Towards Diagrammatic Hypermedia Authoring: Cognition and Usability Issues in Higher Education

Geoff Elliott 1999

Abstract

Lack of skill, time and institutional support are widely cited as obstacles to the development of educational-hypermedia by higher education staff. The problem is how can computer-based tools be made accessible to teachers/academics but at the same time remain sufficiently expressive to produce educationally-effective hypermedia? In this study, human-computer interaction issues relating to the authoring of hypermedia by academics are examined and a range of authoring programs evaluated for their suitability. The main experimental hypothesis of the study was that a diagram-based hypermedia authoring program would help to mitigate against the skills deficit in that it would be very easy to use and hence enable academics with relatively low computer skills to develop hypermedia. The investigation involved a number of research activities including: trials of a diagram-based and a book-based authoring program using sixteen subjects; the development of a full application for use in a higher education teaching scenario; and an evaluation of a number of other representative authoring programs by a focus group of experts in IT training. Analysis of both quantitative and qualitative data indicated that diagram-based authoring programs have significant advantages over existing authoring programs for people with lower computer skills, although this was compounded by spatial relations ability. The main contribution of this study is a causal model of learnability of hypermedia authoring programs derived from the analysis of the empirical data. The model incorporates a number of factors, some already noted in other studies, some identified in this study. The causal model of learnability extends contemporary views of learnability and usability and is underpinned by contemporary learning theories. The main conclusion of the study is that new authoring programs designed for use in higher education should take into account the model of learnability, be based on an accessible diagramming technique such as concept maps for knowledge construction and be integrated with a visual development language that allows the passive hypermedia-knowledge to be supplemented with educationally-effective functionality.

Table of Contents

Appendices

Table of Figures

List of Tables

List of Abbreviations

AI   Artificial Intelligence

CSCW   Computer Supported Collaborative Working

CTI   Computers in Teaching Initiative

EL   Ease of Learning

EU   Ease of Use

GUI   Graphical User Interface

HAP   Hypermedia Authoring Program

HCI   Human-Computer Interaction

HE   Higher Education

HTML   HyperText Markup Language

IT   Information Technology

QDA   Qualitative Data Analysis

RBL   Resource Based Learning

TLTP   Teaching and Learning Technology Programme

UK   United Kingdom
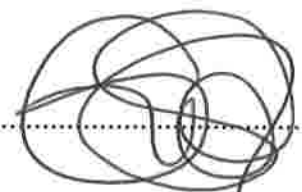
Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed..(Candidate)

Date.......24.6.99.......

This thesis is the result of my own investigations, except where otherwise stated. All sources are fully acknowledged in the text with explicit references. A bibliography is appended.

Signed..(Candidate)

Date.......24.6.99.......

Signed..(Director of Studies)

Signed..(Supervisor)

I hereby give consent for my thesis, if accepted, to be available for photocopying and for interlibrary loan, and for the title and summary to be made available to outside organisations.

Signed..(Candidate)

Date.......24.6.99.......

Acknowledgements

I would like to thank my Director of Studies, Doctor Eleri Jones, whose constant enthusiasm, encouragement and guidance has made this study possible but most importantly enjoyable. I would also like to thank Professor Philip Barker whose help and guidance has been invaluable, despite being based over 200 miles away.

Thanks to the staff in the Faculty of Business, Leisure and Food, UWIC, for their support and inspiration, particularly Greg Dainty and David Botterill, members of the staff who kindly volunteered to be the subjects of my research, without them there would be no new knowledge, Derek Smith for his help in the area of cognitive psychology. and also to Pam Briggs at the University of Northumbria in Newcastle whose advice on usability evaluation was very helpful.

Finally, thanks to my wife, Tracy, for her support and for kindly putting up with me when this study became obsessive. I would like to dedicate this thesis to my daughter Mair, who was the perfect distraction when I needed to wind down, and my son Elis who was born the same week I completed this thesis.

Chapter One

Introduction

1. Introduction

> *One day, when Christopher Robin and Winnie-the-Pooh and Piglet*
> *were all talking together, Christopher Robin finished the mouthful he*
> *was eating and said carelessly: 'I saw a Heffalump to-day, Piglet.'*
> *'What was it doing?' asked Piglet.*
> *'Just lumping along,' said Christopher Robin. 'I don't think it saw me.'*
> *'I saw one once,' said Piglet. 'At least, I think I did,' he said. 'Only*
> *perhaps it wasn't.'*
> *'So did I,' said Pooh, wondering what a Heffalump was like.*

Winnie the Pooh, A. A. Milne.

Hypermedia and other educational technologies have created much interest in higher education and there is considerable research activity taking place to develop and explore the use of the technologies in higher education (HE). The current interest in hypermedia particularly with the advent of the World Wide Web (Web) is similar to the interest that was generated with the advent of cheap video recording and playback equipment - when it was predicted that video would have a profound effect on HE delivery. To date, the effect of video remains minimal - even with large increases in student numbers, most HE in the UK is still delivered in the same way it was 30 or 40 years ago. It is still too early to judge what effect hypermedia and other educational technologies will have on HE, however judging by the impact of the government sponsored research programs, it does appear that change will take decades rather than years.

This study was born out of an enthusiasm for the use of information technology (IT), an experience of teaching large classes and, as Laurillard (1993) similarly describes in her influential book, *Rethinking University Teaching,* personal experiences of learning in HE. Although the government-sponsored research programmes like the Computers in Teaching Initiative (Darby, 1993a) have generated some very interesting material, the effect this has had on most academics engaged in teaching is almost zero. The questions this raises are two fold: How could the situation be improved? and How can academics be persuaded to consider developing their own technology-based teaching and learning materials? One obvious solution would be a revolution in HE - some major change agent that would force academics out of the current delivery paradigm into a new one in which student-learners would be enfranchised and in which academics found teaching fulfilling rather that a chore. Failing the advent of such a revolution how else could academics be encouraged to develop their own materials?

One other possible solution, suggested and evaluated in this study, is the development of a *very* easy-to-use authoring program. If academics had access to such a facility then by a process of 'permeance', the development of materials would gather momentum until interactive learning materials would be the norm rather than the exception. The main experimental hypothesis of this study is that an authoring program based on the use of diagrams as the design metaphor would be easier-to-use and hence superior to other authoring programs.

In evaluating such a *very* easy-to-use authoring program, there is a requirement to evaluate other pre-existing authoring systems in order to ascertain what is good and

bad about them. This study therefore seeks to evaluate existing authoring programs to determine what characteristics a very easy-to-use authoring program should possess. It also seems necessary, when evaluating authoring programs, that cognitive issues surrounding the interaction between user the program and its users need to be examined so that a properly-informed assessment of the needs of a *very* easy-to-use authoring program can be made.

This study was initiated as a part-time research project in January 1994 and has resulted in the publication of a number of research papers and the production of a number of computer programs. The papers derived from this study are listed below and reproduced in full in appendix A.

Elliott, G. J. Jones, E. and Barker, P. G. (1995) Making sense: A review of hypermedia in HE. *Proceedings of EDMEDIA' 95-World Conference on Educational Multimedia and Hypermedia*, Graz, Austria.

Elliott, G. J. Jones, E. and Barker, P. G. (1995) The Application of Concept Mapping to Hypermedia Authoring. *Proceedings of AETT' 95*, Plymouth, United Kingdom.

Elliott, G. J., Jones, Eleri and Barker, P. (1996) Authoring Semantic Hypermedia: A Concept Mapping Approach. *Proceedings of ED-MEDIA' 96,* Boston USA.

Elliott, G. J., Jones, Eleri and Barker, P. (1997) Development of the Computer Skills Metric for Windows. *Proceedings of ED-MEDIA' 97*. Calgary, Canada.
Elliott, G. J., Jones, Eleri and Barker, P. (1997) Which Metaphor for Hypermedia Authoring? *Proceedings of ED-MEDIA' 97*. Calgary, Canada.

Elliott, G. J., Jones, Eleri and Barker, P. (1997). Supporting the Paradigm Shift: Hypermedia Construction with Concept Maps - The Easy Way Forward. *Innovations in Education and Training International,* 34, 4, 294-298.

In addition to the papers that were written, there were a number of other research presentations and activities which also contributed to the contents of this study. These have included the presentation of work at:

- Symposium on Computer-Supported Teaching, Teesside University, December, 1994.

- ITC Palais Du Congres, Paris, 1995.

- Computer-Based Learning in Science, Opava, Czech Republic, June, 1995.

- Association of Learning Technology Conference, Milton Keynes, September, 1995.

- ToolBook' 96 Conference, Wolverhampton, October, 1996.

In addition the author was a participant at the NATO 'Supporting Learning in Computer Environments' Advanced Studies Institute, Heriot Watt University, August 1995.

The description of the study begins in chapter two, a review of the literature to assess the state of the art with respect to all the issues that need to addressed. The outcome of this work is a set of objectives for the study. Chapter three then discusses the methodological choices that need to be made in order to achieve the objectives of the study. In chapter four, the design of an authoring program, utilising a diagram-based design metaphor (*SHAPE*©), is considered and its specification is outlined. Chapter five presents the results of the analysis of the quantitative data collected from the evaluation of *SHAPE*© in comparison with another pre-existing authoring program. Chapter six presents the results of the analysis of the qualitative data collected from

the same evaluation, largely focusing on the cognitive level issues of authoring. Chapter seven provides an analysis of the main types of authoring program used to author hypermedia from three perspectives: first, a desk-based analysis of the functional provisions of existing authoring programs; second, an analysis of the experiences of developing two real applications; and third, the analysis of a focus group of 'experts' considering examples of the main types of authoring program. Chapter eight, the discussion, attempts to synthesis all the results and reconsiders the objectives of the study, paramount in this is the specification of a *very* easy-to-use authoring program. Chapter nine concludes the account of the study.

Chapter Two

Literature Review

2. Literature Review

This chapter considers the issues that are important in the authoring hypermedia learning materials and use of hypermedia in HE. The chapter begins by considering the state of the art of hypermedia, then hypermedia in HE and then focuses in on the problems of hypermedia authoring. Human-Computer Interaction (HCI) is an important factor in the use of hypermedia and is considered next. Throughout the chapter a number of key problems and issues are identified which lead to the establishment of four objectives for this study which conclude the chapter.

2.1 Hypermedia

Interest in hypermedia has grown steadily since Nelson (1967) originally defined the term hypertext. However, perhaps the interest in this area is now 'plateauing' except in the specific context of the Web. Interest in hypermedia can be found in many academic disciplines and much literature has been generated including journals and conferences dedicated to hypermedia, e.g. the Journal of Educational Multimedia and Hypermedia, the annual ACM 'Hypertext' conference, and the New Review of Hypermedia and Multimedia journal.

Hypermedia has been defined by many including Nelson (1980) (modifying his original 1967 definition of the term 'hypertext'), Nielson (1990) and Barker (1993) and there are many ambitious extensions to these basic definitions. For the purposes of this study hypermedia is defined as

The interlinking of computer-based media units/nodes using hyperlinks anchored using hot words, hot regions, hot media etc. which allow users of the hypermedia to navigate around the units/nodes of media.

Computer-based media includes text, sound, video, graphics, movement, pictures, diagrams, i.e. any media which humans use to convey meaning which can be communicated via a computer screen. A distinction is drawn between hypermedia and hypertext, hypertext being the computer-based interlinking of primarily text-based media and hypermedia being the computer-based interlinking of any media including text. This distinction has become rather blurred in the literature and often writers use the term hypertext when hypermedia would be more correct.

The potential applications of hypermedia in education, and in particular HE have been much discussed (e.g. Oliveira, 1992; Lennon and Maurer, 1994; Linn 1992). Hypermedia can be used as a cognitive tool, allowing students to explore and make sense of a knowledge corpus 'constructing' their own meaning in a self-motivated and self-directed fashion and hence developing their metacognitive skills (Jonassen 1992a; Spiro *et al.*, 1991). For many, the use of advanced information technologies in HE is seen as inevitable (MacFarlane, 1990). For governments, hypermedia and other educational technologies can underpin the strategic development of open and distance learning and can support increased efficiency of educational delivery. This is evidenced in the UK by the £33.5 million so far invested in initiatives like TLTP and CTI. (Darby, 1993a; 1993b). The potential impact of hypermedia has even been likened to the impact of the Gutenburg press (Thimbleby, 1992). Despite this enthusiasm, there still seems to be very few HE teaching staff developing hypermedia

applications and the time taken to develop applications is often cited as one of the main reasons for the lack of activity (Hammond *et al.*, 1992).

The phenomenal growth of interest in the Web has resulted in more development but currently this tends to be primarily for hypertext and is of limited functionality because of the limitations of design principles of the Web. There is considerable effort currently taking place to improve the functionality of hypertext markup language (HTML) documents in Web-based hypermedia as described in section 2.1.2 below, but increased functionality requires an ability to write program code like Java and even then the functionality is poor compared to other non-HTML applications.

2.1.1 Hypermedia Taxonomies

In order to place this study in context it was seen pertinent to provide a taxonomy of hypermedia. Halasz (1991) in his closing address to Hypertext 91, the Annual Conference of the ACM on Hypertext/Hypermedia, revised his 3 defining dimensions of hypertext (Halasz, 1988) (see table 2-1) by adding four more (see table 2-2). The new dimensions are classified in terms of the preferences of developers rather than the hypermedia itself.

Table 2-1 Halasz's original defining dimensions of hypermedia

Scope	This really means the size of the hypermedia, i.e. is it designed just for individual use or is it for public access or somewhere in between?
Static versus Dynamic	Is the hypermedia designed for information presentation only with fairly 'static links or is it designed to be amended and supplemented constantly?
Task-specificity	Is the hypermedia designed for a specific task or is it designed to support a number of uses?

Table 2-2 Halasz's additional defining dimensions of hypermedia.

Navigation versus Structure	Navigation-orientated hypermedia places the emphasis on supporting users as they navigate from node to node so much effort is placed on the design of the search engines. The emphasis in structure-oriented hypermedia is on the definition of the structure in terms of the links sometimes at the expense of the node contents. The definition of the structure has become associated with graphical representations.
Literal versus Virtual	In literal hypermedia the links explicitly represent the underlying knowledge. In virtual hypermedia the structure is implicit in the form and content of the nodes.
Node versus Document	Some hypermedia systems are oriented around nodes of specific size and form, others are oriented around the linking together of whole documents and one navigates from one whole document to another.
Composition versus Information Presentation	Some hypermedia are perceived as extensions to conventional documents and their construction is a form of composition and some, the other hand, are perceived as information presentation systems.

It is argued here that 'degree of intelligence' should be added as one further dimension after Rada's idea of combining artificial intelligence (AI) and hypertext into 'Expertext' (Rada, 1991). Basic hypermedia has no intelligence and the onus for navigational choices are on the user, apart from any guidance embedded in the content of the hypermedia. At the other extreme, navigation decisions are made by the hypermedia through the use of embedded intelligence or sophisticated algorithms. The above dimensions represent the extraordinary diversity of what constitutes hypermedia.

Hutchings *et al.*,(1992) have also classified hypermedia in terms of its educational efficacy. Their dimensions are Engagement: (Passive to Active); Control (Teacher/System to Learner); Synthesis (Presentation to Creation) - as shown in figure 2-1 below. They argue that hypermedia should move from hypermedia which is passive and presentation-based to hypermedia which actively engages users and which enables them to become involved in its construction through problem solving type activities. Section 2.1.2 below lists the features necessary for educationally-effective hypermedia.
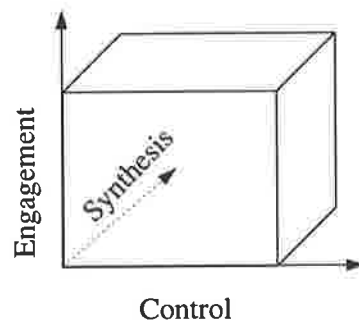
Figure 2-1 An educational taxonomy of hypermedia (Hutchings *et al.*, 1992).

2.1.2 Current Functionality of Hypermedia Applications for HE

The characteristics of a basic hypermedia system are well-documented and there is little to be gained by repeating them here. Jonassen (1989) gives a good overview of the important characteristics of nodes, links and the concept of networks and paths. Halasz (1988) expressed his ideas on what hypermedia systems should contain following his experiences of NoteCards - which are concomitant with Park's (1991) requisite list of features necessary for education and flexibility. Several of these features are already included as standard with development tools, such as ToolBook and KnowledgePro - as shown in table 2-3. What is interesting is the lack of distinction between author and developer-oriented functions and functions oriented towards users of the resultant application.

Table 2-3 Requisite features of hypermedia for education and flexibility

Feature	Example
Interface capability with hardware	Use of Videodisc for motion sequences in applications.
Interface with a high - level language	Use of the HyperTalk language with HyperCard.
Change of Window size and location.	Developers can create applications with windows that can be altered in size, location, using ToolBook or KnowledgePro.
Opening multiple windows	Developers can create applications with many overlaying windows using development tools such as ToolBook or KnowledgePro.

Other features are still subject to research and development as shown in table 2-4.

Table 2-4 Requisite features of hypermedia in experimental phase.

Feature	Research Example
Guidance for Node Selection	Tomek and Maurer (1992) and Kibby and Mayes (1993) both describe systems that use algorithms to help users select the most appropriate link.
Addition of a Browser or Map	HyperCard comes with a summary window facility that enables users to see which cards have been seen. Notecards (Halasz, 1988) used a concept map-based overview map. StrathTutor (Kibby and Mayes, 1993) has a back-track facility so that users can see where they have been.
Node selection by keyword search	The principle of separating the content from the links is becoming increasingly accepted, (Davis *et al.*, 1993; Stubenrauch, Kappe and Andrews; 1993; Mulhauser, 1992). This separation avoids the danger of links becoming outdated and gives the user opportunity to view all the possible links and undertake keyword searches.
Automatic generation of new versions	No evidence of versioning was found.

Another requirement would be a tailorable function to log and audit the nodes visited by users. This plea is made by Horney (1993) that the inclusion of a user navigation pattern logger would allow researchers to better understand how users use a hypermedia corpus and hence improve design. Misanchuk and Schweir (1992) express the view that audit trails could be used for formative evaluation to allow teachers to match material visited or not visited and test results. Kelly (1993) has used a program called SKEIM to monitor student usage of a hypermedia corpus and hence provide useful tutorial feedback. The audit facility should enable students or

teachers to define what needs auditing and how that data is formatted following hypermedia usage, e.g. the facility would allow teachers to select a table or a graphical representation showing the nodes each student uses.

2.1.3 The Problems and Issues of Hypermedia

Halasz's (1991) revised issues that he first set out in his influential 'Reflections on Notecards' paper (1988) categorise the issues under two headings - technology-related and market-related, as summarised in the tables 2-5 and 2-6 below.

Table 2-5 Technology Issues of hypermedia

1. Tyranny of the Link	Halasz argues against the link paradigm and for hypermedia that is based on navigation mechanisms that do not rely on the explicit hyperlink.
2. Open Systems	Halasz argues for increasing the transparency with which hypermedia applications communicate with each other.
3. User Interface	With the size of hypermedia applications increasing almost exponentially facilitating users in 'negotiating' the information space becomes an important issue.
4. Very Large Hypermedia	An associated problem is managing very large hypermedia applications with many different document types and the maintenance of very large databases and sophisticated database engines.
5. Tailorability	This issue has been partially resolved although current tailorability relies on the use of a programming language such as in Hypercard's Hypertalk and ToolBook's Openscript.
6. Collaborative Authoring	Freeman and Ryan (1997) have investigated using concept maps in the collaborative development of hypermedia
7. Computation	

Table 2-6 Market Issues of hypermedia

1. Defining Markets for Hypermedia	There is a mismatch between the enthusiasm of hypermedia developers and the level of take up of hypermedia.
2. Hypermedia Standards	A necessary requisite for any hypermedia development if take up of hypermedia is to increase.
3. Publishing Hypermedia	In the same way that books are published to facilitate take up and distribution there is a growing need to publish hypermedia to make it more available and hence increase the market.

2.2 The Role of Hypermedia in HE

This section considers how hypermedia is used and also how it could be used beneficially in HE.

2.2.1 Educational Benefits of Hypermedia

Ross (1993) draws an interesting parallel between hypermedia and the development of Bloom's higher cognitive skills of analysis, synthesis and evaluation indicating the possible educational importance of hypermedia at each stage of a student's development. Jonassen (1990) also sees the potential of hypermedia to support knowledge acquisition, that is, in expanding a learner's semantic network. In the survey by Laurillard *et al.,* (1993), it was reported that academics perceive that computer-based learning, which includes hypermedia, would allow students to progress at their own pace. Hypermedia offers new ways to learn through the juxtaposition of text, animation and sound and in time it will offer the potential to alter the role of teachers and learners, creating a new dynamic form of interactive

learning for both teachers and students alike (Marchionini, 1988). Laurillard (1993) adds that hypermedia needs additional support from teachers before the benefits of using it become apparent.

2.2.2 Learning with Hypermedia

In terms of computer-based learning, hypermedia is perceived as the vehicle that can offer learners complete control over the viewing of material (Misanchuk and Schweir, 1992). However, much research has shown that only skilled learners benefit from complete learner-control (Steinberg, 1988; McGrath, 1992). Zhao *et al.,*(1993) emphasise the balance needed between complete learner-control and the avoidance of cognitive overload. Most researchers (e.g. Landow, 1990; Whalley, 1990; Beltran, 1993; Laurillard, 1993) agree that some level of direction may be necessary in the use of hypermedia in education if it is to be an effective tool.

The 'direction' in hypermedia comes from the addition of instructional or pedagogical elements to hypermedia. At least four approaches have been tried:

- The incorporation of an intelligent tutor or intelligence within the hypermedia corpus. The Star Guide system (Bruillard and Weidenfeld, 1990) includes an intelligent tutor, whereas the StrathTutor hypermedia system (Kibby and Mayes, 1993) helps direct learners through a hypermedia knowledge corpus by calculating which node is probably the most appropriate to go to next - based on where a learner has already visited and how.

- Creation of hybrid hypermedia systems which contain interactive sequences, (e.g. Beltran, 1993) talks of a hybrid hypermedia model that contains directive sequences.

- Offering learners varying degrees of restrictions through the corpus - depending on their level of understanding of the subject and the medium. The question of who should decide when the level of restriction should be varied is an interesting one.

- Making teachers responsible for giving direction, i.e. placing the use of hypermedia within teaching learning processes, supplementing it rather than supplanting it. Duffy and Knuth (1990) talk of the need for setting 'goals' or authentic tasks in hypermedia interactions, i.e. what assignments need to be set for students to provide an adequate basis for them to engage successfully with the hypermedia system?

Constructivism

The interest and activity in educational technology has been accompanied by an increased awareness in the educational establishment of the advances in the understanding of learning and learning processes, despite the fact that much of the work on learning is quite old, e.g. the work of educational psychologists such as Vygotskii (1962), Piaget (1970) and Gagne (1977). Educational technology is, by its nature, usually amenable to giving learners more control over the pace and topic of learning, concomitant with the ideas expressed in the current understanding of

learning and learning processes. Chaloupka and Koppi (1998) identify that learning materials cannot be functionally isolated from an overall teaching and learning strategy and acknowledge that 'a teacher's locus of control is as important as learner heuristics in the design, development and deployment of hypermedia learning materials. Constructivism which derives from the work of the educational psychologists above, is rooted in the philosophies of Hume, Kant and Wittgenstein. It starts from the principle that deep learning (only) takes place when carrying out 'authentic' tasks (Von Glasersfeld, 1984). If constructivist theories about learning and learning processes are generally accepted, then educational technology, including hypermedia, has a significant contribution to make in supporting learners at whatever level. Constructivism as a theory of learning is contested by theorists who derive their ideas from the behaviourist tradition, for example, Merrill and his ID2 framework (1991). Whatever the theory, most researchers in education would agree that hypermedia has the potential to make a significant contribution to supporting learners at whatever level and whatever they are learning.

2.2.3 Hypermedia as a Cognitive Tool

Nelson and Palumbo (1992) distinguish three different types of use for hypermedia:

- knowledge presentation;

- knowledge construction; and

- knowledge representation.

Arguably, a good presentation system should be a more explicit representation of the underlying knowledge. What is interesting is the use of hypermedia for knowledge construction. Beeman *et al.*(1987) reporting their experiences of 'Intermedia' (a campus-wide hypermedia), identify that the people who learnt most were those most involved in constructing the course material. Reader and Hammond (1994) have also demonstrated that student's post-test scores were enhanced through the use of concept mapping tools alongside hypertext and they argue that students should be encouraged to use cognitive tools to structure their thoughts. There was clear agreement at the NATO Advanced Studies Institute Conference on Cognitive Tools for Learning (Jonassen, 1992a) that hypermedia can be used as a cognitive tool. Reynold and Danserau's (1990) knowledge hypermaps are based on the idea that a hypermedia corpus is a semantic net and hence displays the corpus as a net on the screen.

2.2.4 Resource-Based Learning

A direct consequence of the adoption of the constructivist stance to learning as described above is the need to provide learners with appropriate resources. Learners must be allowed to 'gather' and use whatever materials/tools/artefacts they see as useful to the task in hand. In the past this has been referred to as resource-based learning (RBL) (Clarke, 1982). Hypermedia lends itself to the role of 'resource-provider' and this is particularly true of the Web. There are many references to the use of hypermedia as a resource base, (Spiro *et al.*, 1991). Second generation hypermedia (that is hypermedia with significantly-enhanced functionality) allows not only the production of 'content' but also the creation and use of virtual tools and artefacts (Maurer, 1996). RBL has also been endorsed with the changing role of

libraries (Follet Report, 1996) and their strategic importance is emphasised in the Dearing Report (Dearing, 1997).

2.2.5 Role of Teachers and Students

Very early on in Laurillard's (1993) influential book on the role of education technology in HE she discards the traditional role of teachers as 'imparters of knowledge' in favour of teachers as 'mediators of learning' by drawing on the work of the same pioneering educational psychologists that the constructivists have drawn upon. It is the job of teachers to provide authentic tasks and situations within which learners engage with the ideas, phenomena, etc. which they have to learn. Teachers therefore have a responsibility to create these tasks and situations and, if hypermedia is to be employed, it is the teachers who must create or author the appropriate hypermedia. However, as noted above, Beeman *et al.* (1987) found that the people who learnt most when utilising hypermedia were most involved in constructing the course material, an ironic reinforcement of the need for authentic tasks when concerned with deep learning.

The question of who should be the author of hypermedia in a HE setting is now an academic one, the answer is that both students and teachers should be authors. In terms of passive hypermedia, using Hutchings *et al*'s (1992) taxonomy as described above, Elliott *et al.* (1995a) suggest that teachers may be the initial author of a passive, domain-specific hypermedia but that learners should be able to supplement, superimpose or simply author their own hypermedia based upon their understanding of the domain under study. In terms of authoring hypermedia that is imbued with

active engagement and supports authentic tasks the responsibility falls firmly on the shoulders of teachers, as mediators of learning.

