	2	3	4	5
I will always need the help of demonstrators, irrespective of the quality of knowledge given by the computer application	1	2	3	4	5
I find more teaching material is provided by the computer application	1	2	3	4	5

3. Write in any comments that you would like to make about the application and its use:

Thank you for your co-operation

Appendix 10: Questionnaire of the Second Empirical Study

This appendix presents the questionnaire used during the second empirical study. Please note that the left-hand margin has been reduced to accommodate the layout requirements of this thesis. The effect of this is most notable with the indicators based on Likert scaling where the text appears clustered relative to the actual questionnaires. In all other respects the questionnaire is presented here as it was to the respondents.

The questionnaire was given to the students at the start of the last lecture of the pathology course and collected after the lecture. Students who wished to take the questionnaire away for completion at a later date were given the opportunity to do so and were instructed to return them to the pathology office or their pathology tutor.

Term 2
Cardio-Respiratory Systems Course
Pathology

Computer aided learning was used to present term 2 pathology practicals. In order to evaluate the usefulness of this teaching format, we need to know what you think about it. We would be very grateful if you could take some time to answer this questionnaire. Your answers will be treated confidentially. Feel free to say how you personally feel about the appropriateness and acceptability of the computer programme.

Thank you for your co-operation.

....

Personal Characteristics

1. **Age** *please tick one box*

under 20 ☐
20-25 ☐
26-35 ☐
over 35 ☐

2. **Sex** *please tick one box*

male ☐
female ☐

Attitudes to Pathology Practicals

3. **The following set of statements relate to your attitudes to computers and other learning methods.**

please circle one number only for each statement

	strongly agree	agree	neutral	disagree	strongly disagree
Computers often make more problems than they solve	1	2	3	4	5
Computers can be useful, but I have never found them so	1	2	3	4	5
I sometimes find computers useful	1	2	3	4	5
I find computers useful and efficient tools	1	2	3	4	5
I could not function without computers	1	2	3	4	5
Computers make me uneasy	1	2	3	4	5
I find the computer application a satisfying learning activity	1	2	3	4	5
I do not need the help of demonstrators when I use the computer application	1	2	3	4	5
I felt lost and confused using the computer application	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
I believe demonstrators should be available at set times in the future	1	2	3	4	5
I like the consistency of explanations provided by the computer application	1	2	3	4	5
I like the environment the computer application provides	1	2	3	4	5
I miss demonstrator input whilst using the computer application	1	2	3	4	5
I will always need the help of demonstrators, irrespective of the quality of knowledge given by the computer application	1	2	3	4	5

4. If you have any comments to make about computers and other learning methods, please write them below.

The Role of Demonstrators

5. The following section relates to the role of demonstrators during practicals.

a. In the 'Times per Practical' column of the table below, please indicate how many times on average in each practical you would like to have consulted a demonstrator for each of the reasons. You may use the last rows to enter reasons that have not been mentioned.

b. In the 'Rank Importance' column of the table below, please rank the importance of your reasons for requiring demonstrator assistance during the pathology practicals. Rank 1 is most important, 10 is least important. Only rank those reasons that you experienced personally.

Reason for requiring demonstrator assistance	**Times per Practical**	**Rank Importance**
Discuss pathology material being taught other than Histology		
Explain how to use the multiple choice window		
Discuss Histology		
Discuss background information about any aspect of the material being taught		
Discuss how the material being taught relates to other areas of medicine		
Clarify aims and objectives of practical session		
Explain how to start the application		
Explain how to use the application features (i.e. video, image zoom)		

The role of demonstrators during practicals cont.

c. Please indicate the methods you used to solve the problems encountered during the pathology practicals by placing the letter for each solution in the corresponding 'Enter Solution' column of the table below.

You may use each solution more than once, and more than one solution may be used for every 'reason for requiring demonstrator assistance'.

Reason for requiring demonstrator assistance	Enter Solution	Solutions (Ways you solved problems with the practicals)
Histology		A: text books
Other Material		B: help provided by the pathology application
MCQ window		C: asked in tutorial
Discuss background information		D: asked a member of staff outside of tutorial
Discuss how the material relates to other areas		E: asked another student
Clarify aims and objectives		F: never found the answer
How to start the application		G: read lecture notes / handouts
How to use the video, image zoom etc.		H: found the answer in the application
Enter other reason:		I: Enter other solution:
Enter other reason:		J: Enter other solution:

Pathology Application Improvements

6. **We are continually looking for new ways to improve the pathology application. This section gives you an opportunity to express ways in which you believe the application could be improved.**

Using the table below please rank the importance of the proposed improvements in terms of what you would most like to see incorporated into the application. Rank 1 is most important, 10 is least important. You may use the last rows to enter improvements that have not been mentioned.

Proposed Application Improvement	**Rank Importance**
Glossary of terms	
On-line text books	
Ability to ask questions and receive answers about pathology material	
On-line notepad	
More links to background information about pathology material	
Optional practicals	
More animations like the ECG	
Introduction of sound to listen to patient history etc.	

7. **Write in any others comments that you would like to make about any aspect of the pathology practicals.**

Thank you for your co-operation

Appendix 11: Questionnaire of the Third Empirical Study

This appendix presents the questionnaire used during the third empirical study. Please note that the left-hand margin has been reduced to accommodate the layout requirements of this thesis. The effect of this is most notable with the indicators based on Likert scaling where the text appears clustered relative to the actual questionnaires. In all other respects the questionnaire is presented here as it was to the respondents.

The questionnaire was administered to the control group at the start of the last lecture of the pathology course and collected after the lecture. The same questionnaire was administered to the students in the experimental group after completing practical 6 in the controlled environment. Questionnaires from the experimental group were flagged ("AM/PM" on top right of questionnaire) so they could be coded accordingly when entered onto SPSS. Students who wished to take the questionnaire away for completion at a later date were given the opportunity to do so and were instructed to return them to the pathology office or their pathology tutor.