2.2.6 The Problems and Issues of Hypermedia in HE

Despite the recognition of the role that hypermedia could play alongside other educational technologies in HE, there has not been any great rush towards its development and use. In one survey, Rode and Poirot (1989) found that 65% of even computer-literate staff at the University of North Texas were not disposed to writing educational software, although Rode and Poirot did not examine the reasons. In the surveys reported by Hammond *et al.* (1992) it was noted that the main reason why academics did not use new educational technologies (including hypermedia) in teaching-learning processes seems to be lack of time to develop, followed by lack of training and lack of support staff. These findings are corroborated by findings by Laurillard *et al.* (1993) and Barker and Bannerji (1994). The other major factor noted is the need for educational institutions to become more supportive of staff developing and experimenting with educational technologies such as hypermedia and new teaching strategies. (Laurillard, 1993). The 'IT Assisted Teaching and Learning in HE' report (HEFCE, 1997) commissioned by the HE Funding Council for England noted that academic staff time is increasingly prioritised on research output and not on teaching, hence reducing further the time available for the development of learning materials.

Time

Davis *et al.* (1993 quoting Christie, (1990) estimate that it takes between 100 and 150 hours to produce one hour of hypermedia instruction even for experienced developers. An experienced lecturer preparing a one hour lecture session could produce the requisite material in less than 10 hours. Whilst in the short term a cost-benefit analysis shows traditional approaches to have the edge, in the longer term and in the context of open and distance learning, the balance should change as hypermedia shows the potential for a major increase in flexibility of delivery and erosion of the temporal and geographical constraints which dog traditional teacher-centred delivery.

Computer Skill

Elliott *et al.* (1997) found that there is a wide variation in the computer skill level of academic teaching staff in HE and that the average skill level is quite low. In order to successfully utilise and author effective hypermedia using current authoring tools requires a relatively high skill-level. Along with the other factors noted above this apparent skills gap accounts for the lack of development activity. In contrast some are more optimistic, e.g. Junkala, (1991) and believe that almost anybody can produce college-level courseware.

2.3 Hypermedia Authoring

In this section the issues of hypermedia authoring are considered. It starts by considering who is the author in sections 2.3.1 to 2.3.3 and then compares of the authoring process to knowledge construction in section 2.3.4. Section 2.3.5 considers

the nature of available authoring programmes. This overall section concludes with a summary of the problems and issues identified.

2.3.1 Authoring Foci

The additional four dimensions to Halasz's taxonomy of hypermedia that were indicated above, mean that a range of different hypermedia authoring programs HAPs[1]) are required. The subtleties expressed in Halasz's taxonomy and the requirements of second generation hypermedia have to be produced from a limited number of available HAPs and, more often than not, programming code.

2.3.2 Learners Involvement in Hypermedia Authoring

Many believe that both students and teachers should contribute to the creation and authoring of campus-wide hypermedia information systems (Landow, 1990; Russell, 1990), which should be continually undergoing growth and change (Stubenrauch, *et al*, 1993). The Microcosm system developed at Southampton University (Davis *et al*, 1993) is an excellent attempt to rationalise and provide a co-ordinating framework for this process. Within these growing information systems there is a place for hypermedia applications, carefully and purposely prepared by lecturers, for inclusion in their own particular teaching curriculum - what Halasz would call 'closed hypermedia' (Halasz, 1991). These generally smaller but discrete hypermedia are comparable to the composite nodes of the Dexter model (Hardman *et al.*, 1994), that

---

[1] HAP is used in this thesis for programs designed for authoring hypermedia, sometimes the word program has been used when reference is made to any program.

is, a subsystem sitting within a wider campus system. The educational benefits of these smaller, discrete hypermedia corpora have been open to debate (Stubenrauch *et al.*, 1993) because of their limits but placed within a larger hypermedia, like the Web, these concerns are quelled.

Closed, discrete hypermedia systems are, in some sense, reflections of teachers' understanding of the subject matter - assuming , of course, that it is teachers who have created them. As discussed above, learners using these hypermedia systems would probably not be expected (or even allowed) to modify the existing corpus but would be expected to construct their own 'subset' or view of the corpus under scrutiny - just as they produce their own written notes today. They may, for instance, use a concept mapping tool to reflect their own understanding of the material, extracting material from their teachers' hypermedia corpus if they think it is important. Learners could have the ability to annotate a subset of links and nodes of their 'subset' appropriate to their own understanding of the knowledge corpus. There is an analogy here between the notes a lecturer provides for students and that lecturer's hypermedia system and between the library books given for reference and a campus-wide hypermedia system. Indeed, Davis *et al.* (1993) suggest that when producing applications in Microcosm, hypermedia material should be also added to the campus-wide corpus as explanatory notes.

2.3.3 Collaborative Authoring

There is a model of authoring that has arisen in response to the considerable effort required to develop hypermedia material which relies on the contribution of a number of academics who are interested in a similar subject domain. Computer-supported collaborative work (CSCW) technology has been utilised to provide CSCW-based authoring (Dobson *et al.*, 1993). The DELTA SAFE (D1014) project (Derks *et al.*, 1991) investigated the use of 'learning material components' which can be reused and modified by different groups. Despite these attempts the 'not invented here' syndrome, with some justification, still prevails (Hammond *et al.*, 1993*)*.

2.3.4 Knowledge, and its Representation and Construction

Philosophers as far back as Plato, have been pondering epistemology in an attempt to understand human thinking. More recently, research in AI has grappled with the need to make concrete exactly what knowledge is. Newell (1982), who was one of the major proponents of 'strong AI' (Searle, 1981) presented a theory on knowledge representation in the human mind. Since then there has been very little progress on the nature of this type of knowledge. Some knowledge types have been specifically identified, e.g. procedural, declarative, structural and even meta-knowledge, but how these knowledge types are 'held' or even whether they are distinguishable in the human mind remains unsubstantiated theory and conjecture. Today AI workers are likely to have the more prosaic title 'Knowledge Engineers', reflecting the recognition by the AI community that more fundamental work on epistemology is needed before progress in AI can continue. Knowledge engineering is concerned with the representation of knowledge and subsequent encoding in some computer-based format

following the elicitation of knowledge from humans and elsewhere. Knowledge engineers name the process of knowledge elicitation, representation, and encoding as knowledge acquisition. Successful knowledge engineering has proven to difficult (Hoffman, 1987), and methodologies proliferate, see for example in the book edited by Diaper (1989) which describes numerous techniques of knowledge elicitation and acquisition. A number of computer-based applications have also been developed to help in the task of knowledge elicitation, e.g. KEATS (Motta, 1986) and Pathfinder (Schvaneveldt, 1990) which is based upon accepted techniques of knowledge elicitation. McAleese (1986) identified a number of problems related to the construction of knowledge in authoring environments:

- The problem of exteriorisation - the process by which the author makes explicit, internal knowledge structures;

- The problem of metacognition, i.e. thinking about internal knowledge structures (a problem also identified by Fisher (1992)); and

- The problem of anomalous state metacognition, thinking about 'gaps' in internal knowledge structures.

For the purposes of this study the following definition of knowledge representation expressed by Bench-Capon (1990) is adopted:

> *'a set of syntactic and semantic conventions that makes it possible to describe things'*.

Representation of Knowledge

The AI community have used a number of paradigms to try to express knowledge: production rules; logic - predicate, propositional and fuzzy; structured objects, frames, schema; and semantic nets. (Bench-Capon, 1990). Bench-Capon (1990) has argued that the various paradigms are fundamentally similar in what they can provide in terms of knowledge representation. In terms of being able to utilise the paradigms for knowledge representation, humans are generally poor at representing knowledge using logic, and also to some extent production rules despite their power and clarity of expression. Structured objects have been used partly to overcome this human difficulty by making the representations more 'cognitively ergonomic'. Recently, interest has grown in the area of diagrammatic representation and reasoning and arguments in favour of diagrammatic representations over logic-based ones have been expressed (Kulpa, 1994). In this study diagrammatic representations and structured objects can be considered to belong to the same class of knowledge representation approach. One enduring epistemological issue is whether human knowledge is essentially procedural or declarative in nature (Winograd, 1976; Diaper 1989) - Winograd believing it to be declarative, and Diaper procedural.

Convergence of Knowledge Representation and Hypermedia

There is a convergence between ideas of knowledge and knowledge representation and hypermedia and hypermedia structure (McAleese, 1987; Lawton and Smith, 1993; Nelson & Palumbo, 1992). Jonassen (1993), Carlson and Slaven (1992) and the

Bench-Capon definition of knowledge representation given above comfortably includes hypermedia. The process of hypermedia authoring is often referred to as knowledge construction (Jonassen and Grabinger, 1990). Hypermedia can be associated with the structured object class of knowledge representation paradigms as noted above. The process of hypermedia construction can be viewed as knowledge construction, however the question of whether hypermedia can represent all necessary knowledge types is an unanswered question. Hypermedia seems to correspond with declarative knowledge rather than procedural, although in Winograd's opinion (1976) procedural knowledge can be represented in a declarative form, e.g. using statements like 'A follows B'. Procedural knowledge in this form can also be represented in hypermedia with appropriate hyperlinks however it remains to be seen whether the representation is useful.

Visualisation

The belief that diagrammatic/graphical/pictorial representations of ideas, knowledge domains, etc., is desirable, is enduring and influences the design of many human-computer interactions, for example, visual programming paradigms. The positive ergonomics of visualisation is also supported by empirical studies in psychology. Buzan (1993) cites experiments where subjects exhibit an extraordinary capacity to remember visual images. Semantic nets used in knowledge representations are another example of where visualisation has been used to mitigate the un-intuitive logic representations of knowledge. The problems associated with the navigation of hypermedia have also been addressed with semantic nets in an attempt to help users 'see' where they are in the hypermedia (McAleese, 1987). Semantic nets have also

been used to construct hypermedia (McAleese, 1990) and even Web-based hypertext/hypermedia (Kremmer, 1997; Freeman and Ryan, 1997). Carswell (1998) suggests that people can be categorised on a visualiser/verbaliser dimension although no experimental evidence is cited. The effectiveness of visualisation/diagram-based HAPs to author hypermedia remains to be evaluated and is one of the objectives of this study.

Visualisation has been very successfully used in visual programming languages, like Microsoft's Visual C and Visual Basic. The visualisation syntax employed in visual programming languages enables developers to construct user interfaces and sometimes the structure and logic of the underlying programs as well.

Epistemology

As noted above, any theories on human knowledge representation are still largely conjecture although from the discussion in the previous paragraph it is safe to say that images seem to have some resonance with our memory. The theory of constructivism (Von Glasersfeld, 1984) referred to above, argues that we do not know how knowledge is represented in the human mind and there is no proof of the existence of any piece of objective knowledge which is identical in everyone's mind. What are the implications of constructivism for HE? Teachers are still responsible for ensuring students have access to a body of knowledge and for evaluating their ability to describe, comment and analyse, etc., that body of knowledge. Many academics probably harbour more than a creeping suspicion that their students are not 'getting' the message when they come to read their work, however well designed the

instruction has been. All that one can do is create an environment in which students can successfully engage with a body of knowledge and then evaluate what they have learned. One may also need to test certain behavioural outcomes, e.g. balancing an equation in chemistry although one still needs to test that they have learned the ideas behind balancing equations. What are the implications of constructivism for hypermedia? The answer is probably that hypermedia needs to be designed so that it gives students the best environment for them to create 'suitable' meanings of the body of knowledge under question and does not result in too many misconceptions.

2.3.5 Hypermedia Authoring Programs

There are two perspectives to authoring hypermedia as noted above; large campus-wide systems to which students and lecturers both create nodes, add and modify links in an ever-expanding corpus; and the self-contained, discrete systems designed for inclusion in a particular lecturer's course material that exist within the campus-wide systems. There are several ways of authoring discrete hypermedia applications but the only way non-programmers (the majority of HE teaching staff) will produce applications is with *very* easy-to-use authoring programs.

There are many HAPs: KnowlegePro©; ToolBook©; HyperCard; LinkWay©, to name but a few - Barker (1993) gives a good overview of the available HAPs. Most of these systems already provide the basic functionality outlined above and those that have links to high-level languages can provide the other features albeit with a high programming overhead. A number of Web-based HAPs are also appearing which can

be used to create either discrete or campus-wide hypermedia, e.g. Microsoft FrontPage©, Geo Emblaze©, Macromedia Dreamweaver©. There are a number of HAPs which are designed particularly for education which have mainly come from research projects, e.g. Hyperwave© (Maurer, 1996), Microcosm© (Davis *et al*, 1993), CourseWorks, (Miller, 1995), W3Lessonware (Siviter, 1997).

Taxonomy of Hypermedia Authoring Programs

There is a distinction between multimedia and hypermedia authoring programs, although in essence most contemporary multimedia authoring programs enable hyperlinks to be added and are therefore also HAPs. HAPs can be categorised by the design metaphor used for creating the hypermedia with the most common metaphors used being: control flow (e.g. Macromedia Authorware©); book (e.g. Apple HyperCard© and Asymetrix ToolBook©); music score, (e.g. Macromedia Director©); directory and file structure (e.g. Microsoft FrontPage©). The purpose behind the various metaphors should be to make the HAP as 'transparent' as possible to the user (Shneiderman, 1998). Some HAPs incorporate a scripting language so that any interactivity or functionality not available through the standard facilities can be programmed in. Some HAPs provide overviews so that the structure of the hypermedia as it evolves can be seen, e.g. Microsoft FrontPage©. The overview facility may be an implicit part of the metaphor, for example in the control flow metaphor one can see the flow of the whole application although when levels of flow are nested they become invisible. Fourth and fifth distinguishing factor are the learnability and usability of each HAP, although this factor is considered in more detail in section 2.4.3 below. Giving a rating to the learnability of any particular HAP

is difficult and a number of other factors come into play such as the experience of the user, and this is one of the issues investigated in this study (see section 2.4.3 below). It is roughly true to say that the Macromedia Director© implementation of the music score metaphor is more difficult than the book metaphor of Asymetrix ToolBook©, although this may not be true for every user and one must also bear in mind the power of expression of each HAP with its learnability. One new factor is the degree to which the HAP is 'Web-compatible', i.e. to what degree can applications be converted to Web-deliverable applications. Table 2-7 below summarises some of the attributes discussed above of the more commonly used HAPs. A more detailed analysis is given in chapter seven.

Table 2-7 Taxonomy of most common hypermedia authoring programs.

HAP	Metaphor	Scripting Language	Overview	Learnability/ Usability[2]	Web compatible
Authorware©	Control Flow	Not really	Implicit in the metaphor	Poor	
Director©	Music Score	Yes	Partially	Very Poor	
ToolBook©	Book	Yes	No	Reasonable	Plug-in for browser but no conversion to HTML
FrontPage©	File and Directory	No	Implicit in the metaphor	Reasonable	Aimed at producing Web applications

Developments in Hypermedia Authoring Programs for Education

Barker (1992) suggests that facilities for authoring hypermedia must be easy to use. In an attempt to make the process of producing hypermedia for education easier and to ensure that applications are instructionally effective, a number of models and systems

---

[2] Based on the subjective view of the author prior to the study. See chapter seven for an extended analysis

for hypermedia authoring for education are appearing, e.g. the Nestor model and system (Mulhauser, 1992) and the Hypercourseware model and system (Siviter and Brown, 1992). Other authoring tools have been enhanced with additional functionality, e.g. the NEAT system (Mayer *et al.*, 1993), is an extension to ToolBook that claims to offer the power of a programming language without the need to program, it also provides a variety of metaphors. Midoro *et al* (1992) report a database of learning material resource, so 'units' of learning material can be reused and pulled together with new material to create courseware. The ISAAC system (McAleese and Ching, 1993) integrates instructional design help with the authoring tool. Developers using ISAAC with an authoring tool like ToolBook, receive advice on best instructional practice during design sessions. ISAAC is distinct from other instructional design advisory systems, such as HyperTactics which works off-line (Jonassen and Harris, 1991).

Future Directions for Hypermedia Authoring Programs

Not withstanding the moves towards development environments based on models of good instructional design theory, authoring environments for education need to be very easy to use. As stated above, the academic community considers time to be the major factor affecting their reluctance to consider developing computer-based learning material. Some believe the answer lies in using material developed elsewhere (Laurillard, 1993) and this does seem a good way to proceed although the 'not invented here' syndrome is a formidable barrier. There would be considerable advantages to exploring strategies that would make hypermedia development very easy. If authoring hypermedia became as easy as producing material conventionally,

e.g. using a wordprocessor to produce student notes and lecture material, then perhaps many more people would try it and changes in HE would follow as a consequence.

Expressiveness of Metaphor

There is a trade-off between the expressiveness of an authoring metaphor and its simplicity of operation, Chapanis (1991) calculates that 'versatility' is negatively correlated to other ease-of-use parameters, i.e. programs that are versatile (i.e. more complex) are not usually easy to use. To create educationally-effective hypermedia, authoring tools needs to allow authors the ability to create hypermedia that has the following identified features (Elliott *et al.,* 1995a):

- multifaceted, multi-perspective views of real world problems/issues;

- active engagement;

- flexible support of a number of different structures;

- good search and query facilities;

- support for different learning styles;

- interactive sequences; and

- appropriate node size.

Support for different learning styles is partially dependent upon how hypermedia is applied in a teaching-learning process or setting. As Laurillard (1993) points out, most current hypermedia applications are nothing more than sophisticated information access systems, i.e. electronic page-turners, although this can be mitigated by the way the hypermedia are deployed within a teaching-learning process. Furthermore, authors

need the ability to reflect sufficiently on the knowledge domain underpinning their completed applications. This means being able to display information chunks in whatever form - sound, text, graphics, etc., with associated navigation mechanisms - the system of hyperlinking. Some form of narrative, story-line or guided discovery mechanism is needed in order to 'make sense' of a hypermedia corpus (Elliott *et al.*, 1995b). It is essential also that users have learning objectives and activities embedded within the hypermedia, i.e. active engagement - in the spirit of constructivism.

2.3.6 The Problems and Issues in Hypermedia Authoring

This section has sought to consider the current state of thinking and practice on hypermedia authoring and a number of problems and issues have emerged. The major problems and issues have been:

- How can sufficiently expressive HAPs be constructed to produce the wide variety of hypermedia outlined in Halasz's revised taxonomy and specifically for HE which are easy to learn and use?

- What are the appropriate knowledge types for representing knowledge for HE teaching programs? Is hypermedia able to represent all of them?

- The need for the evaluation of the effectiveness of visual and diagram-based metaphors for hypermedia authoring for developing teaching applications for HE. What factors about users are important to consider in the evaluation of HAPs?

- What are the best metaphors for hypermedia authoring and should a variety of metaphors be used?

- What should hypermedia, based on a constructivist theory of learning, look like and what should the corresponding HAPs look like?

2.4 Human-Computer Interaction (HCI) and Usability Issues of Hypermedia Authoring

When considering the issues and problems of hypermedia authoring, due attention to the developments and issues in the field of HCI is obviously necessary. The issue considered to be of paramount importance in HCI currently is usability, i.e. designing for usability and evaluating usability (Shackel, 1990). Research in HCI is multi-disciplinary involving cognitive psychologists, psychologists, computer scientists, all focusing on different aspects of HCI from modelling users; modelling users' models of the computer system; examining users' actual use of the computer and innovations in HCI. In this section the main issues and developments of HCI are examined with their relevance to hypermedia authoring and HAPs.

2.4.1 The Cognitive Psychology Approach

Cognitive psychology has become closely associated with HCI research because humans are really the focus (Barnard, 1991). The crucial problem for cognitive psychologists is developing adequate theories of behaviour explaining human use of

computers (Shneiderman, 1998) which resonates with but is distinct from Shackel's belief, noted above, that usability is the key HCI issue. In general high-level theories manifest themselves as design guidelines or heuristics for practitioners which are invariably too general to apply, e.g. 'ensuring the interface is consistent'. Buckingham *et al.* (1994) recognise the disparity between theory and practice and suggest a framework for bridging the gap. The lack of an accepted and adequate theory to describe HCI has resulted in a plethora of competing theories, all manifesting their usefulness in general guidelines.

Mental Models

One consistent theme in HCI theory is the need to describe a user's understanding and behaviour in using the computer system in question and then for successful models to be integrated into the software so that they correspond to each user's actions. Theories generally pay attention to the various levels of the interaction, broadly fitting into Foley *et al*'s (1970) early hierarchical model from a high level conceptual approach down to semantic, syntactic, and lastly the lexical level where users use precise commands or actions. At the lexical level, theories attempt to model users' behaviour by their minute actions, including Card *et al*'s (1980) Goals, Operators, Methods and Selection Rules (GOMS). At a higher level are theories like Task Action Grammar (Payne, 1985) which use a well-defined syntax to describe users' interactions from keystrokes up to the semantics of a task. Shneiderman (1998) argues that the hegemony of graphical user interfaces (GUIs) particularly Microsoft Windows and means that the syntactic level has all but disappeared although it is

argued here that the learning of a basic syntax is still required even with GUIs and this can still be a challenge for some.

The usefulness of the theories mentioned above have been criticised - Briggs (1988) for example argues that the 'reductionist view' of HCI idealises it and takes no account of the real behaviour of users with their own foibles and the particular environments in which they work.

2.4.2 The User Interface

Interface design has not advanced significantly since the revolutionary GUI of the Apple Macintosh (1984) with only incremental improvements in particular aspects of GUIs since then. There is considerable work in progress to develop the next generation of interface. In a special issue of the Communications of the Association of Computing Machinery on the future of GUIs, Robertson *et al.* (1993) investigated three-dimensional imagery as a metaphor for the interface, Morse *et al.* (1993) considered the problem of users coping with screen complexity with many objects and windows and other objects interlinked and embedded. In the same issue Nielson (1993a) considered how next generation interfaces will avoid the command orientation of current direct manipulation GUIs, a legacy of command-driven interfaces of the 1970s. It is clear that whatever research is taking place currently, a visual/graphical interface will probably continue for the next 5-10 years, even if voice control is perfected.

2.4.3 Usability and Usability Engineering

As noted in the introduction to this section, usability is an issue of paramount importance. Nielson (1993b) presents a model of system acceptability which encompasses usability as shown in figure 2-2 below.
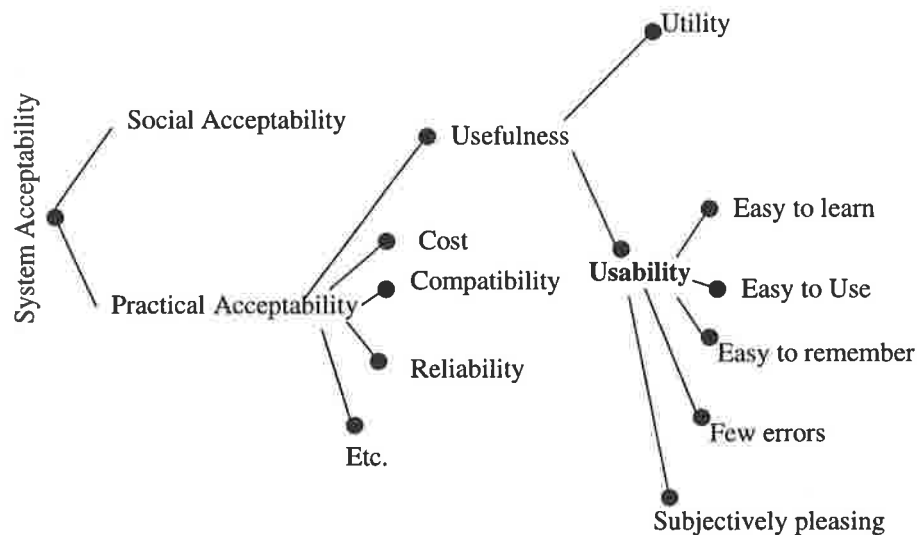
Figure 2-2 Nielson's model of system acceptability (1993).

The model shows that usability is an interplay of a number of factors. Ease of learning and ease of use are generally-accepted terms, other terms like efficiency and memorability are more contestable with many more competing and overlapping concepts including guessability (Jordan *et al.,* 1991), visibility (Gilmore, 1991), acceptability (Nielson, 1993b; Gilmore, 1991, Davis *et al.*, 1989), transparency (Maass, 1983) utility and effectiveness (Sutcliffe, 1995) and task match (Eason, 1984). There are other factors used in usability studies in the literature including Green's (1989) cognitive dimensions which are described below. Reliable and predictive usability evaluation is the recurring focus of research with many methods

and techniques proposed. The main problems with usability evaluation methods is the subjectivity of the evaluation and the interplay of so many factors, including the characteristics of the users, the environment, the sample size of the user group - hence the problem of the isolation of the factors under examination. As Nielson (1993b) points out, the holy grail for usability engineers is the invention of analytical methods that would allow designers to predict the usability of an application before it is has even been tested. This goal is concomitant with cognitive psychologists' goal of a predictive model of users and their interactions with computers. It is telling that when Microsoft was developing Windows 95 the design process avoided all formal usability methods in favour of a highly iterative process with 560 subjects using the prototypes (Sullivan, 1996). What was also interesting about the development of Windows95 was that there were only three main objectives: it should be easy to learn; easy to use; and be completed in 18 months.

The details of the various methods of usability evaluation and metrics are considered in the methodology, chapter three.

Learning and learnability of computer programs.

The process of learning and the learnability of computer programs has been well studied (e.g. Napier, 1989; Allwood, 1990; Wilson *et al.,* 1990), although there does not seem to be a generally accepted model of learning of computer programs in the literature. Clearly, there is an enormous body of research dedicated to human learning but its contextualisation in the learning of computer programs seems lacking at this

stage with Nielson's model of attributes of system acceptability being arguably closest (figure 2-2).

Green's Cognitive Dimensions

In section 2.4.1 above it was stated that there is a plethora of theories to describe HCI although the emphasis in the past has been on low-level reductionist views. Briggs (1988) has argued against such low-level reductionist approaches in favour of more holistic qualitative approaches. Green (1989) has presented his cognitive dimensions approach which responds to a more holistic analysis of usability. Green's dimensions represent the various pertinent attributes of any 'cognitive artefact' that one would wish to analyse, including HCI artefacts. The dimensions are used as prompts for the evaluation of the cognitive artefact and the relative importance of each dimension will depend on the nature of the artefact. Table 2-8 below is a summary of the dimension; reference to 'cognitive artefact' in the descriptions can be replaced by HAP for this particular study. Green's dimensions seem to be a more pragmatic way forward for usability analysis than the reductionist approach of GOMs etc. Interestingly, Green and Petre (1996) have applied cognitive dimensions to the usability of visual programming environments which indicates the current importance placed on visual interfaces.