Term 2 AM/PM

Cardiopulmonary Systems Course
Pathology - Practical 6

Computer aided learning was used to present term 2 Pathology Practicals. In order to evaluate the usefulness of this teaching format, we need to know what you think about it. This questionnaire refers to **Pathology Practical 6**. We would be very grateful if you could take some time to answer the following questions. Your answers will be treated confidentially. Feel free to say how you personally feel about the appropriateness and acceptability of the computer programme.

Thank you for your co-operation.

....

Personal Characteristics

1. Age *please tick one box*

- under 20 ☐
- 20-25 ☐
- 26-35 ☐
- over 35 ☐

2. Sex *please tick one box*

- male ☐
- female ☐

Attitudes to Pathology Practical 6

3. The following set of statements relate to your attitudes to Pathology Practical 6.

please circle one number only for each statement

	strongly agree	agree	neutral	disagree	strongly disagree
I feel confident using a computer to work through Practical 6	1	2	3	4	5
I find working through practical 6 a satisfying learning activity	1	2	3	4	5
I did not need the help of demonstrators whilst working through practical 6	1	2	3	4	5
I found the practical 6 computer application confusing to use	1	2	3	4	5
I believe demonstrators should be available at set times for practical 6	1	2	3	4	5
I liked the consistency of explanations offered by practical 6	1	2	3	4	5
I missed demonstrator input whilst working through practical 6	1	2	3	4	5

	strongly agree	agree	neutral	disagree	strongly disagree
Having worked through practical 6, I will always need the help of demonstrators, irrespective of the quality of knowledge offered by the computer application	1....................	2....................	3......................	4.....................	5

4. If you have any comments to make about computers and other learning methods, please write them below.

The Role of Demonstrators

5. The following section relates to the role of demonstrators during practical 6.

a. In the 'Times per Practical' column of the table below, please indicate how many times on average during practical 6 you would like to have consulted a demonstrator for each of the reasons. You may use the last rows to enter reasons that have not been mentioned.

b. In the 'Rank Importance' column of the table below, please rank the importance of your reasons for requiring demonstrator assistance during practical 6. Rank 1 is most important, 10 is least important. Only rank those reasons that you experienced personally.

Reason for requiring demonstrator assistance during practical 6	Times per Practical	Rank Importance
Discuss pathology material being taught other than Histology		
Explain how to use the multiple choice window		
Discuss Histology		
Discuss background information about any aspect of the material being taught		
Discuss how the material being taught relates to other areas of medicine		
Clarify aims and objectives of practical session		
Explain how to start the application		
Explain how to use the application features (i.e. video, image zoom)		

The role of demonstrators during practical 6 cont.

c. Please indicate the methods you used to solve the problems encountered during practical 6 by placing the letter for each solution in the corresponding 'Enter Solution' column of the table below.

You may use each solution more than once, and more than one solution may be used for every 'reason for requiring demonstrator assistance'.

Reason for requiring demonstrator assistance	Enter Solution	Solutions (Ways you solved problems with the practicals)
Histology		A: text books
Other Material		B: help provided by the pathology application
MCQ window		C: asked in tutorial
Discuss background information		D: asked a member of staff outside of tutorial
Discuss how the material relates to other areas		E: asked another student
Clarify aims and objectives		F: never found the answer
How to start the application		G: read lecture notes / handouts
How to use the video, image zoom etc.		H: found the answer in the application
Enter other reason:		I: Enter other solution:
Enter other reason:		J: Enter other solution:

Pathology Application Improvements

6. **We are continually looking for new ways to improve the Pathology application. This section gives you an opportunity to express ways in which you believe the application could be improved.**

Using the table below please rank the importance of the proposed improvements in terms of what you would most like to see incorporated into the application. Rank 1 is most important, 10 is least important. You may use the last rows to enter improvements that have not been mentioned.

Proposed Application Improvement	**Rank Importance**
Glossary of terms	
On-line text books	
Ability to ask questions and receive answers about pathology material	
On-line notepad	
More links to background information about pathology material	
Optional practicals	
More animations like the ECG	
Introduction of sound to listen to patient history etc.	

7. **Write in any others comments that you would like to make about any aspect of the pathology practicals.**

Thank you for your co-operation

REFERENCES

Ackermann, E. (1995). Construction and Transference of Meaning Through Form. Constructivism in Education. (Eds). Steffe, L. P., Gale, J. Lawrence Erlbaum Associates, pp. 341-354.

Ackerman, M. S., Malone, T. W. (1990). Answer Garden: a tool for growing organizational memory. Proceedings of the ACM Conference on Office Information Systems. New York: ACM, pp. 31-39.

Allinson, L., Hammond, N. (1989). A learning support environment: the hitch-hiker's guide. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 62-74.

Alpay, L. (1992). Alternate Approaches to Medical Reasoning and Diagnosis. Knowledge

Baird, P., Percival, M. (1989). Glasgow online: database development using Apple's HyperCard. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 75-92.

Barry, C. A. (1995). Critical Issues in Evaluating the Impact of IT on Information Activity in Academic Research: Developing a Qualitative Research Solution. Library & Information Science Research, Vol. 17, pp. 107-134.

Beeman, W. O., Anderson, K. T., Bader, G., Larkin, J., McClard, A. P., McQuilan, P., Shields, M. (1987). Hypertext and pluralism: from lineal to non-lineal thinking. Proceedings of Hypertext `87, University of North Carolina, Chapel Hill, pp. 67-88.

Benest, I. D. (1990). A hypertext system with controlled hype. Hypertext: State of the Art, (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 52-63.

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H. F., Secret, A. (1994). The World Wide Web. Communications of the ACM, Vol. 37, No. 8, pp. 76-82.

Bernstein, M. (1998). Patterns of Hypertext. In: Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia, 20- 24 June, Pittsburgh, Pennsylvania, pp. 21-29.

Blattner, M. M. (1994). In Our Image: Interface Design in the 1990s. IEEE MultiMedia, Spring, pp. 25-36.

Bourges-Waldagg, P., Scrivener S. A. R. (1998). Meaning, the central issue in cross-cultural HCI design. Interacting with Computers, Vol. 9, No. 3, pp. 287-309.