Table 2-8 Summary of Green's Cognitive Dimensions

Dimension	Description
Abstraction gradient	What are the minimum and maximum levels of abstraction? Can fragments be encapsulated?
Closeness of mapping	What particular usage strategies implicit in the artefact need to be learned?
Consistency	When some of the cognitive artefacts have been learnt how much of the rest can be inferred?
Diffuseness	How many symbols or graphic entities are required to express a meaning?
Error-proneness	Does the design of the notation induce 'careless mistakes'? Are there places where users need to resort to fingers or pencilled annotation to keep track of what has happened?
Hidden dependencies	Is every dependency overtly indicated in both directions? Is the indication perceptual or only symbolic?
Premature Commitment	Do programmers have to make decisions before they have the information they need?
Progressive evaluation	Can partially complete elements of the artefact be instantiated to obtain feedback on 'How am I doing'?
Role-expressiveness	Can readers see how each component of a program relates to the whole?
Secondary notation	Can programmers use layout, colour, other cues to convey extra meaning above and beyond the official semantics of the language?
Viscosity	How much effort is required to perform a single change?
Visibility	Is every part of the artefact simultaneously visible (assuming a large enough display), or is it at least possible to juxtapose any two parts side-by-side at will? If the code is dispersed, is it at least possible to know in what order to read it?

Cognitive and Other factors of the User

There are a number of other factors that may need to be considered in the usability analysis of HAPs. For instance, the prior skill and experiences of users, especially their computing and prior exposure to similar paradigms might affect their perceptions of a HAP. Perhaps their spatial or diagrammatic reasoning ability might have an effect on their ability to use visual/diagram-based metaphors. Briggs (1988) has argued that environmental factors like the organisation or office in which a computer-based application is to be used are an implicit part of the dynamics of HCI and any analysis that avoids or ignores them is lacking in validity

2.4.4 HCI and Hypermedia

There has been a considerable effort in recent years in terms of evaluating the usability of hypermedia from a user's perspective. Factors that have been debated and analysed at length are the interface metaphor (Mukherjea & Foley, 1995), user/learner control (e.g. Steinberg, 1988; McGrath, 1992; Zhao *et al*, 1993) and navigation where there is a particular concern with the problem disorientation and becoming lost in a hypermedia corpus (Nielson, 1990b; Smith, 1994; Edwards and Hardman, 1993). The interest in the usability of hypermedia has grown in proportion to growth in usage of the Web - see for instance, the special issue of the International Journal of Human Computer Studies (1997) on the usability of the Web. Despite Nielson's (1990c) comment on the lack of analysis of the usability of HAPs no work on the usability of HAPs could be found in the literature. The usability of HAPs raises the questions, what is the best or most appropriate design metaphor for HAPs for particular applications?

2.4.5 The Problems and Issues in HCI

From the discussion above the following problems and issues have been identified:

- The need to develop an accepted, complete and embracing theory of HCI which can predict the usability of an application that:

    i. Determines the right approach for usability analysis, lower level reductionist, high level holistic or a mix.

    ii. Determine the next generation of the user interface.

    iii. Includes a complete model of human learning of computer-based artefacts.

    iv. Identifies what factors need to be taken into account when measuring and designing for usability.

    v. Provides the basis for effective usability engineering methods that will be reliable, valid (in the context of an accepted theory) and can be used for predictive purposes.

- The need to develop a framework for evaluating the usability of HAPs in lieu of the resolution of the issues identified.

- The need to identify the best or right design metaphors for HAPs.

2.5 Summary of the Problems and Issues

The main problems confronting hypermedia is flexibility. How can the various types of hypermedia be constructed from the smallest set of tools with the greatest ease? How can the current inertia in developing and using hypermedia be overcome? The main problems concerning the role of hypermedia in HE are a subset of these main concerns. How can hypermedia be constructed easily with tools that have minimal learning curves so that motivation is maintained? How can hypermedia be authored so that it is educationally-effective? What can be done about the structural problems of HE that mitigate against the take up of educational technologies like hypermedia?

The educational effectiveness of hypermedia raises a number of other problems. What should HAPs, designed for producing flexible learning materials, that are grounded in good teaching and learning theory, look like? How can HAPs be designed so that they can represent all of the requisite knowledge types for HE? What are the most appropriate design metaphors for hypermedia authoring? When designing and evaluating HAPs, the HCI issues of learnability and usability need to be considered. How should HAPs be designed so that they respond to the best theories of HCI? How should HAPs be evaluated and what factors are important?

2.6 Tried Solutions

There is on-going research and development to respond to the issues and problems noted above. In the HE context of the United Kingdom (UK), this has been accentuated by the conclusions of the Follett report (1996) on libraries in HE and the Dearing report (1997) on the structure and funding of HE together with the pressures from central government to squeeze funding at the same time as insisting on increased participation rates. These pressures, which are no doubt shared by the HE systems of other countries, have resulted in the specific UK funding of research to respond to some of the issues raised above. In particular, the work funded by the successive stages of the Teaching and Learning Technology Programme (Darby, 1993b, Darby, 1993c) have addressed a number of the issues and resulted in some recommended solutions. The solutions recommended seem to be falling into a framework that has the following components:

Shared Development: teams of academics involved in generic subjects like basic statistics will collaborate on the production of generic educational technology materials. A number of models for collaborative authoring have been developed e.g. Dobson *et al.* (1993), to respond to this trend. The shared development model mitigates the problems of time and effort although the final product may not be quite what every academic wants so the end product must still be tailorable.

Shopping Trolley: the HE arena for educational technology is now world-wide particularly in the light of the Web. Academics will be

increasingly encouraged to use learning materials developed from all over the world.

Templates: a number of efforts have been made to produce customisable templates, e.g. Neat (Mayer *et al.,* 1993), Nestor (Mulhauser; 1992) and Hypercourseware (Siviter & Brown; 1992). These tools or environments are useful but in attempting to make the authoring process easier they tend to be prescriptive in the nature of materials produced.

2.7 Another Solution

It is beyond the scope of this study to respond to all the issues related to hypermedia and hypermedia authoring discussed above. This study focuses on the usability of HAPs in HE. This study takes account of the solutions suggested above but takes the stance that it may be possible to develop a HAP that is very easy to learn and use but is also capable of creating educationally-effective hypermedia. The underlying philosophy promoted is that of *'permeance'*, i.e. a HAP that is very easy to use will gently permeate HE without necessarily requiring any structural changes. This stance is recognised as idealistic but together with some structural changes in HE the momentum for the provision of educationally-effective learning environments in a system with large student numbers will gather. Microsoft's PowerPoint is a good example of a well-designed program that has permeated HE and is beginning to affect educational delivery.

The proposal suggested here is that a HAP based on a diagram metaphor could be easy to use. A diagram-based metaphor that academic staff can easily understand or with which they are already familiar with would reduce the cognitive load associated with learning the HAP. A diagram-based metaphor that has minimal constraints in its syntax will make it easier for academic staff to manipulate.

Many diagramming techniques have been developed in the last half of this century to help humans understand and represent (sometimes) complex ideas. The use of diagrams has a long pedigree going back to the geometry of Euclid and Pythogoras. In recent years a range of new diagramming techniques have arisen which utilise minimal syntax so that authors can focus on the thing they are trying to represent. All the techniques are fundamentally similar in that they represent the relationships between ideas, albeit some more explicitly than others and many emphasising the desirability of imagery for improving the interpretation and recall of the diagrams. Possible candidates include Mind Maps (Buzan, 1993), Fish Bone, Cause and Effect or Ishikawa diagrams (Ishikawa, 1991), Rich Pictures (Checkland, 1981) and Concept Maps, (Novak and Gowin, 1984).

'Mind Maps' (Buzan, 1993) are an aesthetically pleasing diagramming technique for which Buzan makes some rather impressive claims. The semantic relationships in *Mind Maps* are implicit. *'Fish Bone'*, *'Cause and Effect'* or *'Ishikawa diagrams'* (Ishikawa, 1991) are mainly used in commerce and industry to help people solve business problems usually related to quality. Fish Bone diagrams are essentially the same as *Mind Maps* with semantic relationships represented implicitly, although the emphasis is on determining the hierarchy of causes to an effect or a problem. *'Rich*

Pictures' (Checkland, 1981) are used in computer systems development to enable non-specialists to represent the complexities of business processes. Rich Pictures can contain explicit semantic information however this is at the discretion of the author since they have no pre-defined syntax. Concept Maps (Novak & Gowin, 1984). emphasise the semantics between ideas and although not the most aesthetically pleasing of the diagramming techniques, they contain much richer information.

Jonassen (1990), who has considered the development of associative networks which includes concept maps, notes their similarity to hypermedia because of the explicit representation of the semantics within a particular domain. Concept maps are at one level a very simple construct with very little syntactic constraint. If a HAP was constructed using the concept map metaphor it might be very easy to use. Following their study of metaphors in computer systems, Anderson *et al.* (1994) conclude that metaphors must be closely coupled to the purpose of computer systems and must have minimal *conceptual baggage*. Of the techniques discussed above concept maps seems to be most closely coupled to hypermedia and have less *conceptual baggage* than many techniques. The section below discusses the implications of using concept maps to author hypermedia.

2.7.1 Concept Maps

Concept maps have been used extensively in education as a means of expressing knowledge domains (Okebukola 1992) and evaluating student misconceptions (Ross and Munby 1991). They have also been used for curriculum development (Barenholz and Tamir 1992), as learning tools like the CLASS tool (McAleese, 1994) and even as a form of assessment (Beyerbach and Smith 1990). In the more syntactically-constrained form of semantic nets (Quillian, 1968)[3] concept maps are used as 'front ends' to hypermedia knowledge corpora, (Reynolds & Danserau 1990; Miller 1995). Concept maps are viewed by some as useful ways to access hypermedia (Gaines & Shaw 1995; McAleese 1987). A recent issue of the Journal of Interactive Learning Research had a special issue dedicated to the use of concept maps in learning (JILR, 1997) indicating the continuing importance and usage of concept maps in education and educational research.

Previous work on the use of computer-based concept mapping tools have focused on their use by learners to express understanding but not the developers of teaching/learning materials. Kozma (1992) describes the use of Learning Tool with learners of English composition and found surprisingly that novices were disadvantaged in using the computer-based tool. Fisher (1992) describes SemNet but does not provide any clear evaluation. Kommers (1992) recounts the evaluation of

---

[3] Semantic nets are distinguished here from concept maps in usually having a pre-defined formal syntax of link types, link direction and node types. Concept maps have little syntax other than concepts and relationships, the relationships can be ambiguous, uni or bi-directional and un-typed. Concepts can be anything.

TextVision, a concept mapping tool incorporating sophisticated spatial - semantic algorithms, for use by learners to explore ideas.

Concept mapping could resolve some of the pedagogical and temporal issues associated with the production of educational hypermedia. It is possible to post-process concept maps to produce skeletal hypermedia which can be 'fleshed out' with content material and functionality. Closely linking a concept mapping tool with a hypermedia corpus would enable learners to create their own view(s), extract material as they browse and allow them to form their own notes, to be taken away on suitable magnetic media. It is believed here that simple concepts maps are not sufficiently expressive to allow for the features of educationally-effective hypermedia and additional functionality may be required.

The adoption of a concept mapping approach to hypermedia development raises a number of fundamental issues. Creating a 'view' of a particular subject would seem simple in theory but may be more difficult in practice. A concept map may prove a transitory, rather than definitive, picture of the domain (Jonassen and Mara, 1994). In well-defined subject areas, where major interrelations are generally accepted, concept mapping may be neither difficult nor transitory. Reader (1995) has evaluated computer-based concept mapping tools and warns of the dangers of forcing ideas into inappropriate representations. Therefore concept maps are not always the most appropriate representational system for all knowledge domains and are insufficiently expressive as discussed in chapter four, the design of a map-based HAP called *SHAPE©*.

Expressing knowledge through concept mapping is one thing, doing this directly via a computer is another. The analogy is of people who create with pen and paper but use a computer to re-present their creative work, it is the pen and paper which are the creative medium. With better word-processing packages composition directly at the keyboard can be facilitated. Similarly, to enable the easy development of hypermedia through concept mapping would be a tangible way of responding to the problems set out in the discussions above.

The Users of Concept Maps

An important distinction must be made at the start of this study about users. Previously, most work on concept maps, e.g. Kozma (1992), Fisher (1992), Reader (1995), Beyerbach and Smith (1990), has concentrated on their use by *learners* who are trying to acquire knowledge and skills in a particular domain. Little work has focused on its use by teachers for developing teaching materials, e.g. the MMI110 project described in McAleese (1986). In this study the users are *teachers* who already have (hopefully!) a clear understanding of the knowledge and skills of a domain and who are attempting to transpose that knowledge into a piece of educational hypermedia.

2.8 Conclusions

The current status of research and development of hypermedia, hypermedia authoring, hypermedia in education and the HCI issues of hypermedia have been presented. This

discussion has resulted in a number of problems and issues being identified. The use of concept maps has been proposed as a potential avenue to mitigate some of the problems identified. The implication of suggesting concept maps as a possible solution leads to the need to evaluate the use of concept maps as a paradigm for hypermedia authoring together with the paradigms already being used.

2.9 Objectives

The problems and issues identified above are extensive and possibly intractable so only a subset of them will be addressed in this study. The focus of this study is on the authoring of hypermedia by HE academic staff and in particular the learnability and usability of HAPs. In addition the use of concept maps as design metaphor in hypermedia authoring needs to be evaluated. Therefore the objectives of this study are:

Objective 1: To test the experimental hypothesis that diagram-based, in particular map-based HAPs are easier to learn and use than other HAPs.

Objective 2 To evaluate the learnability/usability of traditional HAPs for use in HE.

Objective 3: To elicit the cognitive level issues to learning and using HAPs by HE staff.

Objective 4: To determine the factors that need to be taken into account in the design of next generation HAPs to ensure their learnability and usability by HE staff.

The evaluation of the educational efficacy of the hypermedia produced by HAPs was considered to be beyond the scope of this present study although it is part of another on going PhD study (Dainty 1997).

2.9.1 Research Questions

This study attempts to understand and examine the issues raised in this chapter but also seeks to 'discover' other cognitive factors which could be important in determining the learnability, usability and utility of HAPs for HE teaching staff. Cognitive factors relate to the dynamics between the author, the interface and the HAP itself. Research questions have been developed to expand on the objectives of the study as shown below:

Objective 1:

- How does a diagram-based program differ from other types of authoring metaphor in terms of their learnability, usability and utility and their resultant hypermedia applications?

- Are explicit knowledge representations useful/helpful to the user? Do such representations enhance learnability usability or utility?

- Do different HAPs result in differently structured material?

- If a diagram- based HAP facilitates learning and using, how does it do so?

- How do the factors identified above interact with the different types of HAP?

Objective 2:

- What attributes do all HAPs share?

- What are the good attributes of HAPs?

- What are the bad attributes of HAPs?

Objective 3:

- What factors are important to learning HAPs?

- How do these factors affect learning HAPs?

- What factors are important to using HAPs?

- What is it about the user interface that makes a difference?

- How do the factors noted above interact with each other?

- What are the factors of users that affect a preference to create on paper or on screen?

- In what mode (textually or visually) do users prefer to create?

- In what way is user knowledge best reflected in hypermedia?

- How do the factors discovered in this study compare and contrast with those cited in the literature?

Objective 4:

- What are the relative effects of the factors of objective 1 on learnability, usability and utility?

- What aspects of the various HAPs help learnability usability and utility?

- What design metaphors are most suitable for hypermedia authoring?

Chapter Three

Methodology

3. Methodology

3.1 Analysis of Alternative Approaches

The objectives of this study and their associated research questions raised at the end of the literature review involve complex and multilayered phenomena due to the interactions of:

- the cognitions of the users;

- the knowledge domain and how it is 'held' in the users' minds;

- the computer interface;

- the operation of the program;

- the representation of the users' knowledge domain in the form of hyperlinks and knowledge chunks; and

- the mapping of users' mental models onto the design metaphor.

Therefore in order to achieve the objectives of this study and to answer the research questions required a multifaceted approach.

Many researchers continue to grapple with the complexities of HCI and there is a considerable body of literature defining models of interaction, user models and methodologies of study synthesised in section 2.4. Alternative approaches to answering the research questions include: (i) experimental; (ii) various usability testing techniques; (iii) other qualitative methods. Each type of approach has

advantages but also drawbacks. Experimental approaches are generally reliable and 'valid' but suffer from being artificial so even with valid results they might not be very relevant or have a narrow focus or not be generalisable. Typical methods of evaluation within usability engineering (Nielson 1993), such as heuristics, are more concerned with finished applications than in the discovery of phenomena of usability in general. The methods used in usability engineering could be adapted and modified for the purposes of this study, i.e. discovering underlying issues and processes. Developments in usability engineering have begun to assimilate the third approach of qualitative methods like 'thinking aloud' and protocol analysis (e.g. Ericsson and Simon, 1984). Qualitative methods have the advantage that they can produce more holistic data that address all the facets of the interaction, however there are two main disadvantages:

I   the difficulty of qualitative analysis; hence,

II the validity of the results of a qualitative data analysis (QDA).

It was decided that to answer the research questions satisfactorily and acknowledging the multifaceted nature of the problem all three approaches would need to be utilised. Nielson (1993b) in fact, recommends the use of a number of approaches so that they can provide different perspectives on the same set of phenomena.

In order to respond to the first research objective the evaluation of a diagram-based authoring program was required, implying the gathering of data on the learning and using of a diagram-based HAP. Appropriate responses to objective three required that observation and questioning of users learning and using HAPs was needed. Objective two was an extension of objective one - to evaluate other HAPs and then compare

with diagram-based HAPs. The investigation of objectives one and two would in turn provide some 'raw data' for the investigation of objective three, however a method for collecting and analysing the qualitative data was needed. It was therefore decided that the methodology would involve data gathered from subjects using authoring programs plus a method of analysis which enabled objective three to be evaluated. It was felt that to ensure a complete evaluation of the diagram-based HAP, an application should be built using it.

A classical experimental approach was considered to be inappropriate since this study was not trying to measure the effect of a 'treatment' on subjects but to analyse the interaction between humans and a computer program. Indeed, it is difficult to imagine the value of an experimental design that included a control group that was not subject to the treatment, i.e. the learning and use of a number of authoring programs. However, an experimental design approach does provide greater rigour and hence provides results of greater validity.

Usability, as discussed in section 2.4.3, incorporates a range of methods of evaluation including task analysis like task-action grammars, heuristics, performance measures, observation, questionnaires, interviews, focus groups, logging actual use and user feedback (Nielson 1993b). Task analysis and performance measures were considered to be inappropriate because of their low-level focus on psychomotor activity which would not provide data that would give a perspective on the cognitive issues. Heuristics are used at the design stage of an authoring tool and do not involve an evaluation of the user. Performance measures could be used although time-based

measures could be better used in subsequent studies. Some kind of measurement or analysis of the output from authoring programs was seen as an important element of the evaluation because of the research questions listed above about knowledge construction and representation. Logging actual use and user feedback are again better suited to beta-products with a large numbers of subjects. Interviews and focus groups were seen as being possibly useful because they elicit qualitative data that could shed light on the complexities of using HAPs. Questionnaires could be utilised because they provide data that can be used in quantitative analysis.

3.2 Selected Approach

It was felt that Kolb's experiential learning cycle (1984) shown in figure 3-1 was an appropriate model for approaching the various aspects of this study because of the iterative nature of study. Each element of the study enabled concepts and generalisations to be made which in turn could inform the next stage of the study. It was decided that objectives one and partially objective two could be addressed by a series of one-to-one training sessions in which subjects would learn and use a number of HAPs. As noted in the pilot study below a compromise was made between a more natural setting for the sessions, and rigorous experimental conditions. This resulted in the same office location being used for all sessions to ensure the validity of analysis and to provide the empirical foundation for the subsequent qualitative analysis. At the inception of this study there were no HAPs based on concept maps available and even now there are only one or two examples known, e.g. CourseWorks (Miller, 1995), and Webmapper (Freeman and Ryan, 1997).

Concrete Experience

Testing Implications
of Concepts in New
Situations

Observations and
Reflections

Formation of Abstract
Concepts and Generalization
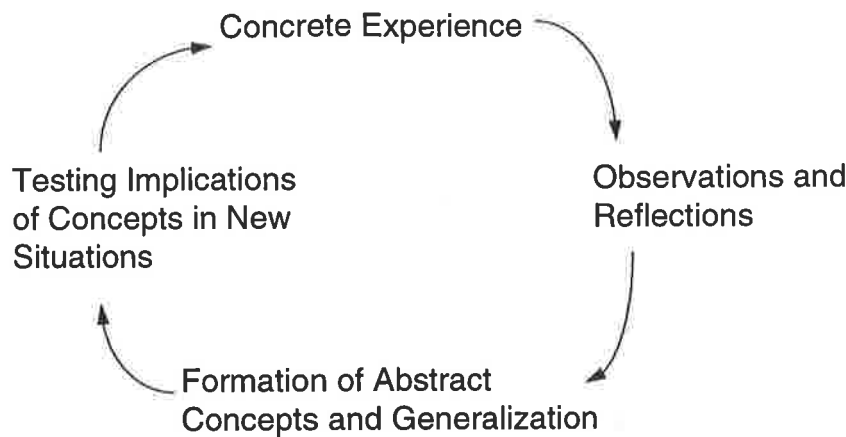
Figure 3-1 Kolb's experiential learning cycle (Kolb, 1984).

It was therefore necessary and desirable to design and build a diagram-based, more specifically a concept-based, authoring program to address objective one, the resultant HAP is called *SHAPE©*. By building a concept map-based HAP[4] a number of the other hypotheses established from studying the literature could be evaluated as discussed in chapter four. It was seen as important to build an actual application using the concept map-based authoring program to establish the feasibility of authoring real and practical applications. Authoring a real application provided a rich source of qualitative data on the use of diagram-based HAPs. Chapter four gives a detailed analysis and design of the map-based HAP together with the underpinning philosophies and hypotheses derived from the study of the literature. Chapter seven describes the construction of applications using the map-based HAP. The one-to-one training sessions incorporated quantitative data collection, in terms of questionnaires to measure users' evaluations of the programs, qualitative data in terms of observations, data from the actual one-to-one training sessions plus any comments

---

[4] concept map-based HAPs are referred to as map-based HAPs from this point onwards

made by the subjects and also a post-session interview to address specific issues of authoring.

Attention to objective three followed on from the analysis of the data from the one-to-one training sessions. It was decided that a focus group of experts would be the best approach to determine the factors that should be taken into account when designing and using HAPs and where the results of the analysis of the one-to-one training sessions could be evaluated in the broader context of other HAPs. The experts chosen for the focus groups were people who train others in the use of software programs, including HE teaching staff.

Figure 3-2 shows the relationship between the chosen methodology and the objectives.
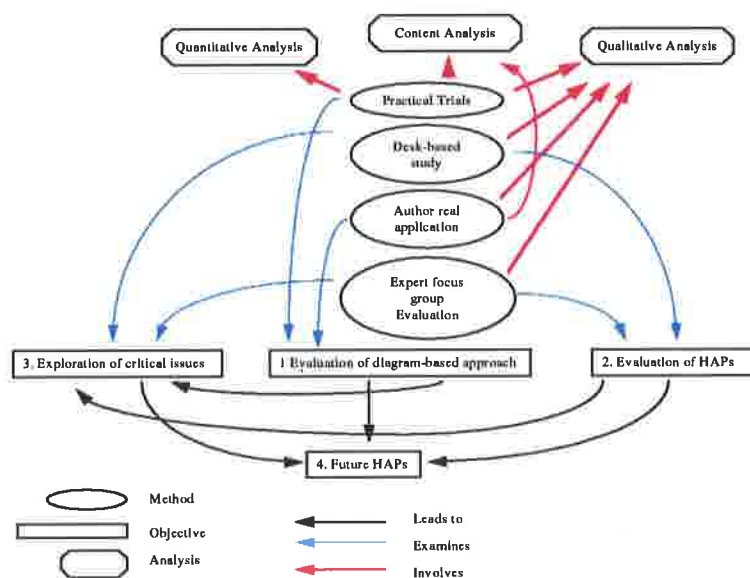
Figure 3-2 Overview of the methodology

The remainder of this chapter describes the design of the one-to-one training sessions and the focus group session together with the methods of data analysis employed.

3.3 Design of the One-to-One Training Sessions

3.3.1 Choice of HAPs

Section 3.2 gave a rationale for the use of one-to-one training sessions in order to examine objectives one, two partially three, this section attempts to provide the foundation for the design of the one-to-one training sessions. It was concluded that at least one other HAP would be needed to compare alongside the map-based one. There are many HAPs that could be used in the one-to-one training sessions, chapter two gives an analysis and taxonomy of the main HAPs in use. It was not practical to test all or even a majority of the HAPs and so a carefully selected subset was required. Objective four aims to determine which factors need to be taken into account in the design of new HAPs for ease of learning and use, so the focus was on HAPs that were already considered to be easier to use. Visual programming languages were rejected because they are not easy to learn or use and do not provide hyperlinking facilities. It was decided that control flow-based HAPs like Authorware would be best considered in the focus group evaluation because of their apparent complexity. This left the music score, and the directory structure and book-based HAPs. The book is humanity's most powerful and pervasive vehicle for externalising knowledge and was therefore a natural choice against which to compare and contrast the map-based HAP. Of the two

well-known book-based HAPs, HyperCard© is only available on a Macintosh platform and so ToolBook© was chosen for this study.

ToolBook© was simplified for the purposes of the one-to-one training sessions to reduce screen clutter, and make it more comparable to the simplicity and functionality of the map-based HAP with which it was being compared. Thus subjects would not have to confront the full extent of ToolBook© but would still experience the book-based HAP. The simplified version allowed a fairer comparison to be made. It is fairer to refer to the simplified program as the 'The Book Metaphor' (TBM) and not as ToolBook©. The modified program allows users to create standardised pages and provided easier hyperlinking to other pages than does the full version of ToolBook©. The modifications meant that instead of using the in-built mechanism for hyperlinking in ToolBook©, users need only type the name of the page to which they wished to link to in a special linking field incorporated with every content-page.

The choice to evaluate only two HAPs made sense for two reasons. First, it was far more practical to compare and contrast just two programs. Second, from the qualitative data perspective, it was felt that by interacting with only two programs subjects would be able to 'construct' a more meaningful comparison and evaluation. This is because the use of two HAPs encourages the development of 'personal constructs' in the same sense as Kelly's repertory grid (1955) by offering the subjects two poles (the two HAPs) which could be compared and contrasted. More HAPs would not have allowed this and may have led to a certain degree of confusion.

3.3.2 Pilot

In order to ensure a reliable and effective design of the one-to-one training sessions a pilot study was undertaken. The design of the one-to-one training sessions was established as comprehensively as possible to ensure that a valid and reliable comparison could be made between the two HAPs. This design was then piloted with 5 subjects who were teachers in HE. A walk-through of the pilot was also conducted with Dr Pamela Briggs at the University of Northumbria at Newcastle, an expert on usability evaluation, who provided many useful comments. The pilot led to many changes in the final design of the sessions. The sections below describe the design of the pilot and the changes made to the final design.

Design

Appendix D contains the initial pilot version of the script for the one-to-one training session which was used with two subjects. The problems with the initial pilot version were:

- Overuse of closed questions
- Lack of scales to objectively measure some of the dependent and independent variables?
- Lack of supporting explanations of:
    - Nature of study
    - Concept maps
    - Hypermedia
    - Hyperlinks
- Un-defined task in the script

- Separation of learning from using an HAP.