Boyle, C. D. B., Snell, J. R. (1990). Intelligent navigation for semistructured hypertext documents. Hypertext: State of the Art, (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 28-39.

Boyle, T. (1997). Design for Multimedia Learning. London: Prentice Hall.

Bradshaw, J. M. (1997). An Introduction to Software Agents. Software Agents, (Ed). Bradshaw, J. M. AAAI Press/The MIT Press, pp. 3-46.

Brailsford, T. J., Davis, P. M. C., Scarborough, S. C., Trewhella, W. J. (1996). Knowledge Tree: putting discourse into computer-based learning. Alt-J, Vol. 5, No. 1, pp. 19-26.

Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C. K., Williams, S. M. (1990). Anchored Instruction: Why We Need It and How Technology Can Help. Cognition, Education, Multimedia: Exploring Ideas in High Technology. (Eds). Nix, D., Spiro, R. J. Lawrence Erlbaum Associates, pp. 115-141.

Brockmeier, J. (1996). Construction and interpretation: Exploring a joint perspective on Piaget and Vygotsky. Piaget - Vygotsky: The Social Genesis of Thought. (Eds). Tryphon, A., Vonèche, J. Psychology Press, pp. 125-143.

Brown, A. L., Metz, K. E., Campione, J. C. (1996). Social Interaction and individual understanding in a community of learners: The influence of Piaget and Vygotsky. Piaget - Vygotsky: The Social Genesis of Thought. (Eds). Tryphon, A., Vonèche, J. Psychology Press, pp. 145-170.

Bryman, A. (1998). Quantity and Quality in Social Research. London: Routledge.

Bryman, A., Cramer D. (1997). Quantitative Data Analysis with SPSS for Windows: A guide for social scientists. London: Routledge.

Bush, V. (1945). As We May Think. Atlantic Monthly, 45, pp. 101-108.

Calverley, G., Watkins, J. (1995). SToMP TLTP Formal Evaluation Report. Centre for Engineering Educational Technology (CEET), University of Surrey.

Carbonell, J. R. (1970). AI in CAI: an artificial intelligence approach to computer-assisted instruction. IEEE Transactions on Man-Machine Systems, Vol. 11, No. 4, pp. 190-202.

Carnegie, A. (1895). Abstract from the address of the dedication of The Carnegie Library of Pittsburgh. Pittsburgh, 5 November.

Carr, L., Hall, W., Davis, H., De Roure, D., Hollom, R. (1994). The Microcosm Link Service and its Application to the World Wide Web. Proceedings of the First International World Wide Web Conference, Geneva, May.

Carr, L., De Roure, D., Hall, W., Hill, G. (1995). The Distributed Link Service: A Tool for Publishers, Authors and Readers. Proceedings of the Fourth International World Wide Web Conference, Boston, Massachusetts, December.

Carr, L., Hall, W., Hitchcock, S. (1998). Link Services or Link Agents? Proceedings of the Ninth ACM Conference on Hypertext, Pittsburgh, USA, June 20 - 24, ACM Press, pp. 113-122.

Carroll, J. M. (1990). The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skills. The MIT Press.

Charniak, E., McDermott, D. (1985). Introduction to Artificial Intelligence. Addison-Wesley.

Clancey, W. J. (1979). Tutoring rules for guiding a case method dialogue. International Journal of Man Machine Studies, Vol. 11, pp.24-49.

Clancey, W. J. (1997). The Conceptual Nature of Knowledge, Situations, and Activity. Expertise in Context: Human and Machine. (Eds). Feltovich, P. J., Ford, K. M., Hoffman, R. R. AAAI Press/The MIT Press, pp. 247-291.

Colbourn, C. J., Cockerton-Turner, T. (1990). Using hypertext for educational "Help" facilities. Hypertext: State of the Art, (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 105-111.

Cooke, P., Williams, I. (1989). Design issues in large hypertext systems for technical documentation. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 93-104.

Conklin, J. (1987). Hypertext: An Introduction and Survey. IEEE Computer, 20, pp.17-41.

Coulouris, G., Dollimore, J., Kinberg, T. (1994). Distributed Systems: Concepts and Design. Second Edition. Addison-Wesley.

Cuban, L. (1986). Teachers and Machines: The classroom use of technology since 1920. New York: Teachers College Press.

Culley, L. (1991). Girls, Boys and Computers. Computers and Learning, (Eds). Boyd-Barrett, O., Scanlon, E. Addison-Wesley, pp. 81-87.

Darrow, B. (1932). Radio: The Assistant Teacher. Columbus, OH: R., G., Adams.

Davis, H., Hall, W., Heath, I., Hill, G., Wilkins, R. (1992a). Towards an Integrated Information Environment with Open Hypermedia Systems. Proceedings of the ACM European Conference on Hypertext, Milan, Italy, 30 November - 4 December, ACM Press, pp. 181-190.

Davis, H., Hall, W., Hutchings, G., Rush, D., Wilkins, R. (1992b). Hypermedia and the Teaching of Computer Science: Evaluating an Open System. Proceedings of the Conference on Development in the Teaching of Computer Science, University of Kent, pp. 122-129.

Davis, H., Hutchings, G., Hall, W. (1993). A Framework for Delivering Large-Scale Hypermedia Learning Material. Educational Multimedia and Hypermedia Annual 1993, Proceedings of ED-MEDIA '93 - World Conference on Educational Multimedia and Hypermedia, Florida, USA, June, (Ed). Maurer, H. ACCE, pp. 115-122.

Davis, H. C., Knight, S., Hall, W. (1994). Light Hypermedia Link Services: A Study of Third Party Application Integration. European Conference on Hypertext Technology, ECHT '94, Edinburgh, 18-23 September, pp. 41-50.

De Corte, E. (1996). Changing Views of Computer-Supported Learning Environments for the Acquisition of Knowledge and Thinking Skills. International Perspective on the Design of Technology Supported Learning Environments, (Eds). Italle, V., De Corte, E., Glaser, R., Mandl, H. Lawrence Erlbaum Associates, pp.129-145.

DeRoure, D. C., Hall, W., Davis, H. C., Dale, J. (1996). Agents for Distributed Multimedia Information Management. Proceedings of the First International Conference on the Practical Application of Agents and Multi-Agent Systems, London, UK, April 22-24.

Duchastel, P. C. (1991). Examining Cognitive Processing in Hypermedia Usage. Hypermedia, Vol. 2, No. 3, pp. 221-233.

Duncan, E. B. (1989). A faceted approach to hypertext? Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 157-163.

Edwards, D. M., Hardman, L. (1989). 'Lost in hyperspace': cognitive mapping and navigation in a hypertext environment. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 105-125.

Engelbart, D. C. (1963). A Conceptual Framework for the Augmentation of Man's Intellect. Spartan Books.

Engelbart, D. C. (1984). Collaboration Support Provisions in Augment. Proceedings of the 1984 AFIPS Office Automation Conference, California, USA, February, pp. 51-58.

Ericsson, K. A., Simon, H. A. (1984). Protocol Analysis, Verbal Reports as Data. Cambridge, MA: The MIT Press.

Ernest, P. (1995). The One and the Many. Constructivism in Education. (Eds). Steffe, L. P., Gale, J. Lawrence Erlbaum Associates, pp. 459-486.

Fenton, N. E., Pfleeger, S. L. (1997). Software Metrics, A Rigorous & Practical Approach, Second Edition. PWS Publishing Company.

Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, D. (1997). Query by Image and Video Content: The QBIC™ System. Intelligent Multimedia Information Retrieval. (Ed). Maybury, M. T. AAAI Press/The MIT Press, pp.7-22.

Fountain, A. M., Hall, W., Heath, I., Davis, H. (1990). Microcosm: An Open Model for Hypermedia with Dynamic Linking. European Conference on Hypertext Technology, ECHT '90, France, November, pp. 298-311.

Fraser, K. (1997). Enhancing the pedagogical value of multimedia products in higher education. International Journal of Academic Development, Vol. 2, No. 1, pp. 60-65.

Furht, B. (1994). Multimedia Systems: An Overview. IEEE MultiMedia, Vol. 1, No. 1, Spring, pp. 47-59.

Furuta, R., Shipman III, F. M., Marshall, C. C., Brenner, D., Hsieh, H. (1997). Hypertext Paths and the World-Wide Web: Experiences with Walden's Paths. Proceedings of The Eight ACM Conference on Hypertext, Hypertext '97, Southampton, UK, April 6 - 11, ACM Press, pp. 167-176.

Gagné, R. M. (1965). The Conditions of Learning. New York: Holt Rhinehart and Winston.

Gagné, R. M., Briggs, L. J. (1979). Principles of Instructional Design. New York: Holt Rhinehart and Winston.

General Medical Council. (1993). Tomorrow's Doctors: Recommendations on Undergraduate Medical Education. Issued by the Education Committee of the General Medical Council.

Grant, J. (1997). Research and Evaluation Methods: The Inside Story. Presented at the UK Conference of Directors and Associate Advisors: General Practice Education, London, UK, 11 June.

Hall, J. L., Colbourn, C., Light, P., Bailey, J. D., Reed, P. A. S. (1995a). Hypermedia in Education: Students' user of a Learning Resource on Phase Diagrams. Proceedings of Computer-Aided Learning in Engineering, Sheffield, UK, 5 - 7 September.

Hall, J. L., Colbourn, C., Light, P. H. (1995b). From Assertion to Evidence: The Use of Educational Technology in an Undergraduate Medical Degree. Proceedings of the Second Groningen International Information Technology Conference, University of Groningen, The Netherlands.

Hall, W., Thorogood, P., Sprunt, B., Carr, L., Hutchings, G. (1990). Is hypermedia an effective tool for education? Computers in Education, pp. 1067-1073.

Hall, W., Colson, F. R. (1991). Multimedia Teaching with Microcosm-Hides: Viceroy Mountbatten and the Partition of India. History and Computing, Vol. 3, No. 2, pp. 89-98.

Hall, W., Davis, H. (1994). Hypermedia link services and their application to multimedia information management. Information and Software Technology, Vol. 36, No. 4, pp. 192-202.

Hall, W., Davis, H., Hutchins, G. (1996). Rethinking Hypermedia - The Microcosm Approach. Kluwer Academic Press.

Hall, W., Hutchings, G., White, S. (1995c). Breaking Down the Barriers: an Architecture for Developing and Delivering Resource-Based Learning Materials. Proceedings of the World Wide Conference on Computers in Education, Birmingham, UK.

Halsall, F. (1992). Data Communications, Computer Networks and Open Systems. Addison-Wesley.

Hammond, N., Allinson, L. (1989). Extending hypertext for learning: an investigation of access and guidance tools. People and Computers V, (Eds). Sutcliffe, A., Macaulay, L. Cambridge: Cambridge University Press.

Hauck, F. J. (1996). Supporting hierarchical guided tours in the World Wide Web. Proceedings of the Fifth International World Wide Web Conference, France, May 6 - 10, pp. 1233-1242.

Hawkridge, D. (1991). Creative Gales and Computers in Third World Schools. Computers and Learning, (Eds). Boyd-Barrett, O., Scanlon, E. Addison-Wesley, pp. 23-38.

Hawkridge, D. (1996). The Next Educational Technology in Higher Education. Innovations in Education and Training International, Vol. 33, No 1, pp. 5-12.

Henderson, J. V. (1995). Needed: new models for network-based learning. Educational Multimedia and Hypermedia 1995, Proceedings of ED-MEDIA '95 - World Conference on Educational Multimedia and Hypermedia, Graz, Austria, June, (Ed). Maurer, H. ACCE.

Hill, G., Wilkins, R., Hall, W. (1993). Open and Reconfigurable Hypermedia Systems: A Filter-Based Model. Hypermedia, Vol. 5, No. 2, pp. 103-118.

Hill, G., Hall, W., De Roure, D., Carr, L. (1995). Applying Open Hypertext Principles to the WWW. Proceedings of the International Workshop on Hypermedia Design '95, Montpellier, France.