- Location of the session

- Length of the session

In addition to the above, it was originally decided to have 'use of pre-structuring of materials' as an independent variable, i.e. some subjects would have the opportunity to structure their ideas with respect to the task before starting the sessions. The problem with this was the number of additional groups that were required to maintain statistical significance so the variable was omitted in the final study.

The final pilot version took account of some of the problems that were identified. The final version pilot is also given in appendix D. The final version of the pilot was used with 5 subjects. The conclusions from the pilot study were:

- Better background explanations of the hypermedia, concept maps

- Use of ordinal scales to measure:

  - Prior experience of concept maps

  - Pre-session understanding of *SHAPE*© and TBM

  - Post-session understanding of *SHAPE*© and TBM

- Inclusion of some redundancy in the open questions used

- Addition of motivation as a dependent variable.

The walk-through of the design of the one-to-one training sessions with Dr P Briggs led to the following changes:

- More open questions to elicit the subjects' experiences and opinions of using the two HAPs;

- A more precise description of the task;

- The addition of a short post session questionnaire that asked subjects to compare and contrast the two HAPs; and

- The addition of 'critical incident' elicitation using a series of screen dumps of the use of the two HAPs at various stages.

The final design of the one-to-one training session is given in appendix D.

Instruments

The instruments that needed to be developed for this study were:

- Ease of Learning

- Ease of Use

- Task Match

- Computer Skill

As noted below Computer Skill was measured using the Computer Skills Inventory of the Computer Skills Metric as detailed in appendix A. The ease of learning and ease of use instruments were adapted from the scale developed by Molnar and Kleke (1996), which in turn was derived from studies by Davies (1989) and Napier *et al.* (1989). Molnar and Kleke's 18 item scale was used to test the usability of a voice-based interface to a spreadsheet package. Although, in principle, the scale was appropriate, certain questions were too specific to the application they were testing, a number of questions considered to be important were not covered and the scale did not distinguish between learning and using. Furthermore, the Molnar and Kleke scale did not address any cognitive issues which was one of the objectives of this study. 6 new questions were devised that examined the cognitive issues of confronting a new

program and 5 that examined the cognitive issues of using a new program. In the end only 6 items unaltered and one rephrased were retained from the Molnar and Kleke instrument. Two separate instruments were created, one oriented around the ease of learning the HAPs and one around the ease of using the HAPs. It was felt unnecessary to use the 7 point scale used in the Molnar and Kleke instrument so a 5 point scale was adopted. In addition to addressing the subjects' experiences and attitudes towards the HAPs, it was also necessary to ascertain how well they thought the HAPs matched the task required of them. It was decided that only two questions were needed to address task match, one that identified the subjects perceptions of how well the application would lead to a final application and one that overlapped with this to some extent and identified how difficult it was to complete the task.. It was decided that the instrument would be enumerated by coding each question with a value of 1 to 5, representing a low to high rating. The values for each question would be aggregated to give overall scores for ease of learning, ease of use and task match, making sure adjustments for reversals of negative responses were made.

Validity of Instruments

The validity of the computer skills metric had already been investigated (Elliott *et al.,* 1997) (reproduced in appendix A). The ease of learning and the ease of using instruments had been partially validated in that they were derived from a long pedigree as noted above. However, in addition to this fact, after the each of the subjects in the pilot had been through the session the values for ease of learning, ease of using and task match were presented to the subjects. Subjects were then invited to comment on whether they felt the values matched their perceptions of the ease of

learning and use and task match of the two HAPs. Subjects were happy with the how the values reflected their perceptions and experiences of the two HAPs, so no further adjustments were made to the instruments.

Procedure

The pilot used the place of work of the subjects as the location for the one-to-one sessions in order to maintain as natural a setting as possible. The problem with this was that the subjects were continually distracted by work colleagues, the phone and students interrupting. Recording the dialogue was also difficult because of background noise. So, although the setting was authentic, it proved to be impossible to conduct the sessions satisfactorily. A compromise between a nature work setting and a controlled experimental setting was made with the same un-used office used for all the sessions in the full study. The pilot sessions took about 3 hours to conduct which proved too much time for the subjects to commit themselves to and left them too tired to satisfactorily complete the session. To mitigate this problem the session was split into two, one session to collect the independent variable data of about a hour, and one session to conduct the one-to-one training session of about 2 hours.

Data Collection and Storage

In the pilot, a paper version of the ease of learning, ease of use and task match instruments were used. However this proved unwieldy so the two measures were computerised to facilitate the speed of data collection. As noted above, a script was developed and refined during the pilot which was then used in the full study. The final script incorporated the elicitation of a number of the variables which needed to be collected. This was done by creating a spreadsheet and typing the results in

directly from the script and alongside those results automatically collected by the computerised instruments.

3.3.3 Final Design

Sampling Frame

The inclusion criteria for subjects were that they:

- were HE teaching staff,

- had been in the Faculty of Business, Leisure and Food for at least 2 years (which school was not an issue).

- were of any gender

- were any age

- had not authored hypermedia before.

The reason for choosing subjects from one faculty was the problem of ensuring that the task to be undertaken was the same for all subjects. The difficulty of task could be controlled, since all the subjects were equally familiar with the faculty. In addition, a more valid comparison of the output from each HAP could be made. Members of the teaching staff were invited to participate in the study and around 40 volunteered. Subjects were chosen from this list on a random basis and assigned to one of two groups alternatively. The two groups were therefore as random as possible in terms of the independent variables discussed below and statistical comparisons of the independent variables showed there to be no significant differences between the two.

Sample Size

According to Nielson (1993), at least 10 subjects are required to ensure sufficiently valid results for usability analysis based on performance measures and therefore it was

determined the minimum sample size would 10. However due to a possible 'Order of Use' effect (discussed below) it was decided that 2 groups were required. So the final design used 2 groups of 8, totalling 16. This addressed Nielson's requirement and ensured statistical significance of the results. The choice of a relatively small samples is defensible for a number of reasons. The use of 16 subjects falls within the sample size of at least two recent PhD studies in a similar field, Reader (1995) used between 8 and 12 in his studies of the use of concept maps, Smith (1994) used 20 in her studies of navigation in hyperspace. Bryman and Cramer (1997) note that the ratio of sample size to population size is not usually relevant and in the 120 studies that they surveyed there was no relationship between sample size chosen and population size. Bryman and Cramer (1997) go on to emphasise that the practicalities of conducting a study is the main determinant of sample size. In this study each subject took about 2.5 to 3 hours to complete the session hence mitigating against the use of larger sample sizes. During the data collection period (about 12 months), a number of exploratory statistical tests were conducted. The trajectories of these tests were noted to determine their stability. After 16 subjects had been treated, it was clear that the results were fairly stable and that increasing the sample size would not significant affect them.

Independent Variables

The independent variables arise from consideration of the cognitive issues in chapter two. It was hypothesised that prior Computer Skills would affect the subjects' ability to learn and use a new program, that is, the higher subjects' computer skill were, the easier they would find the HAPs to use. Therefore, the null hypothesis tested was that

there would not be a positive correlation between computer skill and ease of learning and ease of use. Enthusiasm may be affected by their 'base motivation' (defined below) and their ability to author diagrammatically may be affected by the prior experience of concept maps and possibly their Spatial Relations ability which may be a reflection of their ability to describe ideas spatially. So, along with gender and age, the other independent variables measured were:

Previous Use of Concept Maps:

This was measured with a simple four-point Likert scale from 'No familiarity' through to 'Very familiar'.

Prior Computing Skill:

A subject's experience of computers and their Windows skill was seen as an important variable in evaluating the two authoring metaphors due to the transfer principle (Waern, 1990). The more experienced subjects were with Windows (and Windows-based programs) the more rapidly they would learn and use any new program To measure their Computer Skill, subjects completed the Computer Skills Metric for Windows (CSM(Windows)) (Elliott *et al.*, 1997). In the study Elliott *et al.* (1997) compared two user groups, students and HE teaching staff, where the correlation between the HE teaching staff's declared skill and their actual skill (measured in five simple computerised exercises) was significant but for students it was not. Hence in this study it was felt that it was satisfactory to use the Computer Skills inventory element of CSM(Windows). Appendix A gives an account of the design and testing of CSM(Windows).

Spatial Ability:

It was hypothesised that subjects' ability to think spatially might be positively correlated with the ease with which they constructed spatial representations of a knowledge domain. Therefore the null hypothesis tested was that there would not be a positive correlation between spatial relations ability and the ease with which they constructed spatial representations of a knowledge domain. The ease with which subjects constructed spatial representations of a knowledge domain was measured by the richness of the maps produced, i.e. the number of links, concepts and levels of concept maps. To this end the Spatial Relations element of the Technical Skills Battery (The Psychological Corporation, 1996) was administered prior to the experiment.

Order of Use of Packages:

It was anticipated that after subjects had attempted the task using one of the metaphors they would have developed a mental 'picture' of the subject domain which could influence their perceptions when using the other metaphor. In order to attempt to control for this eventuality subjects were randomly placed into two groups, using the two HAPs in different order.

Base Motivation:

It was anticipated that the subjects' own motivation levels would affect their Motivation to Continue with the HAPs so a measure of it was required. Amabile *et al.* (1994) have validated their work preference inventory which measures both

intrinsic and extrinsic motivation. It was decided to use this test to measure the subjects' base motivation.

Dependent Variables

The dependent variables were: Ease of Learning (EL), Ease of Use (EU), Task-Match and Motivation-to-continue, measured using a Likert scale-based questionnaire. The time taken to complete an activity is seen as an important variable in HCI studies (Jordan *et al.*, 1991) however in this study its relevance was doubtful and making meaningful measurements difficult. As noted at the end of the literature review, the educational efficacy of hypermedia produced from the HAPs was to some extent ignored in this study because of the focus was on authoring not on use. Educational efficacy was measured to some extent in the development of a real application (this is considered in section 7.3).

There are many references to measures for EL and EU for computer interfaces/programs in standard textbooks on usability, (Davis 1989; Napier *et al.*, 1989). This study has adapted the questionnaire developed by Molnar and Kletke (1996) but has distinguished 'EL' from 'EU', as noted in the pilot study (section 3.3.2). The question bank contains two questions that related to how well subjects rated the HAP's task-match to the exercise. Appendix C contains the questionnaire used. For the purposes of the one-to-one training sessions the questionnaire was computerised.

Hypotheses

Subsumed within the overall experimental hypothesis of objective one, outlined in section 2.9 are sub-hypotheses that directly relate to the independent and dependent variables outlined above. These sub-hypotheses are set out below:

Hypothesis H_1: Task Match

There was no reason to believe that either HAP would produce a better hypermedia than the other, since subjects could only construct content-chunks or hyperlinks and both HAPs had mechanisms for doing so. Subjects' rating of the Task-Match of two HAPs to authoring hypermedia would therefore be the same or the difference would be insignificant. Therefore the null hypothesis is:

H_0: There is no significant differences in the task match between the two HAPs.

Hypothesis H_2: Computer Skill and Ease of Learning

Subjects with a higher CSM(Windows) score would rate the EL of **SHAPE**© higher than that of TBM because of their familiarity with graphical interfaces. Therefore the null hypothesis tested is

H_0: There is no correlation between Computer Skill and EL for either HAP.

Hypothesis H_3: Computer Skill and Ease of Use

Subjects with a higher CSM(Windows) score would rate the EU of **SHAPE**© higher than that of TBM. Therefore the null hypothesis tested is

H_0: There is no correlation between Computer Skill and EU for either HAP.

Hypothesis H_4: Order of Use

The order of using each authoring HAP (*SHAPE*©/TBM or TBM/*SHAPE*©) would affect subjects' ratings of EL and EU. More specifically the ratings of EL and EU of *SHAPE*© over TBM would diminish for subjects who used *SHAPE*© first since they would already have developed a better mental image of the knowledge domain than if they had used TBM first. This hypothesis encompasses two sub-hypotheses:

Hypothesis H_{4a}: Order of Use and EL

The Order of Use will affect the EL ratings of each HAP. The null hypothesis is therefore:

H_0: There is no effect of Order of Use on the EL ratings of each HAP.

Hypothesis H_{4b}: Order of Use and EU

The Order of Use will affect the EU ratings of each HAP. The null hypothesis is therefore:

H_0: There is no effect of Order of Use on the EU ratings of each HAP.

Hypothesis H_5: Spatial Relations ability and EL

Subjects' EL rating of *SHAPE*© would positively correlate with their Spatial Relations ability. There is unlikely to be any correlation with TBM because it does not have a spatial interface. Therefore the null hypothesis tested is

H_0: Subjects' EL rating of *SHAPE*© is not positively correlated with Spatial Relations ability.

Hypothesis H_6: Spatial Relations ability and EU

Subjects' EU rating of *SHAPE*© would positively correlate with their Spatial Relations ability. Therefore the null hypothesis tested is:

H_0: Subjects' EU rating of *SHAPE*© is not positively correlated with Spatial Relations ability.

Hypothesis H_7: Prior Experience of Concept Maps and EL

Subjects' EL rating of *SHAPE*© would positively correlate with their Prior Experience of Concept Maps. Therefore the null hypothesis tested is:

H_0: Subjects' EL rating of *SHAPE*© is not positively correlated with their Prior Experience of Concept Maps.

Hypothesis H_8: Prior Experience of Concept Maps and EU

Subjects' EU rating of *SHAPE*© would positively correlate with their Prior Experience of Concept Maps. Therefore the null hypothesis tested is:

H_0: Subjects' EU rating of *SHAPE*© is not positively correlated with their Prior Experience of Concept Maps.

Hypothesis H_9: Motivation to Continue

Since it is anticipated that subjects' will prefer **SHAPE**© they will therefore be more motivated to continue with it than TBM. Therefore the null hypothesis tested is

H_0: The Motivation to Continue with **SHAPE**© will be no different from that of TBM.

The Setting

The setting was designed to be as close to one-to-one tutorials as possible, bearing in mind that evaluations were taking place concurrently. However to maintain consistency the same office location was used for all sessions as noted in the pilot study.

The Task

It was important that the knowledge domain was as familiar as possible to ensure parity of task for all subjects. The task chosen was to construct a skeletal hypermedia application that contained at least three hyperlinks that described the faculty, its courses and staff to potential students. All subjects were situated in the same faculty so their understanding of the term 'faculty' should have been similar as noted in the sample frame. It was not necessary to add content in the form of text and other media as these techniques are the same for both metaphors.

One-to-One Training Sessions Procedure

On an individual basis, subjects were asked to complete a computerised Computer Skills Metric (Elliott *et al.*, 1997). Subjects were then alternatively assigned to one of

two groups, one using TBM followed by *SHAPE©* and the other group using *SHAPE©* followed by TBM. Each session was undertaken consistently by following a script administered by the researcher which can be found in appendix D. Subjects learnt to use each application by completing the task as described above. Following the learning session, subjects were asked to complete the 'EL', 'EU' and task-match computer-based questionnaire. Once subjects had completed the exercise using one HAP they attempted the exercise using the other.

There were 2 scripts used, one for *SHAPE©*/TBM and another for TBM/ *SHAPE©*. The steps followed in the scripts were:

1. The concept of hypermedia and hyperlinks was described to the subject;

2. A hypermedia application was demonstrated so that all subjects were aware of what a hypermedia application was and that they could then see the relevance of hypermedia to their own teaching;

3. Subjects were then asked to rate their understanding of hypermedia and hyperlinking on a scale from one to five (see appendix D for details of this scale);

4. The first package was described:

   - For TBM this meant describing the book metaphor and how hypermedia applications were constructed using it;

   - For *SHAPE©* this meant describing concept maps and how they could be used to construct hypermedia. A picture of a concept map illustrating how concept maps equate to hypermedia was shown;

5. Subjects were asked to rate their understanding of the first HAP and when *SHAPE*© was being used, their prior knowledge of concept maps;

6. The task was explained to the subjects and the details were left available for the subject to peruse during the training session;

7. During the training session the subjects were given control of the computer and were guided through the construction of the task. Subjects were requested to voice any questions, misunderstandings or anything they did not like or like about the program. The training sessions were audio-recorded for later transcription;

8. After the training sessions the subjects restated their understanding of the HAP in question. It was important to note here that they were not asked to rate their skill of using the HAP but their understanding. The degree to which the underlying metaphor or paradigm of the HAP was transparent to subjects was of more interest than their perceived acquired-skill level, which would undoubtedly be low anyway;

9. The subjects were then asked to complete the EL, EU and Task-Match inventories on the computer;

10. Following the inventory, subjects were asked a series of open questions (as detailed in appendix D) to elicit the issues and phenomena of hypermedia authoring. The subjects' answers were tape-recorded;

11. Following the open questions, a series of 'screen dumps' capturing a typical session using of the HAP were shown to the subjects who were then invited to describe what they understood and remembered and also what they disliked or liked about that part of the HAP;

12. The procedure was then repeated for the other HAP;

13. After using both HAPs the subjects were invited to compare and contrast the two.

Qualitative Data Collection

During the one-to-one training sessions all subjects' questions and the guidance given by the trainer was recorded. Subjects were asked a series of open questions after using each HAP to elicit the issues underlying their experiences and evaluations of the HAP. After using both HAPs they were invited to offer any general comments comparing and contrasting the two programs. Subjects were prompted for their responses with the aid of a set of cues shown in appendix D. Each subject's output from each HAP was stored and saved for later analysis as discussed in section 3.5.4 below.

Analysis of Quantitative Data

The independent and dependent variable data collected before and after the training sessions were entered into SPSS (1993). To evaluate the hypotheses a series of statistical tests were carried out as explained below.

It was decided to conduct an exploratory Principle Components Analysis (PCA) (Bryman and Cramer, 1997) on the EU and EL scales in order to detect any underlying factors that might contribute to usability. PCA was considered suitable since Molnar and Kleke (1996) conducted a PCA on their scale, the one from which the EL and EU scales are derived (their results are compared with those found in this study in section 5.2) The results of the Principle Components Analysis are presented in chapter five. The reliability and validity of the use of Principle Components Analysis is subject to some debate. Bryman and Cramer (1997) recommend samples sizes of at least 100 to improve validity of the derived factors. In this study there was only 16 subjects so the

validity of any derived factors was potentially dubious. However, the approach taken was to conduct a separate analysis on the ratings for each HAP and then compare the two sets of data. By taking this approach only those factors which were common to both sets of data were accepted as having sufficient validity and the reliability of the factors would be more assured. Also any factors emanating from the Principle Components Analysis could be corroborated with the qualitative analysis. An SPSS statistical expert was consulted about the above strategy, who gave a conditional affirmation of the appropriateness of this approach. Tabachnick and Fidell (1996) note that where there are less 5 cases per variable, Barlett's Test of Sphericity (Tabachnick and Fidell, 1996) can be used to measure the suitability of using PCA. In this study Bartlett's Test of Sphericity showed that PCA was appropriate for this data (see section 5.2 and Appendix I for values). In addition the Kaiser Meyer Olkin (Tabachnick and Fidell, 1996) measure of sampling adequacy (see section 5.2 and Appendix I for values) was also applied to the data and was significant, indicating the appropriateness of using PCA (Tabachnick and Fidell, 1996).

Statistical Tests Used

Details of the tests were obtained from Bryman and Cramer (1997), Kitchens (1996), Weiss (1995) and Meyers and Grossen (1978). Table 3-1 shows the assumptions required for using various test. All variables were tested for normality using both Shapiro-Wilks and K-S (Lilliefors), where data was found to be normal, parametric tests were applied as appropriate and where data was considered non-normal non-parametric tests were similarly applied. A Cronbach's Alpha test for reliability of

scale was initially applied to the EL, EU and Task-Match scales and indicated the scales were generally reliable.

Table 3-1: Assumptions applied to the use of statistical tests.

Test	Assumptions	Recommended[5] Sample Size
Paired sample t test	Normal distribution/parametric, same variance. Interval or ratio data. Matched data.	Less than 30
Wilcoxon matched pair signed ranks	Ordinal or interval or ratio data. Matched data No assumptions about distribution.	At least 10
Mann Whitney U test	Ordinal or interval ratio data. two sets of data. No assumptions about distribution.	None found
Pearson's correlation coefficient	Two sets, interval or ratio data	None found
Kendall's tau-b correlation coefficient	Two sets of ordinal data No assumptions about distribution.	None found
One way ANOVA	Interval/ratio three or more conditions. Normal distribution.	None found
Kruskall-Wallis test	Ordinal/interval or ratio data, three or more conditions No assumptions about distribution.	5 or greater

The Paired Sample t-test was used for comparing the EU and EL between each HAP. The Wilcoxon Matched Pair Signed-Ranks test for non-parametric comparisons was used to compare Task-Match, Declared-understanding, and Motivation between each HAP. The Mann-Whitney U-test was used for comparing output from each HAP. Pearson's Correlation coefficient was used to check for relationships between:

- EU and EL for each HAP,

- Computer Skill and the dependent variables,

- Spatial Relations and the dependent variables,

---

[5] Recommended sample sizes and assumptions found in Bryman and Cramer (1997), Kitchens (1996), Weiss (1995) and Meyers and Grossen (1978).

- Base Motivation and dependent variables and

- age and dependent variables. The Kendall's tau-b coefficient was used to check for relationships between Previous Use of Concept Maps and dependent variables. The standard one-way Anova was used to measure effect of the Order of Learning on EU and EL for each HAP. The Kruskal-Wallis one-way Anova was used to measure effect of Order of Learning on dependent variables for each HAP. The Kaiser's Criterion was used to qualify factors in the Principle Components Analysis and Bartlett's Test of Sphericity (SPSS, 1994) was used to check the suitability of using Principle Components Analysis.

3.4 Analysis of the Authoring of a Real Application

Ideally a number of real applications would need to be built using *SHAPE©* by a number of subjects for a complete evaluation. In time this will happen, however it was felt that a substantial contribution could be obtained from the authoring of just one complete application. The subject who was chosen has a high level of computer skill and is a lecturer in HE. The reason for selecting such a person was that he would be not only be able to offer constructive comments on the learnability, usability and utility of *SHAPE©* but would be able to consider this in the context of a typical member of HE teaching staff using it. There was an added advantage in that the chosen subject had never constructed hypermedia before. The subject was left to decide how best to author the application and the only help given was in the mechanics of using *SHAPE©*. The application chosen was to construct a open and distance learning hypermedia package to help teach an introductory module on IT.

The subject was asked to note anything dysfunctional about *SHAPE*© and asked to consider how suitable concept maps were for creating the hypermedia.

To encourage an intelligent evaluation of *SHAPE*© by the developer, another similar application was built using another HAP. It was felt that it would be useful to use a HAP that is used for authoring Web documents because of the increasing importance of the Web in HE. Microsoft FrontPage© was selected because it was considered to be the most widely used Web-based HAP in HE (excluding writing HTML).

To allow the subject to relate his experiences of using *SHAPE*© further he was invited to be one of the members of the focus group described in section 3.6. The finished application, later called **CLICK-IT**©, was also examined to evaluate the practicalities of using concept maps to author hypermedia. The analysis of **CLICK-IT**© is given in chapter seven, section 7.3.

3.5 Qualitative Data Analysis (QDA)

Chapter five presents the analysis of a quantitative comparison of *SHAPE*© with TBM and provides empirical evidence that teaching staff preferred concept mapping to the book metaphor for constructing hypermedia. Chapter five also gives a Principle Components Analysis of the quantitative responses of the subjects to the two HAPs and yielded some general factors of usability. The results of chapter five revealed some understanding of user cognition with respect to the suitability and usability of concept maps in order to help teaching staff develop their own hypermedia. However quantitative analysis is limited. Prevailing usability evaluations all have draw-backs:

all forms of task analysis are very involved, psychomotor-oriented and lack a holistic perspective; the experimental analysis is also limited as indicated above. In order to probe deeper into the issues surrounding the use of concept maps by the subjects some form of dialogue with the users was required. A means of analysing that dialogue was then required. Section 3.5.3 describes the chosen approach of QDA.

Mental Models

Section 2.4.1 in the literature review discussed the importance of the role of mental models in designing and evaluating usability. The reductionist approach of GOMS (Card *et al.*, 1983) and Payne's (1985) Task-Action Grammar was rejected in favour of a holistic approach advocated by Green and Petre (1996) and Briggs (1988). Amongst other factors, the case for using a diagrammatic approach to hypermedia authoring depends on the model users have, or develop as a result of using the authoring programs. If the program fits or modifies itself to a user's received model prior to using programs then that user will accept the system more readily. If, as a result of using the program, users' evolving mental models of the program is clear and fits their 'loose model' of how a program should be then again users will accept the new program. 'Loose model', as defined here, refers to all the prior experiences and pre-conditions the user brings with them when they use a new program. The loose model will be rich and idiosyncratic, well beyond the practical possibilities of modelling. It is believed here that the loose model will be to some extent dependent upon the users' prior Computer Skills and in the case of diagrammatic authoring like concept mapping is dependent upon their Spatial Relations ability. The book metaphor was originally chosen as a bench mark in this study because of its ubiquity and therefore its association with users' loose models.

The users' mental models of the system will become rich and complex and deeply affected by their loose model. The quantitative analysis of chapter five would be useful in discovering some aspects and truths of the phenomena underpinning hypermedia authoring but it could be not capable of addressing the issue of the user's mental model of the two programs. This is another reason why a QDA was seen as essential in completing this study.

3.5.2 Objectives of the QDA

It is important to point out that the objective of this study was not just the usability assessment of some HAPs but also the discovery of the issues, principles and phenomena that affect staff authoring hypermedia. The intended outcome of the study was therefore not the evaluation of the usability of the two programs, although this was an outcome, but the elucidation of the cognitive factors surrounding the facilitation of hypermedia development, explored through these programs. The purpose of the QDA was therefore to:

1. identify corroborating evidence for statistical findings of chapter five;

2. identify corroborating evidence for any factors originating from Principle Components Analysis in chapter five;

3. compare *SHAPE©* and TBM on the basis of identified usability factors;

4. determine any interaction between the independent variables and the above objectives;

5. identify underlying principles and themes related to the learnability and usability of HAPs in the context of its use in HE.

It is important to emphasise the exploratory nature of QDA. The fifth objective is considered in the discussion, presented in chapter seven.

3.5.3 Methodology

Quantitative measurement of usability in terms of completion times and success rates has been criticised for its lack of depth. Briggs (1988) argues that the keystroke level assessment of usability *'precludes any possibility of generalisability of results because of the subjective nature of assessment'*. In his account of the design of Windows 95 Sullivan (1996) stresses the need for a holistic approach to usability analysis where subjects used the whole of Windows95 in order to assess how each element of the interface interacted with each other and with the users.

Briggs (1988) recommends methodologies which take account of subjective data and their subsequent analysis. The alternative qualitative methodologies are briefly:

Critical Incident: Briggs has applied Flanagan's (1954) critical incident technique to the study of usability. Users can reveal fundamental issues when asked to relate their experiences when using a computer program of misunderstandings they may have or weaknesses in the usability of some facet of the program under study

Prompted Recall: Prompted recall (Barnard *et al.*, 1986) uses pictures of various screen displays from the program under study as a recall cue.