Hoffer, E. P., Barnett, G. O. (1990). Computers in Medical Education. Medical Informatics: Computer Applications in Health Care. (Eds). Shortliffe, E. H., Perreault, L. E. (Assoc Eds). Wiederhold, G., Fagan, M. Addison-Wesley, pp. 535-561.

Ho-Shing, H., Chan, S. (1997). Hypertext-Assisted Video Indexing and Content-Based Retrieval. Proceedings of The Eight ACM Conference on Hypertext, Hypertext '97, Southampton, UK, April 6 - 11, ACM Press, pp. 232-233.

Hutchings, G. A., Carr, L. A., Hall, W. (1992a). Stackmaker: An Environment for Creating Hypermedia Learning Material. Hypermedia, Vol. 4, No. 3, pp. 198-211.

Hutchings, G. A., Hall, W., Colbourn, C. J. (1993). Patterns of students' interactions with a hypermedia system. Interacting with Computers, Vol. 5, No. 3, pp. 295-313.

Hutchings, G. A., Thorogood, P., Hall, W. (1992b). Hypermedia in biomedical education: a case study. Journal of Audiovisual Media in Medicine, 15, pp. 77-80.

Hutchings, G. A., Hall, W., Thorogood, P. (1994). Experiences with Hypermedia in Undergraduate Education. Computers in Education, Vol. 22, No. 1/2, pp. 39-44.

Illich, I. (1970). Deschooling Society. London: Calder and Boyars.

Ivic, I. (1997). Lev S. Vygotsky[1] (1896-1934). Thinkers on Education: Volume 4. (Ed). Morsy, Z. UNESCO Publishing/Oxford & IBH Publishing, pp. 761-785.

Jacobs, G. (1998). Evaluating Courseware: Some Critical Questions. Innovations in Education and Training International, Vol. 35, No. 1, pp. 3-8.

Jelovsek, F. R., Adebonojo, L. (1993). Learning Principles as applied to Computer-Assisted Instruction. M D Computing, Vol. 10, No. 3, pp. 165-172.

Jonassen, D. H. (1990). Semantic network elicitation: tools for structuring hypertext. Hypertext: State of the Art, (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 142-152.

Jones, S. (1998). Media Use in an Electronic Community. Network & Netplay: Virtual Groups on the Internet. (Eds). Sudweeks, F., McLaughlin, M., Rafaeli, S. AAAI Press/The MIT Press, pp. 77-94.

Kemp, M. W., DuBoulay, C. E. H., Millard, L., Hall. W., Roche W. R. (1997). Microcosm or Microscopy? A Cross-over Controlled Trial. Poster presentation at the 175th Meeting of the Pathological Society of Great Britain and Ireland, Sheffield, 2 - 4 July.

Kemp, M. W., Hall, W., Roche W. R. (1999). An Agent Model for Student-Tutor Discourse and Knowledge Misconception Resolve. Paper presented at CAL 99: Virtuality in Education, The Institute of Education, 29 - 31 March.

Kemp, M. W., Hothi, J., Hall, W., Roche W. R. (1998). Evaluation of Hypermedia Courseware Applications and Tools. Paper presented at LTN 98: Implementing Learning Technologies: Strategies and Experience, Southampton 6 - 8 April.

Kibby, M. R., Mayes, J. T. (1989). Towards intelligent hypertext. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 164-172.

Kidd, M. R., Hutchings, G. A., Hall, W., Cesnik, B. (1992). Applying Hypermedia to Medical Education: An Author's Perspective. Education and Training Technology International, Vol. 29, No.2, pp 143-151.

Kitchenham, B. A. (1996a). Evaluating Software Engineering Methods and Tool, Part 1: The Evaluation Context and Evaluation Methods. Software Engineering Notes, Vol. 21, No. 1, January, pp. 11-15.

Kitchenham, B. A. (1996b). Evaluating Software Engineering Methods and Tool, Part 3: Selecting an appropriate evaluation method – practical issues. Software Engineering Notes, Vol. 21, No. 4, July, pp. 9-12.

Kitchenham, B. A., Jones. L. (1997a). Evaluating Software Engineering Methods and Tool, Part 5: The Influence of Human Factors. Software Engineering Notes, Vol. 22, No. 1, January, pp. 13-15.

Kitchenham, B. A., Jones, L. (1997b). Evaluating Software Engineering Methods and Tool, Part 6: Identifying and Scoring Features. Software Engineering Notes, Vol. 22, No. 2, March, pp. 16-18.

Kitchenham, B. A., Jones, L. (1997c). Evaluating Software Engineering Methods and Tool, Part 7: Planning Feature Analysis Evaluation. Software Engineering Notes, Vol. 22, No. 4, July, pp. 21-24.

Kitchenham, B. A., Jones, L. (1997d). Evaluating SW Eng. Methods and Tools, Part 8: Analysing a Feature Analysis Evaluation. Software Engineering Notes, Vol. 22, No. 5, September, pp. 10-12.

Kitchenham, B. A., Pickard, L. M. (1998). Evaluating Software Engineering Methods and Tools, Part 9: Quantitative Case Study Methodology. Software Engineering Notes, Vol. 23, No. 1, January, pp. 24-26.

Kirkwood, A. (1996). Convergence and Media for Teaching and Learning. Innovations in Education and Training International, Vol. 33, No. 1, pp. 41-48.

Laurel, B. (1997). Interface Agents: Metaphors with Character. Software Agents, (Ed). Bradshaw, J. M. AAAI Press/The MIT Press, pp. 67-77.

Laurillard, D. (1991). Computers and the Emancipation of Students: Giving Control to the Learner. Computers and Learning, (Eds). Boyd-Barrett, O., Scanlon, E. Addison-Wesley, pp. 64-80.

Laurillard, D. (1993). Rethinking University Teaching: A Framework for the Effective Use of Educational Technology. London: Routledge.

Levenson, W. (1945). Teaching Through Radio. New York: Farrar and Rinehart.

Lewis, P. H., Davis, H. C., Griffiths, S. R., Hall, W., Wilkins, R. J. (1996). Media-based Navigation with Generic Links. Proceedings of the ACM Conference on Hypertext, Washington DC, USA, March, ACM Press.

Li, Z., Davis, H., Hall, W. (1992). Hypermedia Links and Information Retrieval. The proceedings of the 14th British Computer Society Research Colloquium on Information Retrieval, Lancaster University, U.K.

Magnus, J. (1996). Distance Education in Sub-Saharan Africa: the Next Five Years. Innovations in Education and Training International, Vol. 33, No. 1, February, pp. 50-56.

Marmann, M., Schlageter, G. (1992). Towards a Better Support for the Hypermedia Structuring: The HYDESIGN Model. Proceedings of the ACM European Conference on Hypertext, Milan, Italy, 30 November - 4 December, ACM Press, pp. 232-241.

Martí, E. (1996). Mechanisms of internalisation and externalisation of knowledge in Piaget's and Vygotsky's theories. Piaget - Vygotsky: The Social Genesis of Thought. (Eds). Tryphon, A., Vonèche, J. Psychology Press, pp. 57-83.

Mason, J. (1994). Linking qualitative and quantitative data analysis. Analyzing Qualitative Data, (Eds). Bryman, A., Burgess, R. G. London: Routledge, pp. 89-110.

Masterton, S. (1998). The Virtual Participant: a tutor's assistant for electronic conferencing. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 249-265.

Mayes, J. T. (1993). Commentary: Impact of cognitive theory on the practice of courseware authoring. Journal of Computer Assisted Learning, Vol. 9, pp. 222-228.

Mayes, J. T. (1995). Learning Technology and Groundhog Day. Hypermedia at work: Practice and Theory in Higher Education, (Eds). Strang, W., Simpson, V., Slater, D. Canterbury, University of Kent Press.

Mayes, J. T. (1997). Dialogue with a dumb terminal. Article published in The Times HIGHER supplement, October 10.

Mayes, J. T., Kibby, M. R., Watson, H. (1988). StrathTutor: the development and evaluation of a learning-by-browsing system on the Macintosh. Computers and Education, 12, pp. 221-229.

Mayes, J. T., Kibby, M. R., Anderson, T. (1990a). Signposts for conceptual orientation: some requirements for learning from hypertext. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 121-129.

Mayes, J. T., Kibby, M. R., Anderson, T. (1990b). Learning about learning from hypertext Designing hypertext/hypermedia for learning. (Eds). Jonassen, D. H., Mandl, H. Heidelburg: Springer-Verlag.

McAleese, R. (1989). Navigation and browsing in hypertext. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 6-44.

McCullagh, P., Roche, W. R. (1992). First year pathology practicals, The Southampton Experience. The Journal of Pathology, Vol. 168, Supplement 148^A, July.

McKendree, J., Mayes, J. T. (1997). "The Vicarious Learner": investigating the benefits of observing peer dialogues. Proceedings of CAL 97, Exeter, England, 23-26 March.

McKnight, C., Dillon, A., Richardson, J. (1990). A comparison of linear and hypertext formats in information retrieval. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 10-19.

McKnight, C., Dillon, A., Richardson, J. (1991). Hypertext in Context. Cambridge University Press.

McKnight, C., Richardson, J., Dillon, A. (1989). The Authoring of Hypertext Documents. Hypertext: theory into practice. Second Edition. (Ed). McAleese, R. Oxford: Intellect, pp. 138-147.

Melly, M. (1995). Cooperative Working in an Open Hypermedia Environment. PhD thesis, University of Southampton.

Mendes, M. E. X. (1999). Metrics for Improving the Quality of Hypermedia Authoring. PhD Thesis, University of Southampton.

Mendes, M. E. X., Hall, W. (1997a). An empirical study of hypermedia authoring for education. Proceedings of CAL 97, Exeter, England, 23-26 March.

Mendes, M. E. X., Hall, W. (1997b). The SHAPE of Hypermedia Authoring for Education. Proceedings of ED-MEDIA & ED-TELECOM 97, Calgary, Canada.

Mendes, M. E. X., Hall, W. (1999). Hyper-authoring for education: a qualitative evaluation. Computers and Education, 32, pp. 51-64.

Monk, A. F. (1990). Getting to known locations in a hypertext. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 20-27.

Moser, C. A., Kalton, G. (1979). Survey Methods in Social Investigation. London: Heinemann Educational Books.

Moyse, R., Elsom-Cook, M. (1992). Knowledge Negotiation: An Introduction. Knowledge Negotiation. (Eds). Moyse, R., Elsom-Cook, M. London: Academic Press, pp. 1-19.

Mukherjea, S., Foley, J. D., Hudson, S. E. (1994). Interactive Clustering for Navigating in Hypermedia Systems. European Conference on Hypertext Technology, ECHT '94, Edinburgh, 18-23 September, pp. 136-145.

Munari, A. (1997). Jean Piaget (1896-1980). Thinkers on Education: Volume 3. (Ed). Morsy, Z. UNESCO Publishing/Oxford & IBH Publishing, pp. 311-327.

Negroponte, N. (1997). Agents: From Direct Manipulation to Delegation. Software Agents, (Ed). Bradshaw, J. M. AAAI Press/The MIT Press, pp. 58-66.

Nelson, T. H. (1988). Managing immense storage. Byte, Vol. 13, No. 1, pp. 225-238.

Nelson, T. H. (1994). Invited talk at MediaActive '94, Liverpool. Available on CDROM , Liverpool John Moores University, Learning Methods Unit.

Nicolson, R. I. (1990). Towards the third generation: the case for IKBH (intelligent knowledge-based hypermedia) environments. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 166-173.

Nielson, J., Lyngaek, U. (1990). Two field studies of hypermedia usability. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 64-72.

Nix, D. (1990). Should Computers Know What You Can Do With Them? Cognition, Education, Multimedia: Exploring Ideas in High Technology. (Eds). Nix, D., Spiro, R. J. Lawrence Erlbaum Associates, pp. 143-162.