Statements made by a user can then be analysed.

Protocol
Analysis:

Protocol analysis attempts to capture the users' thought processes whilst using a program. The problems here are first getting users to 'think aloud' whilst using the program and then attempting to analyse the data. 'Thinking Aloud' as a technique in its own right has been investigated by Ericsson & Simon (1984).

Question-Asking
Protocols:

Kato (1986) gives a critique of the think aloud protocol and concludes it is difficult to apply. Kato's question-asking protocols alternative relies on a more natural 'dialogue' between researcher and user.

Repertory Grid:

Kelly's repertory grid (1955) has been used to explore important concepts and relationships with respect to knowledge elicitation (Boose, 1985) but no record has been found of its use in the evaluation of computer systems.

The Analysis of Qualitative Data

There is a substantial body of literature that describes qualitative data collection as noted above but surprisingly little reference to QDA. The techniques listed above are more to do with data collection than data analysis. In the study by Kato (1986), qualitative data from question-asking protocols is placed into 4 pre-defined categories and indicates a rather limited analysis of the data compared to the extent of the analysis carried out in psychological and sociological studies, e.g. (Weatherburn *et al.*, *1992*; Atkinson and Drew, 1979) This lack of a rigorous approach to QDA may

reflect the size of the task - qualitative data is generally voluminous, complex, and unstructured. Briggs (1988) remarks on the difficulty of analysing such volumes of data. Hierarchical categories (Richards and Richards, 1994) offers a clear approach for analysing unstructured qualitative data but has been used primarily in sociological studies because of the complexity of human interactions however it could be used for QDA.

Hierarchical Categories

Richards and Richard's (1994) paper probably marks the first attempt to articulate a formal methodology for qualitative data analysis with hierarchical categories. Hierarchical categories-based qualitative data analysis (Richards and Richards, 1994) was employed here because it enables the data to be indexed to increasingly fine categories so that, eventually, the 16 individual interviews could be integrated into a single encompassing model that described all the data.

Grounded Theory or Prior Theoretical Framework

There is a debate whether QDA should proceed, based on theory originating either from the researcher or from the literature, or should be based on seeing what 'emerges' from the data which is sometimes called the grounded approach (Strauss and Corbin, 1990). There is a danger that the grounded theory approach may result in failure (Yin 1994) for inexperienced researchers.

The statistical analysis and notably the Principle Components Analysis of chapter five were used as the basis of the QDA as suggested by Yin (1994). This approach was

taken despite reservations by Bryman (1988) that prior theory introduces premature closure. The argument for its use here is that the theoretical framework used is derived from the same data origins and not specifically on a theory from elsewhere. Pre-defined usability factors from the literature, however, were not considered until after the categorisation had taken place. This was done so that it would not influence the analysis of the data and hence the factors that emerged could be compared to those defined elsewhere for verification and comparison

The Adopted Approach of Data Elicitation

In order to obtain sufficient and relevant qualitative data from the subjects whilst undertaking the experiment, a combination of 'Protocol Analysis', 'Prompted Recall', and Question-Asking Protocols' were used. Critical Incident was rejected because it would rely on subjects' experiencing some 'critical' events but the reality might be more prosaic and subjects might not report anything. Repertory grid was rejected because after using it in an early pilot session it was found to be extremely cumbersome.

Selected Techniques for Data Elicitation

During the training sessions a combination of protocol analysis and question-asking protocols were employed. This meant encouraging subjects to articulate their experiences as they learnt and used the two programs, interspersed with prompts by the researcher to comment on particular aspects of the programs. Following the sessions a set of pre-defined questions were used to elicit any experiences, thoughts, impressions, etc. the subjects had in using the two programs. These questions are

described and explained in appendix D. Subjects were then shown a series of pictures on screen depicting the various stages in the use of the two programs. In this fashion a considerable body of data was collected from the subjects.

NUD.IST

NUD.IST (Richards and Richards, 1991) is a computer-based tool for undertaking hierarchical categories-based QDA. NUD.IST allows its users to create a hierarchy of categories, to which qualitative 'evidence' or data can be indexed. For instance all references in the data to a particular phenomena, e.g. 'comfort' would be indexed to a category called 'comfort'. The category 'comfort' could then be sub-categorised into categories like 'pleasure', 'physical', 'mental', etc. so that a hierarchy of categories would 'evolves' from the data. The unit of data is a data bit and constitutes anything from a single spoken phrase, a whole dialogue, through to other literature references. Each category is represented in NUD.IST as a node in a tree of categories as illustrated in figure 3-3. below. NUD.IST also allows users to carry out sophisticated searches like the intersection of all references to a particular phenomena like 'EL' and age so that an analysis of the differences in age can be undertaken.
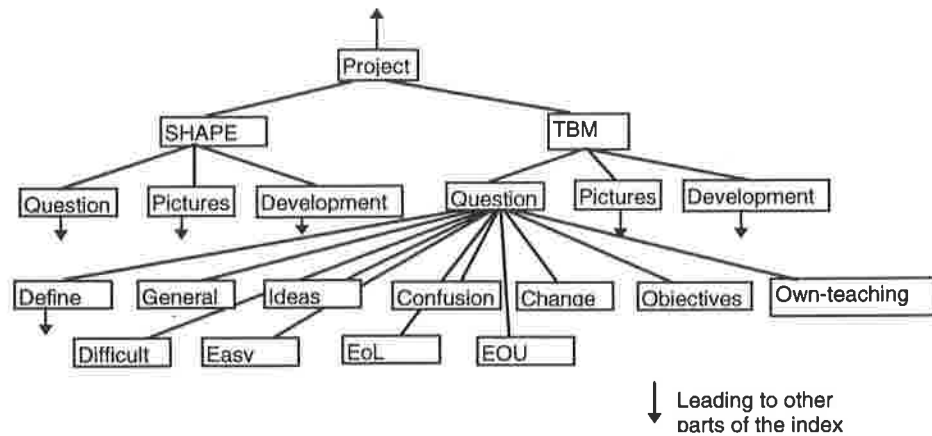
Figure 3-3 Portion of index tree showing the nodes related to the questions asked of subjects

Process of Categorisation

All sessions were taped and then transcribed. Each transcription was imported into NUD.IST and the responses to each question were indexed to a particular node in the category hierarchy. For example, all the responses to the cue which asked subjects to define what they understood to be the purpose of TBM were indexed to a node entitled 'Define TBM'. A portion of the whole hierarchy category tree is illustrated in figure 3-3, the whole tree can be found in appendix E. In accordance with the basis of approach indicated above, a node for each of the factors yielded in chapter five was created and the data searched for corroborating evidence. Once this was done the data was analysed for finer categories. The evidential data for the factors was subcategorised according to which of the two applications it referred to, so that a comparison of the transparency of *SHAPE*© could be compared to that of TBM and therefore conclusions could be drawn. The final part of the process was to synthesise

the findings. The most effective way found to undertake this synthesis was to place all the 'findings' on a large piece of paper so the complete picture could be seen.

The Independent Variables

Subjects' transcripts were also indexed to nodes that reflected quantitative variables, such as computing skill or Spatial Relations ability. This is illustrated in the tree in figure 3-3. This was done so that searches of comments of particular combinations of quantitative variable could be made, for example comments on the EU of *SHAPE*© could be separated according to the subject's computing skill.

3.5.4 Content Analysis of the Output

An important aspect of the analysis of the two programs was to measure the Task-Match of each. That is, how well the programs allowed users to complete the task. The quantitative analysis showed that the subjects thought that the programs were identical in Task-Match. It is necessary however to analyse the outputs from each program to see how well they matched the task. It could be that although both programs matched the task equitably, it was a very low match. There is also a need to analyse the output from each program to identify if they have any negative or positive or spurious effects with respect to the required task.

The problem with analysing the output from the programs is the diversity and synonymity of the data. The quantitative analysis of the outputs from the HAPs revealed the richness of the *SHAPE*© output in terms of links and concepts created, however there is a question of whether those links and concepts are idiosyncratic and have little relevance to the required task, i.e. a low utility (Shackel, 1986). A

satisfactory quantitative analysis of this aspect of the output is impractical - the exercise is essentially a qualitative one - comparing and contrasting the outputs from each HAP for each subject. Content analysis (Berelson, 1954) could be a possible technique to use although it relies on measuring the frequency of the phenomena. This study was not interested in how prominent a particular concept or idea is *per se* but how prominent it was to each subject relative to the other subjects and to the task.

An analysis of the content is also problematic:

- How can the outputs to be compared?

- What concept labels can be considered to be 'official' synonyms and what labels not?

- Should each output for each subject be compared to all the outputs of all the other subjects or should they be compared to some predetermined standard?

- How should the TBM output be compared to the **SHAPE**© output?

Bearing in mind the objectives of this study it was decided that all outputs would be compared to a 'standard model' in terms of the concepts and the links used and any identifiable idiosyncrasies. The standard model was viewed as a kind of generic or idealised representation of what the task-output should look like, the problem was finding the generic picture. By comparing each subject's output for each HAP with the generic model an idea of how well each program produced work that approached the ideal was developed. Clearly, this exercise only tests the output for one particular task and the Task-Match to other tasks might differ, a point considered in the discussion.

The Standard Model

The Standard model was formulated by analysing and synthesising the design of the Web sites of similar HE faculties offering business programmes in the South Western part of the UK. If any of the outputs from the programs was similar in structure to the standard model it would mean that the program was capable of a high Task-Match. Any variances from the standard model would needed to be accounted for, i.e. was the program at fault, was it doing a better job or was it something to do with the idiosyncrasies of particular authors?

Appendix F contains the representations of the structure of the Web sites analysed for this study. The derived standard model is shown in the Fig 3-4 below. The standard model was found to be hierarchical but generally had only two levels and there was little linkage between concepts/nodes of information at the same level.

* Staff mostly refers to academic staff but some sites refer to admin staff as well
** The word School covers faculties where they might have departments or discipline areas instead but they refer to synonymous concepts.
*** Courses sometimes fell under a particular school and were not represented at the top level
+ Other material refers to projects, research interests etc.
++ Some sites had some space devoted to management which sometimes sat at the top level, sometimes it sat at a lower level.

Concepts encapsulating other lower level concepts

Links/hotbutton to lower levels

Lowest Level concept probably represented by a single page

Figure 3-4 The standard model of the structure of a hypermedia describing a business faculty

Method of Content Analysis

The skeleton table shown in figure 3-5 below shows the headings used to compare the outputs from each program with each other and with the standard model for each subject. The actual results were hand written on an A3 sheet of paper (a word-processed version is shown in appendix G).

Subject	TBM versus standard model			SHAPE© versus standard model			TBM versus SHAPE©
	Nodes	Links	Idiosyncrasies	Nodes	Links	Idiosyncrasies	
S1							

Figure 3-5 Structure of table used for content analysis of the output from the one-to-one training sessions

3.6 Design of the focus group session

The value of focus groups in the evaluation of HCI has only recently been recognised. Nielson's (1993) text on usability engineering mentions the use of expert evaluation. O'Donnell *et al.* (1991) describes the use of focus groups in the evaluation of a central heating control interface and analyses the benefits of using focus groups over the more usual approaches of examination of individual usage. O'Donnell *et al.*'s argument for using focus groups is that it avoids the dangers of introspection, i.e. significant personal idiosyncrasies may result from the accounts given by individual subjects and hence no real insight into the cognitive processes taking place elucidated. In O'Donnell's words, focus groups "allow subjects to remind one another of events;

they encourage subjects to reconstruct processes; they explore gaps in the subject's thinking; they overcome the 'not worth mentioning' problem, and importantly for HCI, they allow new solutions to emerge".

Choice of HAPs

The overall purpose of the focus group session was to evaluate the issues surrounding hypermedia authoring by HE teaching staff, uncovered in the analysis of the data from the one-to-one training sessions. Since only two types of HAP were evaluated in the one-to-one training sessions, it was felt that a number of the other key HAP types should be involved in the focus group sessions. From the analysis of the HAP types given in chapter two it was decided that a visual language-based HAP and a control flow-based HAP should be used because, together with the two HAPs evaluated in the one-to-one training sessions, they represented the majority of the programs used to author hypermedia. On investigation it was found that many visual languages do not support the construction of hyperlinks. The use of visual languages was therefore abandoned. There are two well known control flow-based HAPs: IconAuthor$^{@}$ and Authorware$^{@}$. It was decided that only one HAP was required in this evaluation. At the time of the focus group session a small number of other map-based authoring HAPs had become available as a result of research into Web authoring so it was decided to include at least one other HAP based on concept maps. Similarly a number of visual HTML editors had also become available, so it was decided that at least one of these should be included in the focus group session. Therefore the final list of HAPs included in the focus group session were:

- *SHAPE©*

- A book-based HAP (The full version of ToolBook©).

- One other map-based HAP (Webmapper, (Freeman and Ryan (1997)).

- A control flow-based HAP (Authorware©).

- A visual HTML-based HAP (FrontPage©).

No examples of HAPs for authoring open hypermedia, such as Microcosm and Hyperwave as discussed in chapter two, were included because the emphasis in the study was on discrete hypermedia applications designed for particular teaching and learning scenarios. The analysis of the chosen HAPs is given in chapter seven together with the analysis of the results of the focus group.

3.6.2 Procedure

The participants were given the context text given in appendix H to read and were then told the following:

- Thank you for agreeing to participate in this study. You are invited to comment at any stage in the session.

- The context for this study are the barriers confronting HE staff developing hypermedia.

- The objectives of this study are to:

    1. Assess the usability, learnability and utility of a number of hypermedia authoring programs based on different paradigms;

2. Investigate what makes a program suitable for HE staff to author hypermedia;

3. Investigate the issues surrounding creating-on-paper versus creating-on-computer;

4. Establish the relevance of the factors developed in the previous study.

The paradigm and process of authoring with each program was explained to the focus group participants and any comments they made were recorded. A demonstration exercise was conducted by attempting to build a specific application. After the demonstration exercise the discussion proceeded by following the set of questions given in appendix H. The questions were constructed from the conclusions of the analysis of the data from the one-to-one training sessions. Participants were also invited to ask if they were unsure of any terms used or the meaning of the questions. The four programs were assessed in one session. After examining each program the discussion was opened up for debate, mediated by the researcher and using the cues listed below:

- How do the HAPs compare in terms of the usability, utility and learnability?

- What makes knowledge construction/hyperlinking easy?

- How can you make a hypermedia authoring program more usable for HE staff?

- How can you get staff to create at the screen?

- What do you think of the concept of *Transparency of Operation*?

- What do you think of the concept of *Operational Momentum*?

- What do you think of the concept of *Transparency of Purpose*?

3.6.3 The participants of the focus group

Each of the five participants had experience of IT training of HE staff and students. One of the participants was also the person chosen to construct a real application using *SHAPE©* so that his experiences of *SHAPE©* could be shared and contribute to the focus group deliberations.

3.6.4 Data Analysis

The focus group discussions were recorded and transcribed. The transcripts were then imported into NUD.IST and analysed in the same fashion as the qualitative data from the training sessions as described above. Each of the transcripts for each HAP was imported into NUD.IST as a node. Each HAP node was then analysed and data bits were categorised into appropriate categories, which could be pre-existing categories or new ones as appropriate.

3.7 Summary

The objectives this study focus on the issues of hypermedia authoring, and specifically hypermedia authoring by HE teaching staff. The research questions were developed from the objectives and provide the trajectories by which the objectives were achieved. The research questions were developed from the analysis of the current state of the art as depicted in the literature. The hypotheses were derived from the research questions plus an understanding of the dynamics of HCI and knowledge representation and knowledge construction.

To answer the research questions and resolve the hypotheses three main methods were employed, practical one-to-one training sessions, authoring a real application and a focus group. Both qualitative and quantitative analysis was employed in the evaluation of the hypotheses and the research questions.

Figure 3-2 gave an overview of the methodology in relation to the objectives, Figure 3-6 illustrates the final design of the methodology chronologically. Month one in figure 3-6 was approximately September 1995, month 17, the end of the one-to-one training sessions, was January 1997 and the analysis was completed by the end of July 1997. Figure 3-7 illustrates the relationship between the stages of the methodology based on time flowing from left to right.

Section of Thesis	Stage	Month: 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	3.3.2	Design pilot	▓	▓																						
2	3.3.2	Implement pilot																								
3	3.3.2, 3.3.3	Make changes to full study				▓																				
4	3.3.3	Implement full study													▓											
5	3.4	Analyse quantitative results																▓								
6	Chapter 5	Development of a real application								▓																
7	3.3.3, 5.2	Conduct Principle Components Analysis																		▓						
8	Chapter 6	Analyse qualitative results																								
9	3.5.4	Design content analysis																								
10	6.12	Conduct content analysis																						▓		
11	7.2	Conduct desk-based comparison of HAPs																								
12	7.3	Analyse development of a real application																						▓		
13	3.6, 7.4.1	Design and implement focus group																								
15	7.4.2, 7.4.3	Analysis of focus group																								

Figure 3-6 Stages and timescale of the implementation of the methodology.
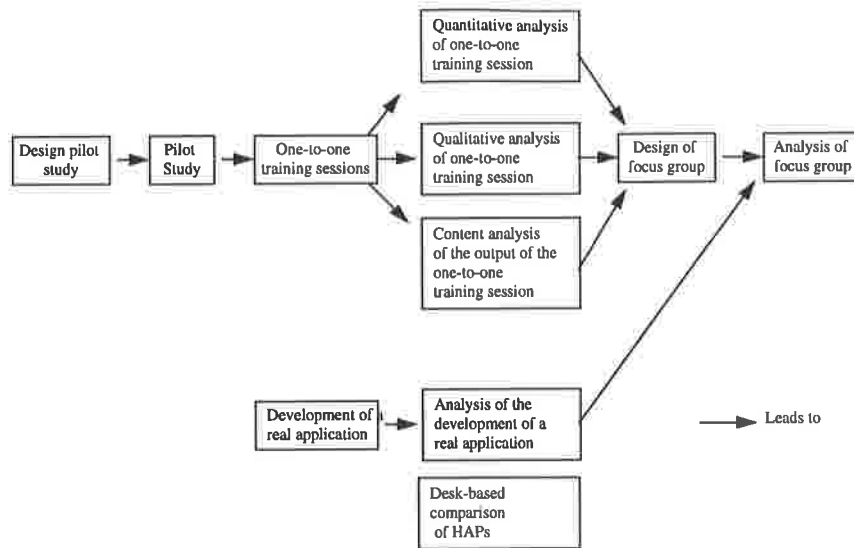
Figure 3-7: Relationship between the stages of the methodology.

Chapter Four

The Design of *SHAPE*©: Semantic Hypermedia Authoring Program for Education.

4. The Design of *SHAPE©*: a Semantic Hypermedia Authoring Program for Education

In chapter two the convergence of knowledge representation and hypermedia was discussed and in section 2.3.4 the case for using visualisation as a vehicle for knowledge representation was presented. In section 2.7 the case for using concept maps as a particular instance of a visual/diagrammatic metaphor for hypermedia authoring was developed. This chapter extends the argument for using concept mapping and presents the design for a prototype HAP called *SHAPE©*.

4.1 The Correspondence of Concept Maps to Hypermedia

Concept maps correspond well with basic hypermedia as illustrated in table 4-1 and figure 4-5 (later in the chapter) and thus provide an appropriate paradigm for hypermedia development. Users could develop their concept maps of a domain which could in principle be converted into the corresponding hypermedia object. However some have argued (see, for example, Halasz, (1991) and Kaplan and Moulthrop, (1994)) that the node and hyperlink model of hypermedia is outmoded and should be supplanted by more sophisticated models of hypermedia which rely on search spaces, generated links and overlapping domains. Concept maps also converge with hypermedia in that they are considered to be examples of tools for knowledge representation and refinement or cognitive tools, as expressed in the NATO Advanced Studies Institute on Cognitive Tools (NATO ASI, 1992).

Heeren (1992) has listed the functionality of stand-alone concept mapping tools shown in Table 4-2, although, not all these functions can be found in one extant tool.

Table 4-1 Mapping concept maps to hypermedia

Concept Map		Hypermedia
Feature	Attribute	
Knowledge Chunk or Concept	Identification	Name of object Name of page Name of composite object Name of window
	Size and Complexity	Whole screen/page Part of screen/page Set of objects Single object Single window Set of windows e.g. child/parent
	Annotation	Material attached to object, window, etc.
Links	Identification	Name of button, hotword, hotregion, pull down menu option
	Direction	Single/Two way button Single/Two way hotword Single/Two way hotregion
	Annotation	Interim displayed content between hyperlinks
Semantic proximity of concepts(represented by thickness of directed arc or spatial proximity of concepts)		Explicit hyperlinks rather than selection from a list

Table 4-2 List of functionality of concept mapping tools (Heeren, 1992)

• Sub-map hierarchical	• Graphics can be attached to concepts
• Submap-zoom	• Text can be attached to relations
• Outliner	• Graphics can be attached to relations
• List of Concepts	• Selective representation of concepts and relations
• List of relations	• Computation and representation of concept centrality
• Logical find function	• Dynamic path presentation
• Fileboxes for organising concepts hierarchically	• Formulate and answer questions (self test)
• 3D representation	• Mask concepts (self test)
• Text can be attached to concepts	• Mask relations(self test)

Some of the features required for educationally-effective hypermedia partially map to Heeren's list of concept mapping tool functionality as shown in Table 4-3. The output from any HAP, including ones based on concept mapping, must ensure that the product is educationally-effective, if it is designed to be used in some sort of teaching and learning setting.

Table 4-3 Partial mapping of requisite feature of educationally-effective hypermedia with the generic functions of concept mapping tools

Requisite Hypermedia Feature	Concept Map Function
Multifaceted, multiperspective views of real world problems	• File boxes for organising concepts hierarchically • Submap-hierarchical • Selective representation of concepts and relations
Active engagement	• Text and Graphics can be attached to concepts • Text and Graphics can be attached to relations
Flexible support for a number of different structures	• Submap-hierarchical • File boxes for organising concepts hierarchically
Good search and query facilities	• List of concepts • List of relations • Logical find function
Support for different learning styles	• Selective representation of concepts and relations
Selective representation of concepts and relations	• Formulate and answer questions (self-test)

It can be seen from table 4-3 that the features requisite for educationally-effective hypermedia are only superficially provided for in currently-available concept mapping tools or HAPs. Of these features, active engagement is probably the most important feature, but the ability to append text and graphics allows only superficial engagement. The provision of multiple perspectives and views of a problem or domain would require the overlaying of concept maps, each representing a different perspective. Hypermedia produced from concept maps must result in applications that accommodate as wide a variety of learning styles as possible. McAleese (1990) suggests that maximising the number of links between the nodes of a hypermedia application facilitates multiple learning styles. Multiple links in the hypermedia material would correspond directly to multiple semantic connections in the corresponding concept map.

So, although concept maps are a possible vehicle for producing hypermedia which could be easy to use and allow the diagrammatic representation of knowledge, with their basic form, it is believed in this study that they lack the expressiveness to embody the features of educationally-effective hypermedia.

4.2 Comparison with Other Hypermedia Authoring Programs

Table 2-7, chapter two shows a comparison of the leading HAPs where it was argued that HAP usability and learnability is at best 'reasonable' - which is not good enough for many HE staff. It is hypothesised in this study that because concept maps are a simple idea, with little formal diagrammatic syntax, a HAP based on them would be easier to learn and easier to use than other HAPs (the null hypothesis tested is that it concept map-based HAPs are no easier to learn and use than other HAPs). Since the inception of this study, interest in concept maps as a design metaphor for hypermedia has grown considerably with at least three HAPs based on concept maps being developed (i.e. CourseWorks (Miller, 1995); Webmapper (Freeman and Ryan, 1997), and CLASS, (McAleese, 1994; CLASS, 1998)). All these map-based HAPs, including *SHAPE©*, derive from Notecards (Halasz *et al.,* 1988). They are all superficially the same but differ in the preferences the designers have given to the functionality that the HAPs should possess. Although for some of these HAPs, particularly Notecards, there has been extensive work on the output from them. As indicated in chapter two there has been little or no usability/learnability evaluation (Nielson, 1990), hence the emphasis in this study on the evaluation of the learnability/usability of map-based HAPs has been a major objective.

4.3 Knowledge Representation

Concept maps can be drawn for literally any domain to any degree of detail. The lack of a constraining syntax means that concept maps can represent many levels of detail and semantically-remote ideas in the same map with considerable ambiguity. This property is both a great advantage (because of the flexibility of expressiveness of the representation from such a simple syntax) and a considerable problem (in terms of - processing the concept map for further use). Bench-Capon (1990) notes the problems of representing a semantic net as production rules in LISP where a concept has many similar but different relationships. The question is whether hypermedia made from concept maps, which have been drawn 'un-rigorously', e.g. mixing levels of meaning, inconsistency in the use of link names, has dubious value? If a concept map represents the creator's view of the domain, is that sufficient or does it lead to the production of 'faulty' hypermedia?

When Gowin and Novak (1984) started using the term 'concept map' as a technique for knowledge representation in a teaching and learning context they emphasised the importance of hierarchy in constructing concept maps even going as far as to say that the mind worked hierarchically. Minsky's (1975) idea of frames resonates with concept maps as a means of knowledge representation and he too stresses the idea of hierarchy. Levesque and Mylopoulos (1979) also appreciated the importance of hierarchy when they introduced the idea of classes into semantic nets. From a pragmatic or even aesthetic point of view representing domains hierarchically is desirable because of the problem of clutter. In complex domains with many relations and concepts, representation in the same plane becomes 'messy'. Deconstructing a

domain hierarchically resolves the clutter problem although it also results in the loss of a single view of the domain (see figure 4-5).

Another important question is whether concept maps are able to represent all knowledge domains that are likely to be used in a HE teaching and learning setting? Jonassen (1993) notes that concept maps are useful for representing 'structural' knowledge, i.e. the declarative aspects of a domain the question is whether they are capable of representing procedural aspects? In a straightforward way, procedural knowledge can be represented by a string of concepts in which one concept points to the next. At the domain-specific level, the issue is whether concept maps can represent subjects like languages, mathematics, etc., sufficiently easily and expressively to be useful? McAleese (1987 and 1986) notes the problems users have in expressing their knowledge of a domain with concept maps and then reconciling the various views generated by a map-based HAP and the original internal understanding. Taking into account, all these issues may mean that again there is a need to supplement basic concept maps to make them useful to HE teaching staff.