Oppenheim, A. N. (1998). Questionnaire Design, Interviewing and Attitude Measurement. London: Pinter.

O'Shea, T., Self, J. (1983). Learning and Teaching with Computers. Brighton: Harvester Press.

Papert, S. (1980). Mindstorms: children, computers and powerful ideas. Brighton: Harvester Press.

Papert, S. (1990). Introduction. Constructionist Learning, (Ed). Harel, I. Cambridge, MA: A MIT Media Laboratory Publication. pp. 1-8.

Papert, S. (1993). The Children's Machine: Rethinking School in the Age of the Computer. New York: Harvestor Wheatsheaf.

Pentland, A. (1997). Machine Understanding of Human Behavior in Video. Intelligent Multimedia Information Retrieval. (Ed). Maybury, M. T. AAAI Press/The MIT Press, pp. 175-188.

Perelman, L. J. (1992). School's Out: Hyperlearning, the New Technology and the End of Education. New York: William Morrow.

Petre, M., Carswell, L., Price, B., Thomas, P. (1998). Innovations in large-scale supported distance teaching: transformation for the Internet, not just translation. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 97-116.

Piaget, J. (1963). The Child's Conception of the World. New Jersey: Littlefield Adams.

Piper, A., Wright, V., Hall, W., White, S. (1995). Using Hypermedia for Language Learning: A Resource-Based Environment. Innovations in Education and Training International, Vol. 32, No. 3, pp. 220-228.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., Carey, T. (1994). Human-Computer Interaction. Addison-Wesley.

Pressman, R. S. (1994). Software Engineering: A Practitioner's Approach (Third Edition, European Adaptation). McGraw-Hill.

Ramsden, P. (1992). Learning to Teach in Higher Education. London: Routledge.

Richmond, A. (1996). Enticing online shoppers to buy - A human behaviour study. Proceedings of the Fifth International World Wide Web Conference, France, May 6 - 10, pp. 1469-1490.

Ritchie, J., Spencer, L. (1994). Qualitative data analysis for applied policy research. Analyzing Qualitative Data. (Eds). Bryman, A., Burgess, R. G. London: Routledge, pp. 173-194

Rouet, J. F. (1992). Cognitive Processing of Hyperdocuments: When Does Nonlinearity Help? Proceedings of the European Conference on Hypertext Technology, ECHT '92, Milano Italy, November 30 - December 4, pp. 131-140.

Saettler, P. (1968). A History of Instructional Technology. New York: McGraw-Hill.

Schank, R. C. (1994). Active Learning Through Multimedia. IEEE MultiMedia, Vol. 1, No. 1, Spring, pp. 69-78.

Schank, R. C., Cleary, C. (1995). Engines For Education. Hillsdale: Lawrence Erlbaum Associates.

Scott, P., Eisenstadt, M. (1998). Exploring telepresence on the Internet: the KMi Stadium Webcast experience. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 154-172.

Scott, P., Phillips, M. (1998). Developing Web-based student support systems: telling student stories on the Internet. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 79-95.

Self, J. A. (1989). The Case for Formalising Student Models (and Intelligent Tutoring Systems generally). Paper presented at the 4th Conference on Artificial Intelligence and Education, Amsterdam, Netherlands.

Self, J. A. (1992). Computational Viewpoints. Knowledge Negotiation. (Eds). Moyse, R., Elsom-Cook, M. London: Academic Press, pp. 21-40.

Shadbolt, N., O'Hara, K. (1997). Model-Based Expert Systems and the Explanation of Expertise. Expertise in Context: Human and Machine. (Eds). Feltovich, P. J., Ford, K. M., Hoffman, R. R. AAAI Press/The MIT Press, pp. 315-337.

Shoham, Y. (1997). An Overview of Agent-Oriented Programming. Software Agents, (Ed). Bradshaw, J. M. AAAI Press/The MIT Press, pp. 271-290.

Shortliffe, E. H. (1976). Computer-Based Medical Consultations: MYCIN. New York: Elsevier.

Shortliffe, E. H. (1990). Clinical Decision-Support Systems. Medical Informatics: Computer Applications in Health Care. (Eds). Shortliffe, E. H., Perreault, L. E. (Assoc Eds). Wiederhold, G., Fagan, M. Addison-Wesley, pp. 466-502.

Siegel, S. (1956). Nonparametric Statistics for the Behavioural Scicences. McGraw-Hill.

Simpson, A., McKnight, C. (1990). Navigation in hypertext: structural cues and mental maps. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 73-83.

Sinclair, B., Bacon, R. (1997). Learning with SToMP: flexibility in using a complete environment. Proceedings of CAL 97, Exeter, England, 23-26 March.

Skinner, B. F. (1953). Science and Human Behaviour. New York: MacMillan.

Skinner, B. F. (1960). Teaching Machines. Teaching Machines and Programmed Learning. (Eds). Lumsdain, A. A., Glaser, R. Nat'l Ed. Assoc., Washington, D.C., pp. 137-158.

Smith, C. B., McLaughlin, M. L., Osborne, K. K. (1998). From Terminal Ineptitude to Virtual Sociopathy: How Conduct is Regulated on Usenet. Network & Netplay: Virtual Groups on the Internet. (Eds). Sudweeks, F., McLaughlin, M., Rafaeli, S. AAAI Press/The MIT Press, pp. 95-112.

Smith, D. C., Cypher, A., Spohrer, J. (1997). KidSim: Programming Agents without a Programming Language. Software Agents, (Ed). Bradshaw, J. M. AAAI Press/The MIT Press, pp. 165-190.

Smith, J., Weiss, S. (1991). Hypertext. Computers and Learning, (Eds). Boyd-Barrett, O., Scanlon, E. Addison-Wesley, pp. 231-234.

Smith, L. M. (1997). B. F. Skinner (1904-90). Thinkers on Education: Volume 4. (Ed). Morsy, Z. UNESCO Publishing/Oxford & IBH Publishing, pp. 519-531.

Speedie, S. M., McKay, A. B. (1990). Pharmacy Systems. Medical Informatics: Computer Applications in Health Care. (Eds). Shortliffe, E. H., Perreault, L. E. (Assoc Eds). Wiederhold, G., Fagan, M. Addison-Wesley, pp. 298-323.

Spiro, R. J., Jehng, J. C. (1990). Cognitive Flexibility and Hypertext: Theory and Technology for the Nonlinear and Multidimensional Traversal of Complex Subject Matter. Cognition, Education, Multimedia: Exploring Ideas in High Technology. (Eds). Nix, D., Spiro, R. J. Lawrence Erlbaum Associates, pp. 163-205.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., Coulson, R. L. (1995). Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction for Advanced Knowledge Acquisition in Ill-Structured Domains. Constructivism in Education. (Eds). Steffe, L. P., Gale, J. Lawrence Erlbaum Associates, pp. 85-107.

Squires, D. (1997). Exploring the use of interactive information systems in academic research: borrowing from the ethnographic tradition. Education for Information, 15, IOS Press, pp. 323-330.

Stanton, N. A., Stammers, R. B. (1990). Learning styles in a non-linear training environment. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 114-120.

Stark, H. A. (1990). What do readers do to pop-ups, and pop-ups do to readers? Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 2-9.

Stevens, A. L., Collins, A., Goldin, S. (1979). Misconceptions in Students' Understanding. International Journal of Man-Machine Studies, 11, pp. 145-156.

Stratfold, M. (1998). Promoting learner dialogues on the Web. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 119-134.

Summer, T., Taylor, J. (1998). Media integration through meta-learning environments. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 63-78.

Suppes, P. (1980). The Future of Computers in Education. The Computer in the School: Tutor, Tutee, Tool. (Ed). Taylor, R. P. New York: Teachers College Press.

Sutcliffe, A., Hare, M., Doubleday, A., Ryan, M. (1997). Empirical Studies in Multimedia Information Retrieval. Intelligent Multimedia Information Retrieval. (Ed). Maybury, M. T. AAAI Press/The MIT Press, pp.449-472.

Thorndike, E. L. (1931). Human Learning. New York: Century.

Turban, E., Aronson, J. E. (1998). Decision Support Systems and Intelligent Systems. New Jersey: Prentice Hall.

Voiskounsky, A. E. (1998). Investigation of Relcom Network Users. Network & Netplay: Virtual Groups on the Internet. (Eds). Sudweeks, F., McLaughlin, M., Rafaeli, S. AAAI Press/The MIT Press, pp. 113-126.

von Glaserfield, E. (1995). A Constructivist Approach to Teaching. Constructivism in Education. (Eds). Steffe, L. P., Gale, J. Lawrence Erlbaum Associates, pp. 3-15.

Vygotsky, L . S. (1978). Mind in Society: The Development of Higher Psychological Processes. (Eds). Cole, M., John-Steiner, V., Scribner, S., Souberman, E. Harvard University Press.

Wenger, E. (1987). Artificial Intelligence and Tutoring Systems: Computational and Cognitive Approaches to the Communication of Knowledge. Morgan Kaufmann.

Whalley, P. (1998). Collaborative learning in networked simulation environments. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 47-61.

Wiederhold, G., Perreault, L. E. (1990). Clinical Research Systems. Medical Informatics: Computer Applications in Health Care. (Eds). Shortliffe, E. H., Perreault, L. E. (Assoc Eds). Wiederhold, G., Fagan, M. Addison-Wesley, pp. 503-534.

Witmer, D. F. (1998). Practicing Safe Computing: Why People Engage in Risky Computer-mediated Communication. Network & Netplay: Virtual Groups on the Internet. (Eds). Sudweeks, F., McLaughlin, M., Rafaeli, S. AAAI Press/The MIT Press, pp. 127-146.

Witmer, D. F., Katzman, S. L. (1998). Smile When Your Say That: Graphic Accents as Gender Markers in Computer-mediated Communication. Network & Netplay: Virtual Groups on the Internet. (Eds). Sudweeks, F., McLaughlin, M., Rafaeli, S. AAAI Press/The MIT Press, pp. 3-11.

Woolf, B. P., Hall, W. (1995). Multimedia Pedagogues: Interactive Systems for Teaching and Learning. IEEE Multimedia, Vol. 28, No. 5, May, pp. 74-80.

Wright, P., Lickorish, A. (1990). An empirical comparison of two navigation systems for two hypertexts. Hypertext: State of the Art. (Eds). McAleese, R., Green, C. Oxford: Intellect, pp. 84-93.

Yan, T. W., Jacobsen, M., Garcia-Molina, H., Dayal, U. (1996). From user access patterns to dynamic hypertext linking. The proceedings of the Fifth International World Wide Web Conference, Paris, France, May 6 - 10, pp. 1007-1014.

Yankelovich, N., Meyrowitz, N., van Dam, A. (1991). Reading and Writing the Electronic Book. Computers and Learning, (Eds). Boyd-Barrett, O., Scanlon, E. Addison-Wesley, pp.235-249.

Yankelovich, N., Haan, B. J., Meyrowitz, N., Drucker, S. M. (1988). Intermedia: The Concepts and the Construction of a Seamless Information Environment. IEEE Computer, January, pp. 81-96.

Zdrahal, Z., Domingue, J. (1998). The World Wide Design Lab: an environment for distributed collaborative design. The Knowledge Web: Learning and Collaborating on the Net, (Eds). Eisenstadt, M., Vincent, T. London: Kogan Page, pp. 225-238.

Zhang, H. J., Low, C. Y., Smoliar, S. W., Wu, J. (1997). Video Parsing, Retrieval and Browsing: An Integrated and Content-Based Solution. Intelligent Multimedia Information Retrieval. (Ed). Maybury, M. T. AAAI Press/The MIT Press, pp.139-158.

Zizi, M., Beaudouin-Lafon, M. (1994). Accessing Hyperdocuments through Interactive Dynamic Maps. European Conference on Hypertext Technology, ECHT '94, Edinburgh, 18-23 September, pp. 126-135.