4.4 Functionality - Instructional Design and Pedagogical Provision. versus Syntactic Constraint

Following on from the discussion above is the question of how much functionality can be derived from a concept map. As noted above concept maps can represent structure but in their simplest form they probably cannot embody instructional design and may thus have lower pedagogical value. However, if the syntax of concept maps can be supplemented to carry 'meta-information' like 'parametric instructional design objects', then there is a danger of them losing their simplicity and hence their

usability. Figures 4-1(a) and 4-1(b) show the trade offs between usability of a tool versus its expressiveness as a function of the level of constraint of the HAPs. Programming languages are highly unconstrained and allow great expressiveness and flexibility, their disadvantage is that they are not very usable for most users. At the other extreme there are authoring templates or wizards which are very easy to use but are not very flexible and hence have more limited expressiveness, e.g. NEAT (Mayer *et al*, 1993). Commercial HAPs, such as ToolBook© and Authorware©, fall somewhere in between, although it is argued in this study that they still do not have a higher usability rating from the perspective of an average HE teacher. What is desirable is a HAP that is both expressive and flexible and also *very* easy to use.
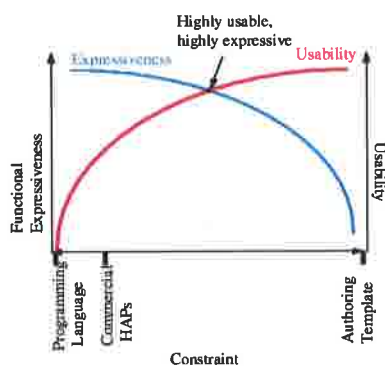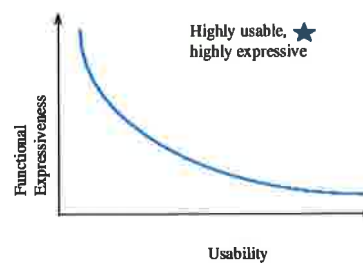
(a)   (b)

Figure 4-1(a) The degree of constraint of a HAP versus its expressiveness.[6] Figure 4-1(b). The corollary of 4-1(a) is a plot of expressiveness versus usability which shows that usability and expressiveness are inversely related.

---

[6] Expressiveness is a measure of its ability to produce educationally effective-hypermedia. The usability is measured from the perspective of HE teachers, although the curve is likely to vary depending on skill particularly Computer Skill. The graphs depicted are only illustrative and are not based on empirical data.

The star on figure 4-1(b) represents the ideal HAP with higher functional-expressiveness and high usability.

4.5 Design of *SHAPE©*

The design of **SHAPE©** takes into account the above discussion although compromises on some of the principles mentioned were inevitable. The high-level requirements specification for **SHAPE©** was to:

- Enable hypermedia to be created from concept maps in a two stage process where in stage one a concept map is developed and in stage two the concept map is post processed and supplemented with content.
- Embody the principle of hierarchy of knowledge representation.
- Use un-constrained un-typed concepts and links.
- Wherever possible embody pedagogical principles.

4.5.1 The Syntax of the Concept Maps

Concept Types

It was decided, again for reasons of simplicity and usability, that there would be a minimum number of concept types in the **SHAPE©** prototype. One could create many classes of concept as desired. However it was considered that this would lead to a barrier to usage since users would be subjected to the higher cognitive overhead of understanding the meaning of each type of concept. In this sense there was no concept classes and users could label as they wished - these concepts were called *atomic*.

There was a need to fulfil the requirement, as discussed in section 4.3, that the concept maps should be hierarchical, none of the other HAPs based on concept maps (Freeman and Ryan, 1997; Miller, 1995) existing allowed for hierarchical representation. For this reason, a *composite* concept was defined as one which subsumed a number of subordinate concepts, i.e. a parent concept. An example might be the concept 'family' which subsumes the subordinate concepts 'mother', 'father', 'sister' etc. Composite concepts could have other composite concepts as subordinates, for example the concept 'family' might have 'uncles' as a subordinate which is itself a composite and so one. The decision of whether a concept was *composite* or *atomic* was left to its users and was a function of the degree to which users wished to deconstruct a subject and the level of detail they wished to explore. There was no limit placed on the hierarchy of composite concepts, users could deconstruct a domain top-down as far as they wished, in accordance with the ideas discussed in section 4.3 above.

Link Types

The chosen syntax for concept maps in *SHAPE*© was straightforward. In terms of links there was no directionality apart from that implied in the link name. The reason for deciding on no directionality was because all semantic relations are reciprocal as illustrated in figure 4-2 below. In figure 4-2(a) biology has subcategories whereas in figure 4-2(b) Human and Animal belongs to Biology. In a hypermedia, generated from these concept maps, the relations become hyperlinks and there seems little pedagogical reason for not allowing the user to return to last node, i.e. either from the Biology node to the Human node or vice versa.

Biology, subcategory, has subcategory, Animal, has subcategory, Human

Biology, Belong to, Animal, Belong to, Human

(a)   (b)
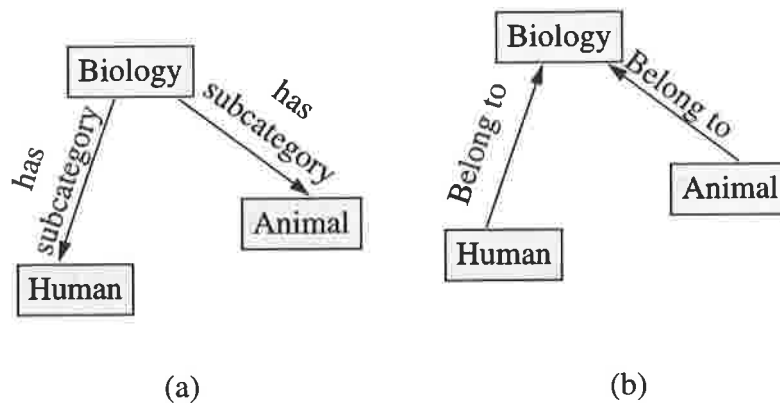
Figure 4-2 Reciprocity of relationships in concept maps

It was decided that there would be no link types other than the free format labels so that the resultant HAP was as unconstraining as possible. This decision contrasts with earlier concept mapping tools which advocated the use of typing (McAleese, 1987; Kozma, 1992; Reader, 1995), but is supported by others (Fisher, 1992). Users would be allowed to label their links with whatever text they wished even no link name. This is very interesting point to consider, why should someone want to link two concepts but not explicitly label the link? The reasons might be either laziness or simply an inability to articulate the name of a particular semantic relation. McAleese (1986) identifies this problem as anomalous state metacognition. Another reason might be that the link label might unnecessarily limit the rich meaning behind the relationship. The inability might be a function of the relationship itself in that there is no convenient phrase or 'sound bite' that encapsulates the relationship.

4.5.2 Database

In order to accommodate the requirements for concepts and links above a database was designed. Figure 4-3 shows the design of the database tables.

Links Table

The links table includes seven fields as indicated in figure 4-3. The label field refers to the text a user enters for the name of that link. The source and destination concepts are the two concepts that the user wishes to link. The screen objects refer to the items which represent the link on the screen, this information is important when one of the concepts associated with the link is moved. The status refers to whether the link has already been post-processed to create a skeletal hypermedia. The 'parent page id' refers to the screen page object that holds the link.

Concepts Table

Most of these fields have similar meanings as for links, but in addition, the class field indicates whether the concept is *composite* or *atomic*.

Figure 4-3 The *SHAPE*© database.

4.5.3 Mapping to Hypermedia

Once the concept maps have been drawn they need to be converted into hypermedia. Section 4.1 discussed the correspondence between hypermedia and concept maps. For the purposes of the prototype version of **SHAPE**© it was decided that each concept would map to a group of objects that would form a hypermedia node. In later versions of **SHAPE**© it might be possible to determine what sort of content objects a concept should map to but it was felt that this would introduce expressiveness to **SHAPE**© that might be a distraction for users at this stage. It was decided that links would simply map to a hyperlink at this stage even though provision for some annotation of objects was included in the database design. It was decided that each concept, whether it was a *composite* or *atomic*, would map to a hypermedia content node.

The hyperlinks represented were simply hot words to a content-page. Other media to media links were omitted at this stage. In the prototype linking of concepts was only possible between concepts in the same parent page. It was recognised that it would be desirable to be able to link to any concept wherever it was in the hierarchy of concept maps.

Figures 4-4(a) and 4-4(b) show the mapping between hypermedia and concept maps and figure 4-5 shows the resultant hypermedia.

(a)   (b)
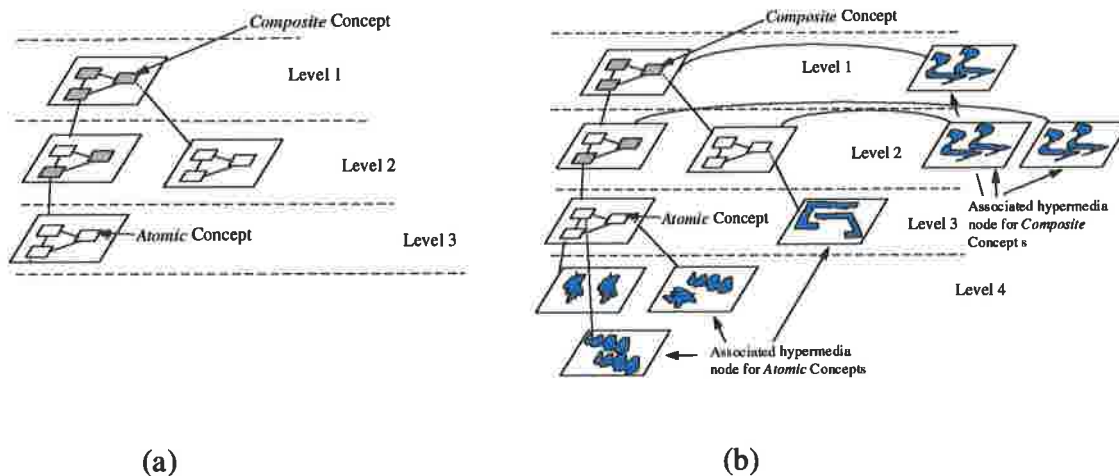
Figure 4-4 The hierarchy of concept maps before mapping and after mapping to hypermedia nodes and hyperlinks.

Level 1

Level 2

Level 3

Level 4

Hierarchy of concept maps

Corpus of hypermedia nodes with hyperlinks
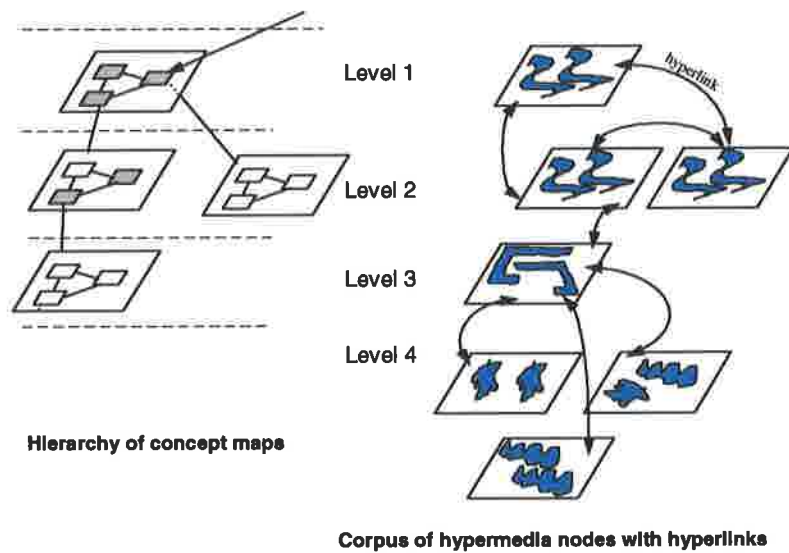
Figure 4-5 The hypermedia corpus generated from the hierarchical concept map

Templates and Annotations

Each concept maps on to a collection of screen objects which are instances of a template as shown in figure 4-6 below. The word 'instance' is used in the same sense as it is in object-oriented programming. Provision was made in the database to allow users to select from or create their own templates. Templates can be considered to be the basic design of contents of each node. The prototype implementation used only one template. The corpus of the instantiated template pages (the hypermedia nodes) formed a hypermedia book in which the semantic links of the concept map become hot word hyperlinks between each hypermedia node as illustrated in figure 4-6.
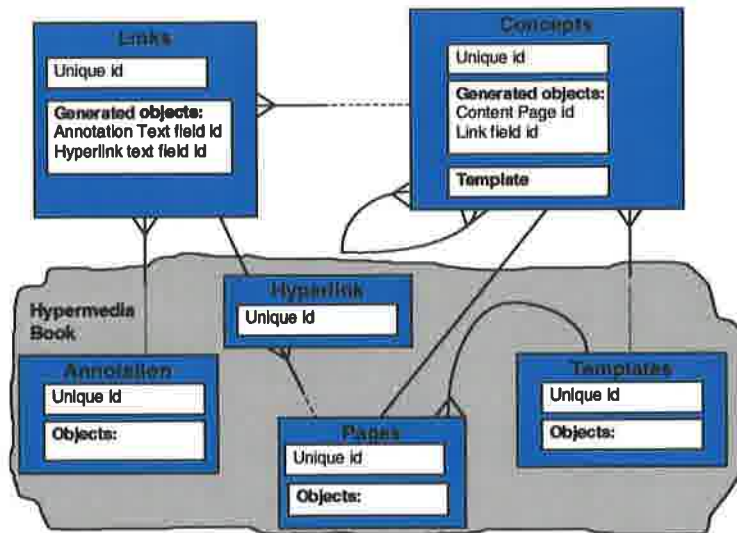
Figure 4-6 Supplementary fields to the database after mapping to hypermedia. Diagram utilises the 'crows feet' (CCTA, 1990) syntax for representing entity-relationships.

Annotation of links was not implemented in the *SHAPE*© prototype but was seen as an essential component in a full-scale implementation of *SHAPE*©. When users selected a link there should be opportunity to explain the link before jumping to the selected hypermedia node.

4.5.4 Implementation

SHAPE© was built using Asymetrix ToolBook© and produces skeletal ToolBook© 'books'. *SHAPE*© was built using an existing HAP rather than a language like C++ or Visual Basic because its facility for creating hyperlinks and integrating multimedia could be called from *SHAPE*©. ToolBook© was selected to base *SHAPE*© on in 1994 when it appeared to be the *de facto* standard for developing educational multimedia in

the UK. Since then, the Web has become an important delivery medium and it is unlikely the same decision would be made today.

The database of links and concepts means that concept maps could be implemented in any environment that supports hypermedia. including the Web.

4.5.5 The Operation of *SHAPE*©.

Figure 4-7 shows ***SHAPE***© being used to develop a concept map of the subject 'IT'. Concepts are input via the concept input box at any time during the authoring session. Concepts can be moved wherever its user wishes. Links are added by using a simple keystroke sequence which selects the concepts to be linked. Once the concepts have been linked as desired it is possible to make some concepts *composite* with a simple keystroke as indicated by the shading of the concept 'technology' in figure 4-8. *Composite* concepts allow subordinate concepts in a subordinate concept map to be created. Concepts in the subordinate map can also be changed into *composite* concepts, hence the knowledge domain in question can be hierarchically deconstructed in a progressive fashion.
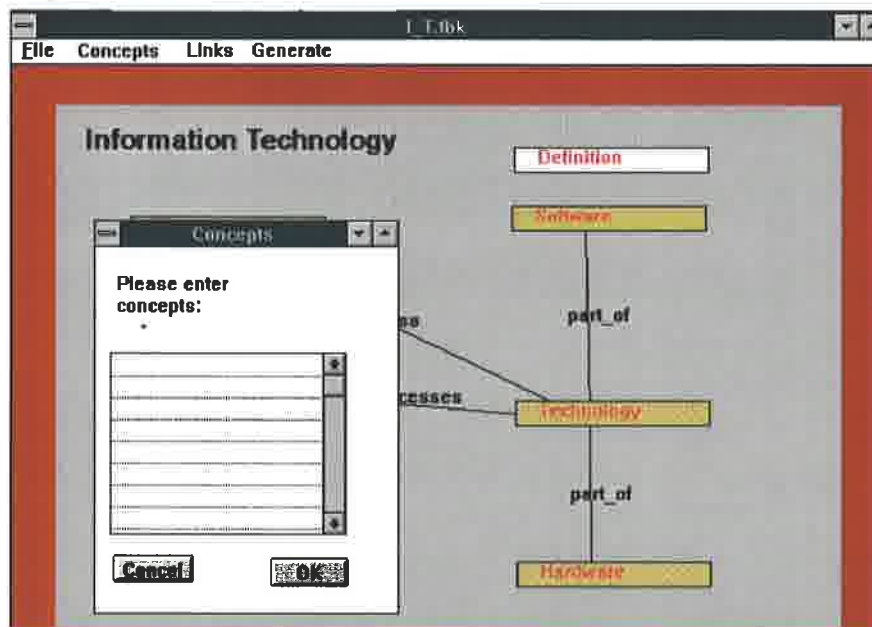
Figure 4-7 *SHAPE*© being used to develop a concept map for the subject-domain 'IT'

The top-level concept map is then processed to create the skeletal ToolBook© book. Each concept name is converted into a hotword that leads user to a newly-generated template (a ToolBook© page). Figure 4-8 below shows the atomic page for 'Definition'. The composite page for 'technology' is a blank second-level concept map with a hyperlink back to the top-level concept map. The author subsequently adds the content into the page for each concept.
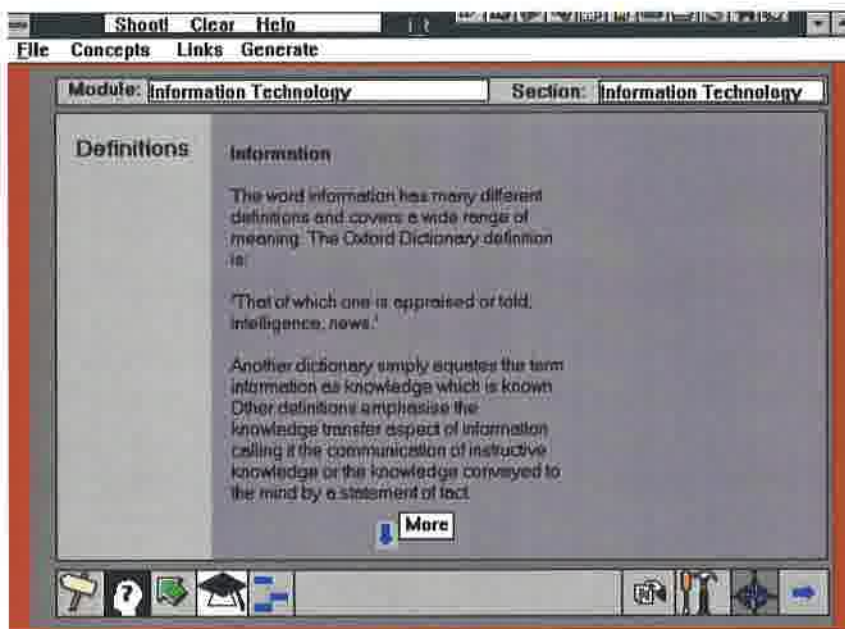
Figure 4-8 Atomic page for a concept called 'Definition' in *SHAPE*©

The semantic links between concepts are manifest as a list of optional hyperlinks to other pages as shown in figure 4-9 below for the concept 'Printers'.

The final product is a ToolBook© 'book' comprising of concept maps and corresponding content pages with hyperlinks corresponding to the semantic links from the concept maps. The examples shown here are from an application called **CLICK-IT**©, the development of which is described in chapter seven.

Hyperlinks
generated from
links in the
concept map.
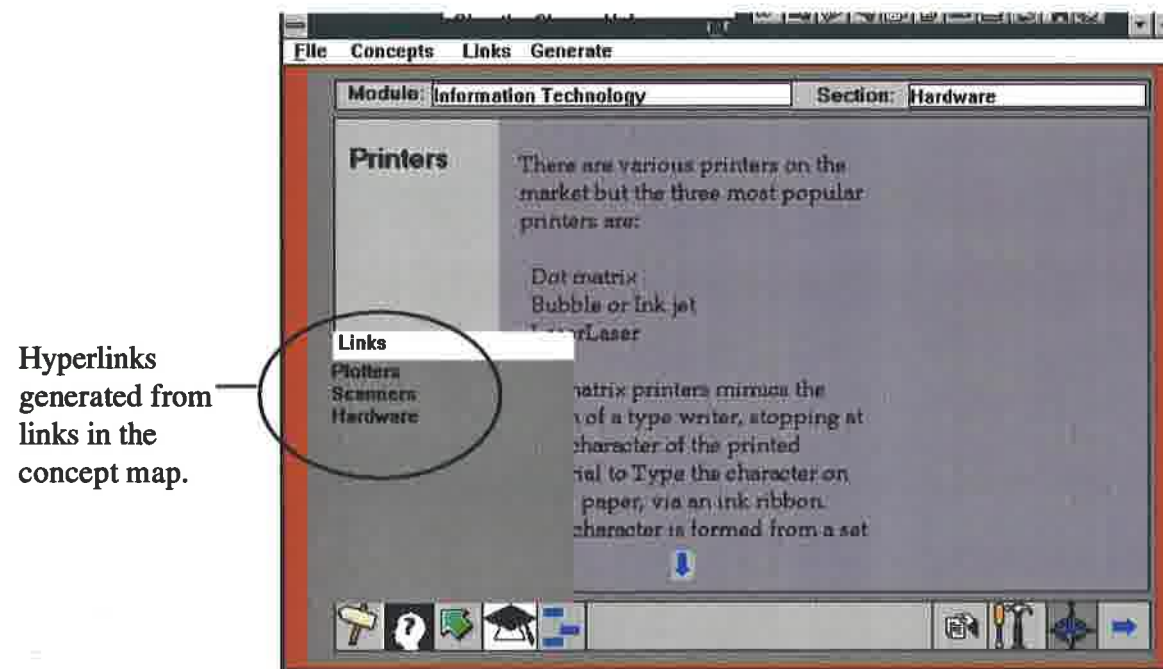
Figure 4-9 Hyperlinks in a page generated from relations in a concept map created in *SHAPE*©.

4.6 Conclusions

SHAPE© is a working prototype based on the idea of using concept maps as metaphors for hypermedia authoring. *SHAPE*© has been used to build two hypermedia applications for teaching, one for Strategic Management and the other for basic IT. The development of the material for teaching IT called **CLICK-IT**© formed part of the methodology of this study, **CLICK-IT**© and its development is analysed in chapter seven.

Chapter Five

Quantitative Analysis of the Use of TBM and *SHAPE*©

5. Quantitative Analysis of the Use of TBM and *SHAPE*©

5.1 Analysis of the Independent and Dependent Variables

In this section the results of the one-to-one training sessions are presented through analysis of each of the dependent variables and their interaction with the independent variables. The raw data with summary values are detailed in appendix B. The section concludes with an evaluation of the hypotheses established in the methodology.

5.1.1 Reliability and Validity

A Cronbach Alpha test was carried out on the EL, EU and Task-Match data, the results are shown in table 5-1 below, to check their reliability.

Table 5-1 Reliability measures of the dependent variables using the Cronbach Alpha test

Scale-Variable	Cronbach Alpha
EL - *SHAPE*©	0.92
EL - TBM	0.91
Ease of Using - *SHAPE*©	0.92
Ease of Using - TBM	0.92
Task Match *SHAPE*©	0.35
Task-Match TBM	0.66

All scales appear to be reliable except Task-Match *SHAPE*© which has a low Cronbach Alpha reading. The Task-Match was measured using only two items which has a large effect on the reliability of a scale so a high Cronbach Alpha is more

unlikely therefore the Cronbach Alpha for *SHAPE*© is deemed to be acceptable in this context.

The dependent and independent variables were checked for normality as shown in table 5-2 using the Shapiro-Wilks and K-S Lilliefors tests for normality. Where the probability is greater than 0.05 a normal distribution can be assumed. In this study if either of the two tests used showed a probability of less than 0.05 a non-normal distribution was assumed.

Table 5-2 Normality tests for the independent and dependent variables used in the study.

Variable	Shapiro-Wilks		K-S (Lilliefors)	
	Value	Prob	Value	Prob
Age	0.945	0.442	0.117	>0.200
CSI	0.954	0.523	0.117	>0.200
Spatial Relations	0.982	0.958	0.079	>0.200
Ease:				
EL - *SHAPE*[©]	0.909	0.123	0.117	>0.200
EL -TBM	0.934	0.346	0.095	>0.200
EU *SHAPE*[©]	0.945	0.440	0.126	>0.200
EU TBM	0.945	0.443	0.121	>0.200
Task-Match *SHAPE*[©]	0.830	<0.010	0.324	0.000
Task-Match TBM	0.922	<0.244	0.165	>0.200
Motivation:				
Extrinsic Motivation	0.932	0.328	0.131	>0.200
Intrinsic Motivation	0.975	0.879	0.090	>0.200
Challenge	0.964	0.698	0.152	>0.200
Compensation	0.950	0.487	0.126	>0.200
Enjoyment	0.964	0.706	0.122	>0.200
Outward	0.914	0.172	0.157	>0.200
Motivation to Continue with *SHAPE*[©]	0.833	<0.010	0.194	0.108
Motivation to Continue with TBM	0.903	0.092	0.152	>0.200
Prior Experience of Concept Maps	0.796	<0.010	0.337	0.000
Output:				
Number of concepts - *SHAPE*[©]	0.766	<0.010	0.185	0.144
Number of Pages - TBM	0.744	<0.010	0.258	0.005
Number of Links - *SHAPE*[©]	0.860	0.017	0.253	0.007
Number of Links - TBM	0.935	0.351	0.201	0.083
Understanding:				
Pre session - *SHAPE*	0.741	<0.010	0.321	0.000
Pre session - TBM	0.820	<0.010	0.271	0.003
Post session - *SHAPE*[©]	0.806	<0.010	0.204	0.074
Post session - TBM	0.907	0.109	0.134	>0.200

Many of the dependent variables were derived from four or five point Likert scales so the possibility of the data being normally distributed is low. This is the case with many of the dependent variables in this study, as indicated in table 5-2. In cases where the variables could not be classified as normally distributed non-parametric tests were used. Since the EL, EU and Task-Match scales were the aggregation of several items, they were treated as interval data in the analysis.

5.1.2 Ease of Use (EU), Ease of Learning (EL)

The items in the computerised question bank for measuring EU and EL were aggregated by adding the score for each item, taking into account whether the item was negatively or positively weighted. A paired sample t-test showed that the EL and EU of *SHAPE*© were significantly higher than for TBM as shown in table 5-3 below.

Table 5-3 Paired sample t-test comparing EU *SHAPE*© EU TBM. EL *SHAPE*© with EL TBM.

Variable	Mean	SD	t value	Significance (2-tailed)
EU - *SHAPE*©	43.750	5.459	3.61	0.003
EU - TBM	36.875	6.672		
EL - *SHAPE*©	39.562	5.240	3.44	0.004
EL - TBM	32.812	5.492		

Correlation between EL and EU

The EU and EL for *SHAPE*© and for TBM correlated significantly with each other as shown in table 5-4 below showing that there is probably a relationship between the EL and EU. There was no correlation between the EU of *SHAPE*© and the EU of TBM and no correlation between the EL of *SHAPE*© and the EL of TBM shown in table 5-4.

Table 5-4 Correlations between EL and EU for each HAP and also between HAPs

Variable	r^2	Significance (2 tailed)
EU and EL of *SHAPE*©	0.891	0.000
EU and EL of TBM	0.749	0.001
EU *SHAPE*© and EU TBM	0.222	0.408
EL *SHAPE*© and EL TBM	-0.068	0.803

5.1.3 Understanding

Subjects were asked to rate their understanding of the two HAPs before using them but after a standard explanation was given and then again after using the HAPs. Table 5-5 shows the results of a non-parametric Wilcoxon test. The tests indicate that subjects found *SHAPE*© more understandable even before the one-to-one training sessions and remained that way after the session.

Table 5-5 (a) and (b). Wilcoxon matched-pairs signed-ranks test, comparing subjects' rating of their understanding of *SHAPE*© and TBM prior to and after the one-to-one training sessions.

Variable	Median	Range	Z	Significance (2 tailed)
Declared understanding of *SHAPE*© prior to use	3	3	4.03	0.006
Declared understanding of TBM prior to use	2	3		

(a)

Variable	Median	Range	Z	Significance (2 tailed)
Declared understanding of *SHAPE*© after use	3.5	2	4.27	0.0028
Declared understanding of TBM after use	3	3		

(b)

There was a significant increase in the understanding of both HAPs after each session as shown in table 5-6.

Table 5-6 Wilcoxon matched-pairs signed-ranks test, comparing the difference in understanding of each HAP after training session.

Variable	Median	Range	Z	Significance (2 tailed)
Declared understanding of *SHAPE*© prior to use	3	3	-3.283	0.017
Declared understanding of *SHAPE*© after use	3.5	2		
Declared understanding of TBM prior to use	2	3	-2.716	0.007
Declared understanding of TBM after use	3	3		

There was no significant difference in the increase in understanding between each HAP as shown in table 5-7 below. The value for relative increase in understanding were calculated by subtracting the rating of understanding prior to the training session from the rating of understanding after the training session for each subject.

Table 5-7 Wilcoxon matched-pairs signed-ranks test, comparing the difference in the increase of understanding of each HAP after training session.

Variable	Median	Range	Z	Significance (2 tailed)
Relative increase in understanding of *SHAPE*© post-training session	0.5	3	-0.340	0.713
Relative increase in understanding of TBM post-training session	0.5	2		

Figure 5-1 illustrates the finding in table 5-7 that there was no relative change in understanding of each HAP before and after the one-to-one training sessions.
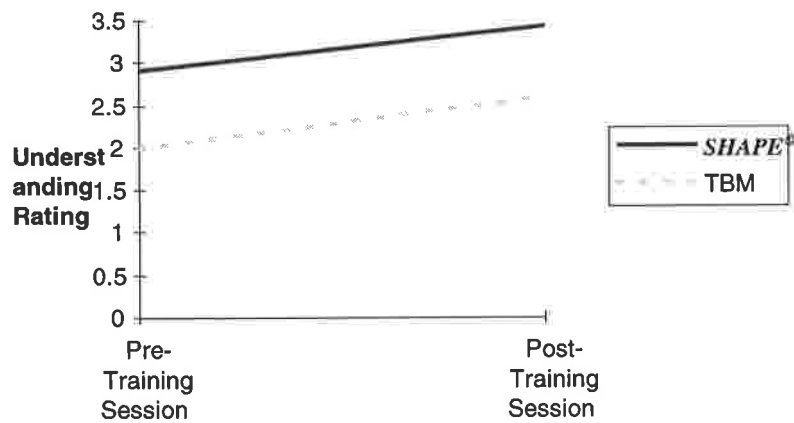
Figure 5-1 Comparison of ratings of understanding pre-and post-training sessions

5.1.4 Output

The number of concepts and semantic links created in *SHAPE*[©] were compared with the number of pages and hyperlinks created in TBM. Table 5-8 (a), (b) and (c) below show the results of the t-test for independent samples. Results were recalculated taking in to account the order in which each HAP was used as shown in tables 5-9, 5-10 and 5-11. Tables 5-8(a), 5-8(b) and 5-8(c) show that the number of concepts/pages created with *SHAPE*[©] was significantly higher than with TBM - regardless of the Order of Use. Tables 5-9, 5-10 and 5-11 indicate the same result for links created.

Table 5-8(a) Mann-Whitney U test for all cases, comparing the number of concepts/pages created for each HAP

Variable	Number of Cases	Median of number of concepts/pages	Range	Mann-Whitney (U, Z-Value)	Significance (2 tailed)
SHAPE[©]	16	16.5	51	38.00, -3.401	0.0007
TBM	16	5.5	23		

Table 5-8(b) Mann-Whitney U test comparing the number of concepts/pages created cases where the HAP was used first.

Variable	Number of Cases	Median of concepts/pages	Range	Mann-Whitney (U, Value)	Significance (2 tailed)
SHAPE[©]	8	18.5	46	9.5	0.018
TBM	8	7	22		

Table 5-8(c) Mann-Whitney U test comparing the number of concepts/pages created cases where the HAP was used second.

Variable	Number of Cases	Median of concepts/pages	Range	Mann-Whitney (U Value)	Significance (2 tailed)
SHAPE©	8	13.5	20	8	0.011
TBM	8	5	13		

Table 5-9 Mann-Whitney U test comparing the number of links created for all cases for each HAP.

Variable	Number of Cases	Median of number of links	Range	Mann-Whitney U, Z-Value	Significance (2 tailed)
SHAPE©	16	6.5	17	(33.00, -3.61)	0.000
TBM	16	2.5	5		

Table 5-10 Mann-Whitney U test comparing number of links created where HAPs were used first for each HAP.

Variable	Number of Cases	Median of links	Range	Mann-Whitney (U Value)	Significance (2 tailed)
SHAPE©	8	7	17	7	0.008
TBM	8	3.5	5		

Table 5-11 Mann-Whitney U test comparing number of links created where HAPs were used second.

Variable	Number of Cases	Median of links	Range	Mann-Whitney (U Value)	Significance (2 tailed)
SHAPE©	8	6	6	7	0.007
TBM	8	2	3		

5.1.5 Task-Match

The Task-Match scores were found from summing the scores for Items 8 and 9 of the EU scale found in appendix C. The measured values of subjects' perceived Task-

Match were not normally distributed so the non-parametric Wilcoxon matched Signed-Ranks test was used instead. The results are shown in table 5-12 and indicate that there was no significant difference in Task-Match.

Table 5-12 Wilcoxon matched-pairs signed-ranks test, comparing the Task-Match of each HAP measured after training sessions.

Variable	Median	Range	Z	Significance (2 tailed)
Task-Match *SHAPE*©	6	3	-1.778	0.075
Task-Match TBM	6	4		

5.1.6 Computer Skill

As indicated in table 5-13, Computer Skill was correlated with EL, EU, number of concepts and pages created and number of links created with each HAP. In terms of learning the two HAPs, there was a significant correlation between the EL of TBM and Computer Skill but not the EL of *SHAPE*©. In terms of using the two HAPs the opposite was true, there was a significant correlation between the EU of *SHAPE*© and Computer Skill but not the EU of TBM, although at the 90% significance level the EU of TBM also correlated with Computer Skill. Figure 5-2 below shows the scatter diagram of Computer Skill and EU for both HAPs together with regression lines significant at the 90% level.

Table 5-13 Pearson correlation of Computer Skill with dependent variables

CSI correlated with:	r^2	Significance
EL - *SHAPE*©	0.399	0.160
EL - TBM	0.568	0.022
EU - *SHAPE*©	0.501	0.048
EU - TBM	0.458	0.074
No of Concepts - *SHAPE*©	0.043	0.875
No of Pages - TBM	0.395	0.130
No of links - *SHAPE*©	0.177	0.527
No of links - TBM	0.006	0.982
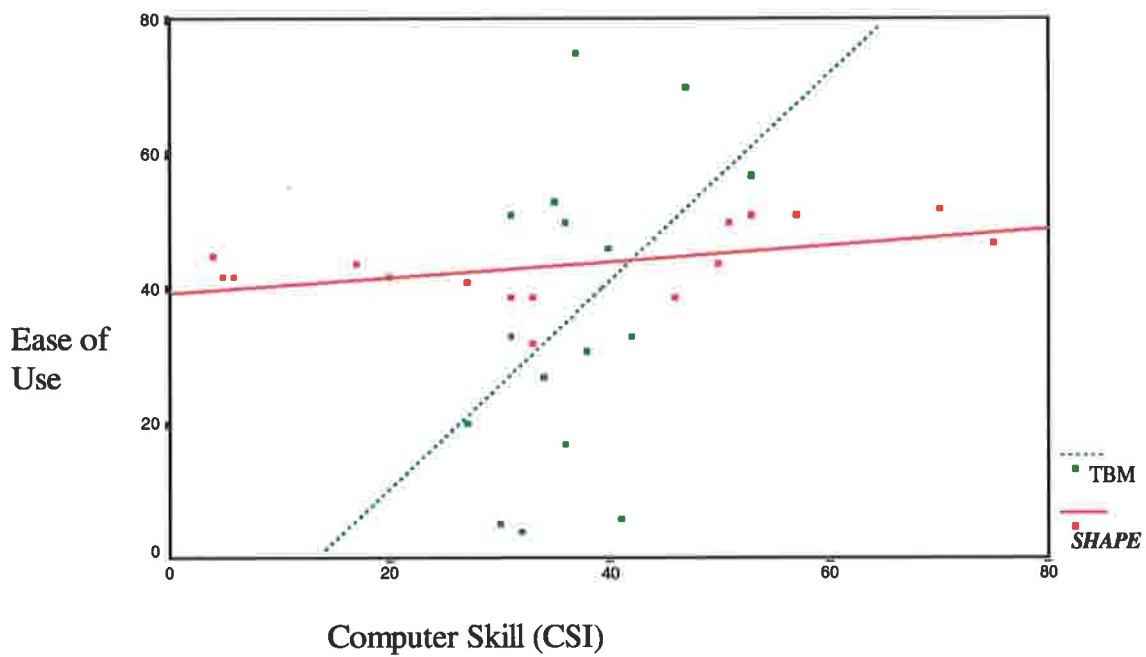
Figure 5-2 Scatter diagram of EU with Computer Skill for each HAP, together with regression lines (P<0.1).

5.1.7 Order of Use

The order in which each HAP was used was seen as a potential cause of interaction in the measurement of the dependent variables. It was therefore necessary, to measure the effect of the Order of Use by using multivariate analysis. The measurement of EL, EU, understanding, output, Motivation to Continue and Task-Match were retested - taking into account the Order of Use. Where the variable was normally distributed a one way ANOVA was used, where the variables were not normally distributed, a Kruskal-Wallis one-way ANOVA was used. Table 5-14 shows the results of the one way ANOVA test and tables 5-15 and 5-16 shows the results of the Kruskal-Wallis one-way ANOVA. The conclusion from tables 5-14, 5-15 and 5-16 is that the Order of Use has little significant effect on the differences in the dependent variables for each HAP. The only effect Order of Use had was to depress the EL rating of TBM when it was used after *SHAPE*© as shown in figure 5-3.

Table 5-14 One way ANOVA to measure effect of the Order of Use on EL and EU of each HAP

Variable	n	Order 1= TBM/*SHAPE*© 2= *SHAPE*©/TBM	Mean	SD	F Ratio	F Probability
EL -*SHAPE*©	8	1	38.250	6.497	1.004	0.333
	8	2	40.875	3.563		
EL -TBM	8	1	36.250	2.053	10.050	0.007
	8	2	29.375	5.780		
EU *SHAPE*©	8	1	42.500	6.024	0.829	0.378
	8	2	45.000	4.899		
EU TBM	8	1	39.125	4.486	1.933	0.186
	8	2	34.625	7.981		

Table 5-15 Kruskal-Wallis one-way ANOVA to measure effect of the Order of Use on dependent variables of each HAP

Variable	n	Order 1=TBM/*SHAPE*© 2=*SHAPE*©/TBM	Median	Range	χ^2	Significance
Number of Concepts - *SHAPE*©	8	1	13.5	20	2.019	0.155
	8	2	18.5	46		
Number of pages -TBM	8	1	7	22	2.371	0.124
	8	2	5	13		
Number of Links -*SHAPE*©	8	1	6	6	1.354	0.244
	8	2	7	17		
Number of Links -TBM	8	1	3.5	5	1.310	0.252
	8	2	2	3		
Task-Match - *SHAPE*©	8	1	6	2	1.681	0.195
	8	2	6.5	2		
Task-Match -TBM	8	1	6	2	2.972	0.85
	8	2	5	4		

Table 5-16 Kruskal-Wallis one-way ANOVA (for ordinal variables) to measure effect of the Order of Use on Motivation to Continue and Understanding of each HAP.

Variable	n	Order 1= TBM/*SHAPE*© 2= *SHAPE*©/TBM	Median	Range	χ^2	Significance
Motivation to Continue - *SHAPE*©	8	1	3.50	1.00	0.003	0.956
	8	2	3.50	2.00		
Motivation to Continue - TBM	8	1	3.25	2.00	2.283	0.131
	8	2	2.50	3.00		
Pre-session understanding - *SHAPE*©	8	1	3.00	1.00	1.801	0.179
	8	2	3.00	3.00		
Pre-session understanding -TBM	8	1	2.50	3.00	0.800	0.371
	8	2	1.50	2.00		
Post-session understanding - *SHAPE*©	8	1	3.75	2.00	0.080	0.777
	8	2	3.25	1.00		
Post-session understanding - TBM	8	1	3.00	2.00	1.526	0.217
	8	2	2.75	2.50		
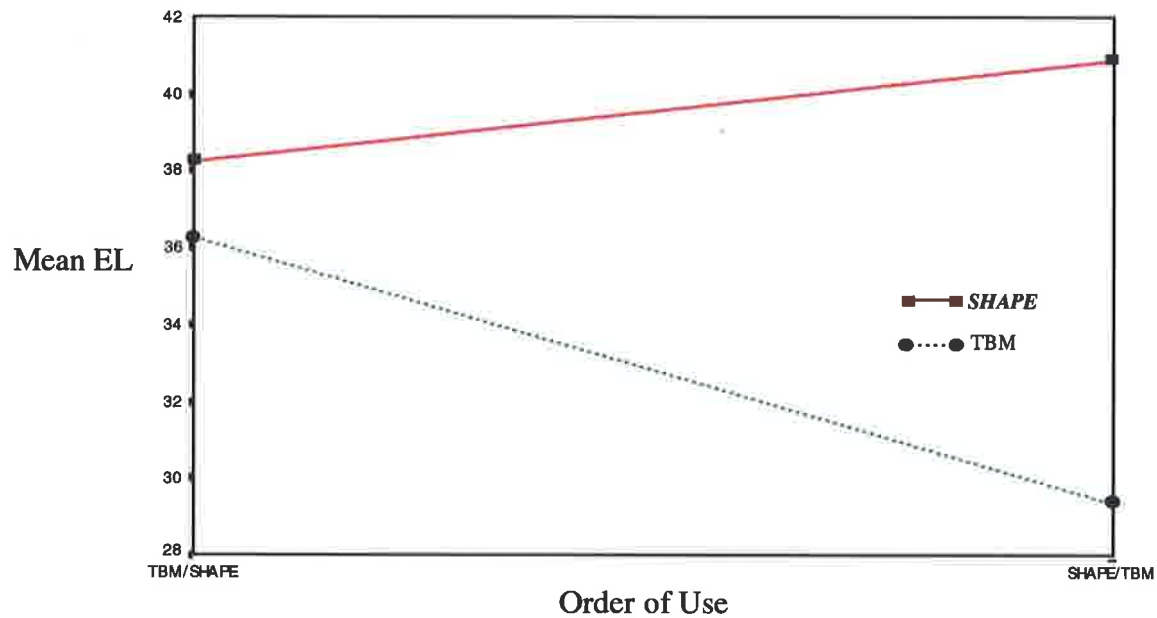
Figure 5-3 Mean EL plotted against Order of Use.

5.1.8 Spatial Relations

Table 5-17 shows the correlations of subjects' Spatial Relations scores with their ratings of EL, EU of each HAP, the output variables and the understanding variables. Spatial Relations ability has no apparent effect on EL or use and no effect on the number of pages and links created in TBM. However, there is a significant correlation with the number of concepts and links created in *SHAPE*©. Figure 5-4 shows the scatter diagram of Spatial Relations ability versus number of links and concepts with the regression lines superimposed ($P<0.1$). Further analysis showed that there could be an effect created by the Order of Use but larger samples would be required to prove this.

Table 5-17 Pearson's correlation of Spatial Relations correlated with EL and EU and output variables.

SR correlated with:	r^2	Significance
EL - *SHAPE*©	0.280	0.293
EL - TBM	0.108	0.690
EU - *SHAPE*©	0.197	0.464
EU - TBM	0.221	0.412
No of Concepts - *SHAPE*©	0.494	0.052
No of Pages -TBM	0.347	0.187
No of links - *SHAPE*©	0.654	0.006
No of links - TBM	-0.672	0.805
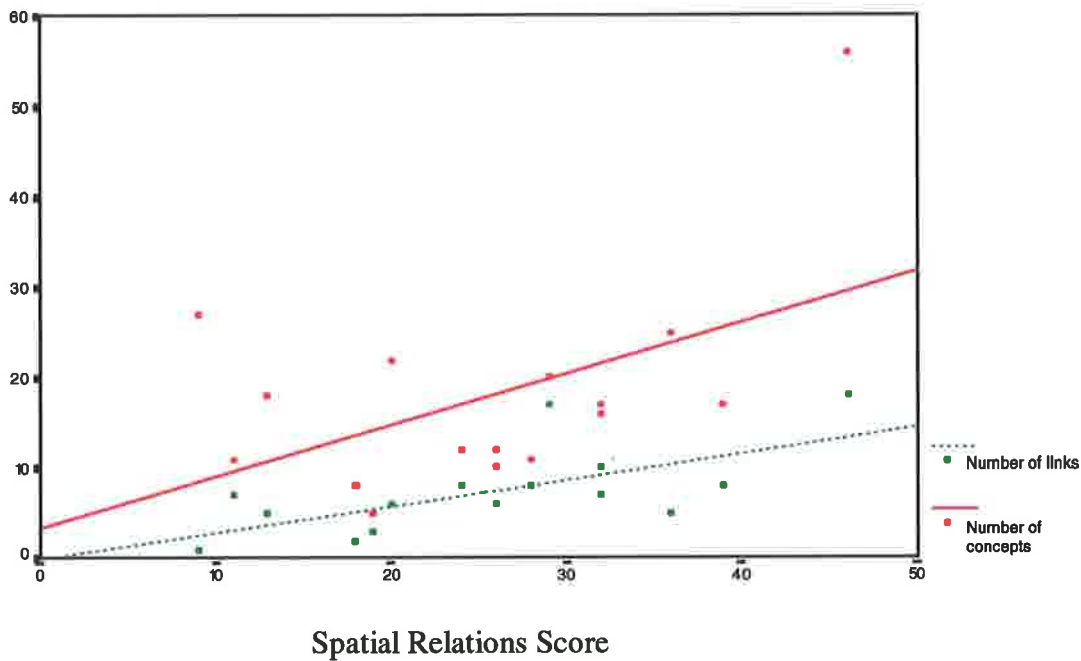
Figure 5-4 Scatter diagram of Spatial Relations ability and number of concepts and Links created with *SHAPE*©. with regression lines (P<0.1).

5.1.9 Motivation

Subjects' Motivation to Continue with each HAP was compared. The results are given in table 5-18. Subjects' work preference inventory scores were correlated with the ease variables, the understanding variables and the output variables, shown in table 5-19. The results presented in table 5-19 tend to indicate that neither subjects' intrinsic nor their extrinsic motivation had any bearing on the dependent variables of this study. The one possible exception is the correlation between EU of *SHAPE*© and intrinsic motivation and between Motivation to Continue with *SHAPE*© and intrinsic motivation with 90% significant positive correlations.

Table 5-18 Wilcoxon matched pairs signed ranks test comparing Motivation to Continue for each HAP.

Variable	Median	Range	Z	Significance (2 tailed)
Motivation to Continue with *SHAPE*©	3.50	2.00	-2.49	0.013
Motivation to Continue with TBM	3.00	3.00		

Table 5-19 Pearson correlations between dependent variables and Intrinsic and Extrinsic Motivation

Variable	Intrinsic Motivation		Extrinsic Motivation	
	r^2	P	r^2	P
EL - *SHAPE*©	0.478	0.061	0.045	0.869
EL - TBM	-0.028	0.992	0.089	0.742
EU - *SHAPE*©	0.389	0.137	-0.100	0.711
EU - TBM	0.023	0.932	0.393	0.132
No of Concepts - *SHAPE*©	0.147	0.593	0.150	0.580
No of Pages - TBM	0.135	0.619	0.275	0.303
No of links - *SHAPE*©	0.133	0.623	-0.122	0.652
No of links - TBM	-0.058	0.832	-0.352	0.181
Motivation to Continue with *SHAPE*©	0.433	0.094	0.226	0.399
Motivation to Continue with TBM	0.177	0.511	0.235	0.382

5.1.10 Miscellaneous Results

There are several miscellaneous results which need to be considered, these are presented in table 5-20. Analysis of gender was omitted because of the limited number of females in the sample. Table 5-20 indicate that age was not a confounding factor of the main effects, although interestingly, the older the subject the more concepts and pages they created. Prior Experience of Concept Maps had an effect on the EL and EU of *SHAPE*©.

Table 5-20 Correlation between Age (Pearson) and dependent variables and between Prior Experience of Concept Maps (Kendall tau-b) and dependent variables.

	Age		Prior Experience of Concept Maps	
	r^2	P	r^2	P
EL - *SHAPE*©	-0.380	0.147	0.358	0.09
EL - TBM	0.050	0.854	0.062	0.766
EU - *SHAPE*©	-0.325	0.219	0.438	0.037
EU - TBM	-0.169	0.531	-0.020	0.921
No of Concepts - *SHAPE*©	-0.432	0.095	-0.010	0.961
No of Pages - TBM	-0.507	0.045	0.284	0.178
No of links - *SHAPE*©	-0.298	0.262	0.198	0.345
No of links - TBM	0.290	0.277	-0.203	0.356
Motivation to Continue with *SHAPE*©	-0.227	0.399	0.201	0.373
Motivation to Continue with TBM	-0.004	0.988	0.044	0.839

5.2 Principle Components Analysis

Exploratory PCA of the EL and EU scales of each HAP was undertaken. The Principle Components Analysis was conducted on the set of data collected for each HAP and using Kaiser's criterion (Bryman and Cramer, 1997), i.e. factors with Eigen values of one or greater were retained. PCA was followed by a varimax rotation of the raw factors and the results of this can be found in appendix I. As noted in section 3.3.3 the use of PCA with such a small sample could lead to a dubious analysis, however a number of factors mitigate against this outcome. It can be seen from Appendix I that the values of Bartlett's Test of Sphericity indicated that using PCA on the data EL of *SHAPE*©, the EL of TBM, the EU *SHAPE*© and the EU of TBM was appropriate. The Kaiser-Meyer-Olkin measure of sampling adequacy for each of the data sets is above 0.5, which for such small sample sizes is a good value. Since similar groupings of factors emerge from the two EL and EU data sets, the data sets

can be combined (Tabachnick and Fidell, 1996). When the data sets are combined, the same grouping of factors emerge but with a Kaiser-Meyer-Olkin measure of sampling adequacy of 0.8 which is greater than the 0.6 recommended by Tabachnick and Fidell (1996).

The items of each factor derived from the analysis of each HAP were compared to find common factors and common items. This analysis is documented in appendix J and results are shown in table 5-21.

Table 5-21 Common items of factors between *SHAPE*© and TBM .

Factors	Scale Items ('+' positive) ('-' negative)	Item Descriptors (X is either *SHAPE*© or TBM)
EL Factors:		
1	+2	I liked learning 'X'
	+1	'X' was easy to learn
	+4	I felt comfortable learning 'X'
	-10	It took too much time to learn 'X'
2	-6	The ideas behind 'X' were difficult to appreciate
	-9	I often became confused learning 'X'
	-10	It took too much time to learn 'X'
	-5	I found 'X' difficult to understand
	+7	When 'X' was explained it was obvious what to do
3	+3	I gained a lot learning 'X'
	+8	'X' is no more difficult than other Windows based programs
EU Factors:	-	
1	+24	I felt comfortable using 'X'
	-14	I liked using 'X'
	-13	-It was easy to use 'X'
	+26	It took too much time to use 'X'
2	+18	It was obvious what to do next
	+16	The set of operations one needed to use were easy to remember
	+15	I had no difficulty understanding how to use 'X'
	+26	It took too much time to use 'X'
3	-22	I really felt I had accomplished something using 'X'
	-23	I felt frustrated using 'X'
	+25	'X' was fun

Items 20 and 21 relate to Task-Match and were omitted from PCA. The factors shown in table 5-22 were identified from the item descriptors in table 5-21, the exact process of identifying the factors is subjective in nature so corroborating evidence was sought in the qualitative analysis which is given in chapter six. The identification of two of factors have some resonance with the factors discovered in the PCA of the scale from which the scales used in this study were derived (Molnar and Kleke, 1996). Molnar and Kleke's factors were 'attitude towards the process' and 'attitude towards the software', i.e. Accommodation and Transparency respectively, hence providing some validity to their existence as factors independent of this study.

Table 5-22 Identification of the common factors of EL and EU

Factors	Identification
EL Factors:	
1	Accommodation
2	Transparency
3	Accomplishment
EU Factors:	
1	Accommodation
2	Transparency
3	Accomplishment

The symmetry between the EL and EU factors is striking but not unlikely since other results have shown that they appear to be closely related. The duplication of the factors from EL and EU scales provides for their validity and substantiates their existence. The meaning of the factors are discussed in greater detail in chapter six together with the corroborating qualitative results.

5.3 Hypotheses

The hypotheses are laid out in section 3.3.3 In this discussion the main hypothesis is considered last because it is the summation of the other hypotheses.

5.3.1 Hypothesis H_1: Task-Match

The Wilcoxon test analysis given in table 5-12 indicates that there was no significant difference in the perceived Task-Match between *SHAPE*© and TBM hence the null hypothesis was accepted.

5.3.2 Hypothesis H$_2$: Computer Skill and Ease of Learning

The null hypothesis was that there would be no correlation between the EL for

SHAPE© and TBM and Computer Skill measured by CSI. The Pearson correlations

values given in table 5-13 indicate that there are significant correlations between EL-

TBM and CSI scores but not between EL-*SHAPE*© and CSI scores. Hence hypothesis

two is partially supported. The lack of a correlation between EL-*SHAPE*© and CSI

score is an important deviation from hypothesis two and indicates that although the

mean EL of *SHAPE*© is higher than that for TBM, it is independent of Computer

Skill, i.e. subjects found *SHAPE*© easier to learn irrespective of their computer skill.

5.3.3 Hypothesis H$_3$: Computer Skill and Ease of Use

The null hypothesis was that there would be no correlation between the EU for

SHAPE© and TBM and Computer Skill measured by CSI. The Pearson correlations

values given in table 5-13 indicate that there are significant correlations between EU -

SHAPE© , EU - TBM. Hence the null hypothesis is rejected and the experimental

hypothesis accepted. However, this discovery is partially mitigated by the regression

lines for the correlations for EU and Computer Skill found in figure 5-2 indicating that

although there is a significant correlation between EU and Computer Skill, it is almost

flat for *SHAPE*© (i.e. the effect of Computer Skill is significant but weak) but very

steep for TBM (i.e. the effect of Computer Skill is an important determinant in the

perceived EU of TBM).

5.3.4 Hypothesis H_4: Order of Use

In most cases the Order of Use had no effect on dependent variables (see tables 5-15, 5-16 and 5-17). More specifically, the EL and EU of *SHAPE*© and the EU of TBM were unaffected by the Order of Use. However, subjects who used *SHAPE*© first, rated the EL of TBM significantly lower that when they used *SHAPE*© second pointing to an unanticipated phenomenon. Analysis of the transcripts indicated that, instead of subjects being better prepared in terms of their mental picture and understanding of the faculty when using *SHAPE*© first, subjects became subsequently more frustrated when learning TBM, regardless of whether they already had a mental picture of the faculty.

Hence evaluating the sub-hypotheses:

Hypothesis H_{4a}: Order of Use and EL

The null hypothesis is rejected (see table 5-14) and the experimental hypothesis accepted, although the direction of the effect of Order of Use was opposite to that anticipated.

Hypothesis H_{4b}: Order of Use and Ease of Use

The null hypothesis is accepted (see table 5-14) and the experimental hypothesis rejected, the Order of Use had no effect on the EU ratings of either HAPs.

5.3.5 Hypothesis H_5: Spatial Relations and EL

Table 5-17 indicates that there was no correlation between Spatial Relations and EL hence the null hypothesis is accepted and the experimental hypothesis rejected.

5.3.6 Hypothesis H_6: Spatial Relations and EU

Table 5-17 indicates that there was no correlation between Spatial Relations and EU hence the null hypothesis is accepted and the experimental hypothesis rejected.

Despite the rejection of the experimental hypothesis of the effect of Spatial Relations ability on EL and EU, there was a significant positive correlation between the output from the use of *SHAPE*© and Spatial Relations ability. Showing that although Spatial Relations ability did not affect the subjects opinions of either HAP it did affect their productivity.

5.3.7 Hypothesis H_7: Prior Experience of Concept Maps and EL *SHAPE*©

Subjects' Prior Experience of Concept Maps was positively correlated with their rating of EL *SHAPE*© as given in table 5-20. The null hypothesis is rejected and the experimental hypothesis accepted.

5.3.8 Hypothesis H_8: Prior Experience of Concept Maps and EU *SHAPE*©

Subjects' Prior Experience of Concept Maps was positively correlated with their rating of EU *SHAPE*© as given in table 5-20. The null hypothesis is rejected and the experimental hypothesis accepted.

5.3.9 Hypothesis H₉: Motivation to Continue

Subjects were more motivated to continue with *SHAPE*© than with TBM (see table 5-18) hence the null hypothesis is rejected and the experimental hypothesis accepted. The effect of the subjects' own base motivation as measured by the Work Preference Inventory (Amabile *et al.,* 1994) on EL and <u>EU of each HAP was</u> insignificant (see table 5-19).

5.3.10 Main Hypothesis

In aggregating the results of the nine hypotheses discussed above and the other results presented in this chapter, it can be concluded that subjects preferred and were facilitated in learning and using the HAP based on concept maps - *SHAPE*©. The main null hypothesis is therefore rejected, although this result is confounded by the subjects' Computer Skill, Spatial Relations ability and Previous Use of Concept Maps as indicated in the evaluation of the 9 sub-hypotheses above. The Order of Use of each HAP did not have much impact on the overall result.

5.4 Conclusions

This chapter has presented the results of the quantitative analysis of the one-to-one training sessions. The dependent variables have been compared for the two HAPs in relation to the independent variables and the hypotheses related to the objectives of the study evaluated. Based on this analysis the main hypothesis of this study has been shown to be true - map-based HAPs can facilitate staff in developing hypermedia. The Principle Components Analysis on the EL and EU scales have indicated the

existence of three emergent properties potentially important in learnability and usability.

Chapter Six

Qualitative Analysis of the Use of TBM and *SHAPE*©

6. Qualitative Analysis of the Use of TBM and *SHAPE*[©]

There is no prescribed way of presenting qualitative results, partly because the account forms a main part of the analysis. Compared with a statistical test where the test itself is the analysis, reporting the account of the QDA forms a major part of the analysis (Dey, 1993; Coffey and Atkinson, 1996). QDA often involves the coding of data into categories which are synthesised into patterns to make a coherent picture. It is important that the emerging picture is placed in the context of the state of the art from the literature. Dey (1993) even likens QDA to telling a story. The qualitative results from this study are therefore presented as a discussion in this chapter. Due to the lack of examples of qualitative analysis of the usability of computer programs there is no precedent for reporting it here. The chosen method of presentation was to present the categories identified in NUD.IST, followed by an analysis of how they interact and then a discussion on how this work relates to other work cited in the literature.

The validity and reliability of the QDA was considered in the methodology described in chapter three. However the importance of the 'triangulation' of results is an accepted part of this study. Therefore the QDA presented in the previous chapter, the QDA described in this chapter and the analysis of the focus group study (presented in the next chapter) have, wherever possible, been compared and contrasted.

6.1 Results of the Principle Components Analysis

The qualitative data supported the defining of a number of factors with some data bits loading negatively onto the factors, some loading positively as shown in the record of the evidence in appendices L, N and O. A considerable body of evidence was also found to support the concept of *Transparency*, discovered in the Principle Components Analysis of chapter five, and indeed, the word is used by two of the subjects while another uses the term 'obviousness'. The Principle Components Analysis pointed to the possibility of *Transparency* being subdivided into *Transparency of Operation* and *Transparency of Purpose* and there was clear evidence to support the existence of these two new sub-factors in the qualitative data.

The concept *Transparency of Operation* is defined as:

> *The summative aspects of the HAP that allowed users to find, understand and then use rapidly and easily the functions of the HAP to achieve a task or sub-task.*

The definition of *Transparency of Purpose* is:

> *the summative aspects of the HAP that at any point during its use gives users an 'image' of the end-product of the HAP.*

There was more evidence to support *Transparency of Operation* than *Transparency of Purpose*.

The factors related to learning are considered in section 6.11 below.

6.2 Usability Factors

Usability is considered in detail in the literature review (see section 2.4.3 chapter two). Shackel's (1986) four defining criteria of usability (effectiveness, learnability, flexibility and attitude) are widely accepted, although they appear a little crude now. Kellog (1987) and Tullis (1986) have introduced the idea of 'consistency' as an attribute of usability and there are references to more general, inclusive concepts such as 'ease of use', which were considered in chapter two, (Eason and Damordaran, 1981). Sutcliffe (1995) uses the idea of 'utility' which is similar to Davis' (1989) 'perceived usefulness'.

Shackel's (1986) four defining criteria, although true, do not give any clues to what it is about a program that will satisfy these criteria - only that if a program or interface possesses these attributes it will be usable. For example, how is a 'flexible' interface achieved? Instilling the right attitude in the user comes as a consequence of having these attributes in place but how can one actively imbue the right attitude when using a program? Similarly how does one instil in the user the perception that a program has a high utility (Sutcliffe, 1995) and an increased perception of usefulness (Davis, 1989)? Green's (1996) cognitive dimensions framework (considered in more detail in section 2.4.3, chapter two), provides a more useful agenda/tool for attending to the attributes listed above. Green's cognitive dimensions map to Shackel's criteria as indicated in figure 6-1. It is evident that the cognitive dimensions largely contribute to Shackel's flexibility criterion. It is broadly true that improved flexibility should improve effectiveness and learnability (Roberts and Moran, 1983; Whiteside *et al.*, 1985). However, increased flexibility can make programs more complex and, based

on the results of this study, for certain user types (particularly those with lower Computing Skills) may make the programs less effective and less learnable. In this context it is not surprising that the cognitive dimensions tend to correspond to more than one of Shackel's criteria. It will also be noticed that 'attitude' does not correspond to any of the cognitive dimensions - this is because cognitive dimensions address aspects of the program not the user.
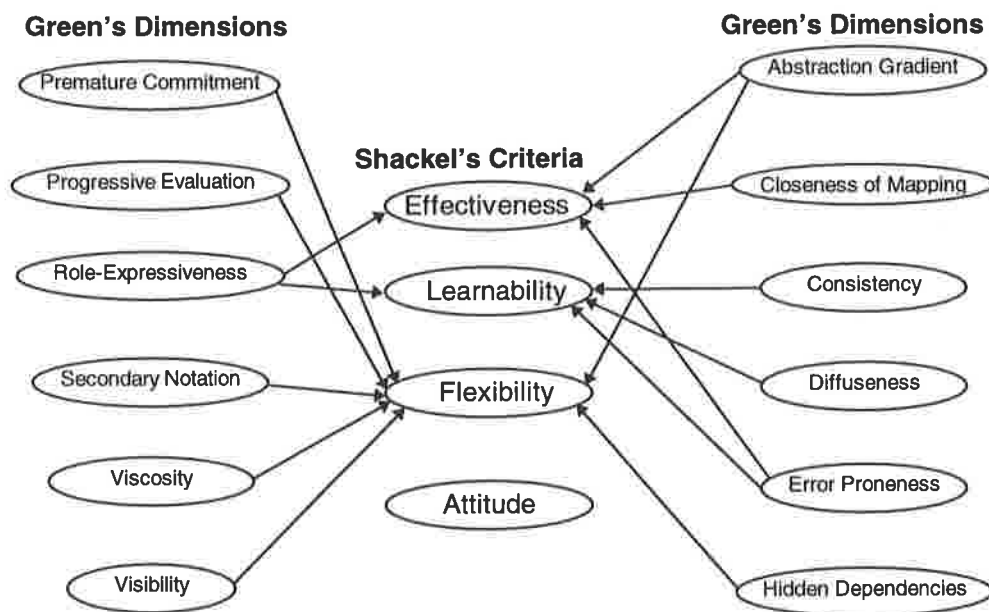
Figure 6-1 Mapping Green's (1989) cognitive dimensions to Shackel's (1986) criteria.

Dix *et al.* (1998) present a framework for usability based on the attributes of a system which are broadly in line with Green's cognitive dimensions framework. The problem with Dix *et al's* framework is that it does not seem to have any empirical evidence to support its existence. For this reason the Dix *et al's* framework was disregarded in this analysis but reference to this framework will be made in the discussion of chapter seven.

Green's framework has more resonance with the factors discovered in this study due to the nature and objectives of this research, i.e. the underlying factors affecting the authorability of hypermedia, and because of the rather general nature of the other factors found in the literature. The following sections investigate the factors discovered in this work and their parity with those in the literature.

6.3 Transparency of Operation

The elements of *Transparency of Operation* are discussed below and were identified from the data bits. A complete list of the evidential data bits for each element is given in appendix K, but table 6-1 provides some examples of evidence. *Transparency of Operation* is an holistic property subject to the elements identified below. The concept of *Transparency of Operation* does not appear to have much correspondence with Shackel's criteria of usability. However, a HAP that has the property of being highly transparent will improve its effectiveness and probably its learnability. Seely Brown (1986) considers a concept called *Opacity* which is essentially the opposite of the same concept. Green's dimensions correspond better with the elements of *Transparency of Operation* because they themselves are elements that contribute to the property of *Transparency of Operation*. Section 6.10 below discusses the interplay of the sub-factors of *Transparency of Operation*.

Table 6-1 Examples of evidence defining the sub factors of *Transparency of Operation* taken from the transcripts of the one-to-one training sessions and categorised in NUD.IST.

	Examples of Evidential Data Bits
Operational Momentum	"What I quite liked was when I typed the words in it came up automatically with a series of boxes - you could see what was going on. I wasn't quite sure how I was going to move them around but I felt I had to move them around so I could see the next logical stage." "It's straightforward it's knowing what to do next"
Logic of Operation	"Straightforward to follow" "the confusion I had was the operations - what I had to press"
Noise/ Economy of dialogue	"I understand the concept but I don't like the screens, I don't like the way that it says I am on page 3, what's page 3? what's page 1 and 2?"
Mental Model Match	"It's the way I think I suppose when I'm creating structure - I like to keep the structure in my mind and this is a structure I'm creating" "I automatically wanted to structure it with main headings and subheadings but nothing allowed me to do that."
External Consistency	"Consistent with other programs"
Internal Consistency	"A program that once you've learned the fundamentals, the bells and whistles follow along the same pattern"

6.3.1 Operational Momentum

Operational Momentum is defined as the summative aspects of the HAP that 'leads the user on' to the next stage, sometimes iteratively if need be. There does not seem to be a direct correspondence between *Operational Momentum* with any listed factors although there is some cross-over. Green's dimensions of *Hidden Dependence, Premature Commitment, Viscosity* and *Visibility* undoubtedly contribute to creating *Operational Momentum*. Any *Hidden Dependencies* will reduce *Operational Momentum* because changes will occur unknown to users, who will not know why certain events occurred due their actions. *Premature Commitment* can result in

backtracking to remedy the situation. *Viscosity* results in slow progress when changes are required. The lack of *Visibility* means users cannot see where they are and where they are going in the context of the task. In the same sense, the factors defined below also contribute to *Operational Momentum*. Comments by subjects that the fear of getting lost or remembering what to do next were significant. Where these fears are mitigated subjects felt that the HAP 'didn't seem to interrupt the thinking'. A comment was made that *'you had to use the HAP* [TBM] *more'* which is effectively a comment on Green's *Viscosity* dimension. In addition to the attributes indicated by Green's dimensions, there are other functions of a program that contribute more directly to *Operational Momentum* . The comment made by the subject in table 6-1 above indicates a HAP characteristic which does indeed 'carry the user on'. *Operational Momentum* probably contributes to Shackel's criterion of effectiveness however *Operational Momentum* does not necessarily increase flexibility since the more flexible a HAP, the less momentum is generated. Increased *Operational Momentum* could lead to increased prescription in a HAP, i.e. to improve *Operational Momentum* one might reduce the flexibility available to the user.

The data bits that contributed to the formulation of the concept of *Operational Momentum* were positively loaded with respect to *SHAPE*© and negatively loaded with respect to TBM. It would be fair to conclude that *SHAPE*© has a greater degree of *Operational Momentum*, which was also a result of differences in the *Logic of Operation, Noise/Economy of Dialogue*, sub-factors as described below.

6.3.2 Logic of Operation

Logic of Operation is defined as the degree to which the completion of tasks and sub-tasks whilst using the HAP have an internal logic so that users can almost guess what action they need to take next. This quality is obviously affected by a user's mental model of how the HAPs should work. There does not seem to be a good correspondence between *Logic of Operation* and any of Shackel's criteria, although it is probably true that the *Logic of Operation* will affect learnability and a HAP that has clear logic in its operations will be more effective. To some extent *Logic of Operation* is a function of the mental model a user has of the HAPs. However in a particular cultural context, e.g. western democracy with its mathematical and philosophical tradition, there will be a 'received' sense of logic used in designing HAPs which is likely to be shared by the users of the HAP. Green's dimensions of *Abstraction Gradient*, *Closeness of Mapping* and *Role-expressiveness* affect the perceived logicality of the program.

There was some evidence that TBM was less logical in its operation.

6.3.3 Noise/Economy of Dialogue

Noise/Economy of Dialogue is defined as the perceived level of unnecessary information or complexity in completing a task or sub-task (evidential data bits are found in appendix K). This factor corresponds to Kozma's (1992) conclusion that the complexity of tool must be kept to a minimum. Again there is no direct correspondence with Shackel's criteria, only that HAPs with minimal noise and economic dialogues will improve learnability and effectiveness. *Noise/Economy of*

Dialogue partially corresponds to Greens' *Abstraction Gradient*, although increased abstraction while increasing the power of the program may also confuse the user. *Diffuseness* also partially corresponds to *Noise/Economy of Dialogue*. A HAP which has minimal noise or diffuseness will probably result in less mistakes being made, i.e. Green's *Error-Proneness*.

Subjects found TBM very 'noisy' but did not comment on the noisiness of *SHAPE*©.

6.3.4 Mental Model Match

Mental Model Match is defined as the degree to which the HAPs can accommodate a range of user models of the HAPs. Much has been said about the importance of mental models as indicated in section 2.4.1. A HAP designed with a very clear idea of who the users are should increase its learnability. Mental model match corroborates McAleese's (1985) claim that authoring tools need to adapt to a variety of authors. Getting the right level of Green's abstraction gradient and closeness of mapping relies on a good understanding of the users of the system. The evidence found in this study was that the level of computing skill was seen to affect the understanding of how the HAPs 'should' operate (see appendix L). To design a HAP so that it responds to a range of generic mental models seems very difficult. Responding to levels of computing skill seems both obvious and achievable but responding to a whole gamut of user parameters, e.g. learning style, culture, class, motivation, is probably unrealistic.

Subjects found **SHAPE**© to be a better match with how they expected the HAPs to work than TBM. This is true despite the familiarity of the book metaphor. Subjects liked the diagrammatic representation of knowledge and would have preferred TBM to have offered this facility. It could be that diagrammatic representations are becoming more common as mechanisms for interacting with computer programs. The philosophy underpinning concept maps is the natural preference of the human brain to remember images (Buzan, 1993) and therefore work with them. The preference for a diagrammatic interface is a key difference between **SHAPE**© and TBM and is considered again in the discussion of chapter seven.

6.3.5 External Consistency

External Consistency is defined as the degree of consistency of the HAPs with other programs. Typically this means how well the HAPs conforms to standard user interfaces like Windows or X Windows-based applications. This factor 'fuels' the mental model of a user who will have different levels of experience of the end-interface software. Green also uses the term *Consistency* but it is in the context of *Internal Consistency* as discussed below. *External Consistency* will influence some of the features of *Internal Consistency*, for example the file menu in Windows-based applications. There is little that can be done in terms of *External Consistency* apart from ensuring the software adheres to the 'norms' of the end-user interface.

Subjects made no comment about the *External Consistency* of *SHAPE*© and TBM but they did recognise the importance that *External Consistency* had on the usability and learnability of a HAP.

6.3.6 Internal Consistency

Internal Consistency is defined as how consistent the various functions/aspects of the HAPs are with each other, e.g. that buttons with the same name do the same or similar thing. As noted in the preceding section, *Internal Consistency* does to some extent rely on *External Consistency*. Green's concept of consistency is the same as *Internal Consistency* defined here. It is widely accepted (e.g. Kellog, 1987; Tullis, 1986) that *Internal Consistency* is a good program attribute for improving its *Transparency of Operation*. No differences were identified between *SHAPE*© and TBM.

6.3.7 Interaction of Sub-Factors

Transparency of Operation is the emergent property of the interplay and interaction of the identified factors of this study and other factors like those described by Green (1996). *Transparency of Operation* is the property of a HAP that 'lets' users complete their tasks rapidly and easily - as defined in section 6.3 above. 'Good' *Transparency of Operation* of a HAP comes about because the HAP has 'good' *Operational Momentum* and *Logic of Operation*. 'Good' *Operational Momentum* and *Logic of Operation* are partially achieved by ensuring the HAP is consistent internally and externally, the dialogue with users is economical - but not to the point where they need to maintain some prior knowledge and the amount of information on

the screen at any point is just sufficient. 'Good' *Transparency of Operation* also depends upon the users' models of the HAP - as discussed in the section on *Mental Model Match* above. Figure 6-2 shows the causal relationship between factors suggested from the analysis of the factors.
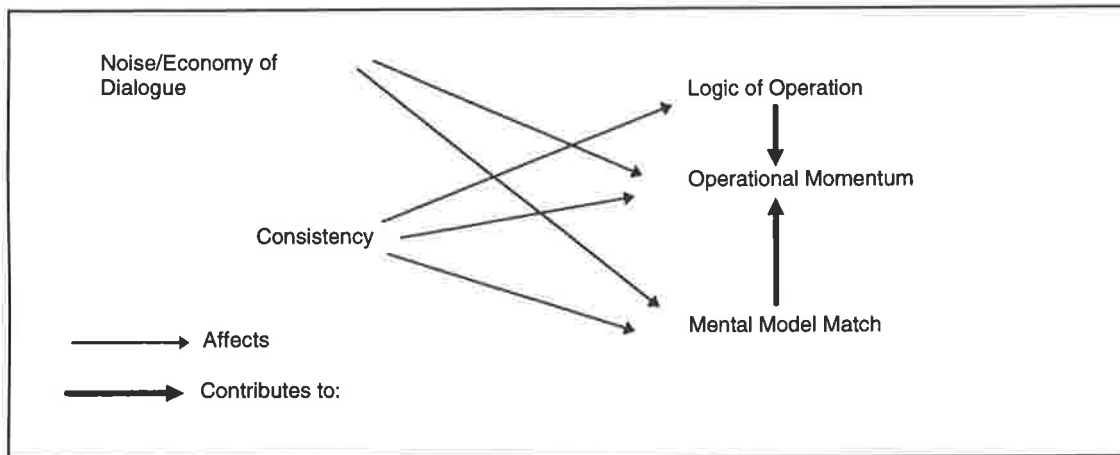
Figure 6-2 Interaction of sub-factors

6.4 Transparency of Purpose

Transparency of Purpose is defined in section 6.4 above. Table 6-2 gives examples of evidential data bits that corroborate the concept of *Transparency of Purpose*. A complete listing of the evidence can be found in appendix M. Two sub-factors were also identified as part of *Transparency of Purpose, Task Match* and *Instantaneity*. *Task Match* is an accepted concept (Eason, 1984), however no similar concept of *Instantaneity* was found in the literature.

Table 6-2 Examples of evidential data bits for the sub-factors of *Transparency of Purpose*

Element	Examples of Evidence
Task Match	"The result of the prototype didn't give me what I wanted"
Instantaneity	"It doesn't seem to offer so much so quickly as *SHAPE*©" "It got there eventually"

Sutcliffe's idea of 'utility' (1995) and Davis' (1989) similar idea of 'perceived usefulness' correspond to the sub-factor *Task Match* rather than *Transparency of Purpose*. Utility is more a combination of *Transparency of Operation* and *Task Match*. *Perceived Usefulness* is a little closer to the idea of *Transparency of Purpose* but does not convey its full meaning. Green's *Progressive Evaluation* dimension maps to *Transparency of Purpose* to some extent. Programs which allow *Progressive Evaluation* means that users have a clearer idea of how well the program is realising the end-product. *Transparency of Purpose* is a function of the *Task Match* of the HAPs, within the constraints of the mental model of users and the conveyance of that match to users at any point during the HAP's use. An analogy here might be the difference between a good lecturer who conveys the students through the argument so that they can see where it is going, and bad a lecturer, who although correct in what he or she says, loses the students.

Instantaneity is the perceived property of the HAPs that gives users a continual sense of the end-product of the HAPs. This property is a function of the complexity of the task and the more complex the task the more difficult it is to imbue *Instantaneity* in

the HAP. *SHAPE*© was perceived as having a greater degree of *Instantaneity,* i.e. subjects could 'see' where it was going.

The *Task Match* of the two HAPs was seen to be the same, however there was an issue over the type of structures subjects created with *SHAPE*©. Some subjects thought that *SHAPE*© was fine for describing hierarchical domains but were unsure about non-hierarchical domains. Some subjects thought that the hierarchical nature of *SHAPE*© was too constraining. Section 6.9 below discusses the issue of linking.

6.5 Subjects' Mental Models

A related concept to *Transparency of Purpose* is the perception the subjects' had of the purpose of the two HAPs, i.e. part of their mental models. Subjects were given an explanation of the purpose of the each HAP prior to each training session and were then asked at the end to give their understanding of the purpose of the HAPs. The answers the subjects gave were collated in several ways according to the independent variables - spatial relations ability, computing skill and order of learning. There are a number of other ways the data bits could be collated, e.g. gender, age or cross tabulated with the other independent variables, but this was seen as not worthwhile following the lack of any coherent pattern emerging from the effect of spatial relations ability.

SHAPE© was perceived to be more of a tool for developing ideas with an emphasis on the 'author as user' rather than the 'end-user as user' of a hypermedia product. TBM was seen as a HAP for creating information and an emphasis on communication

with the end user. Surprisingly there were few comments on the similarity of the end-products of both HAPs, even though the *Task-Match* statistic was the same for both HAPs. It is probably to be expected that subjects saw *SHAPE*© as a tool for developing ideas because of the original purpose of concept maps and (as section 6.9 indicates) that the semantics of linking was the focus of attention when using *SHAPE*© rather than the mechanics of linking in the case of TBM. Bearing in mind the objective of the study was to investigate ways of improving the ability to author hypermedia for educational purposes, it is interesting that *SHAPE*© was viewed in this way. Given the understanding the subjects had of *SHAPE*© as a personal ideas developer, raises the question of how suitable is the resultant hypermedia as a means for conveying knowledge about a subject to an end-user, i.e. students? Another interesting point is the idea that subjects viewed the book metaphor only as a creator of an information resource. These points are explored in the discussion, presented in chapter seven.

6.5.1 Effect of Spatial Relations Ability and Order of Learning

There was no recognisable difference in the understanding the subjects had of the two HAPs according to their spatial relations ability or the order of learning although in the quantitative results there was a significant correlation between the number of links and concepts created and spatial relations ability.

6.5.2 Effect of Computing Skill

Not surprisingly, subjects with higher computing skills gave more precise descriptions of the two HAPs. What is more interesting is that subjects with lower computing skills gave clearer descriptions of *SHAPE*© than TBM suggesting that *SHAPE*© had a higher level of *Transparency of Purpose*, reflecting a more acceptable user model. This is also very interesting bearing in mind the points made in section 6.3, concerning the supposed ubiquity of the book metaphor.

6.6 Accommodation

Accommodation was one of the factors identified in the Principle Components analysis in chapter four. There was evidence to support the concept of *Accommodation* (as shown in appendix N). *Accommodation* is the degree to which a HAP puts a user at ease and is therefore a function of *Operational Momentum* and *Instantaneity*. *Accommodation* manifests itself as a set of attitudes developed in users while they are learning and using an HAP. Shackel (1986) defines *Attitude* as the degree of satisfaction with a program. *Accommodation* is distinguished from 'satisfaction' because it conveys a greater sense of being 'put at ease' by the HAPs, i.e. *Accommodation* is more a property of the HAP than its users. A HAP that is satisfying may not be accommodating, i.e. the end result might be a good task match but was difficult to use. Subjects used words like 'enjoyable', 'comfortable', 'fun', 'interesting', 'confidence', but the word satisfaction did not appear. It seemed subjects were more concerned about feeling comfortable than satisfied - particularly those subjects with lower computing skills. Satisfaction is more closely related to *Accomplishment* is discussed in section 6.7.

6.6.1 **Differences between *SHAPE*© *and* TBM**

Although in general ***SHAPE***© was perceived to be more accommodating, some felt less comfortable with it - seemingly because they felt challenged by the task of labelling the semantics links between concepts. ***SHAPE***© was seen as more enjoyable and fun and some found TBM unexciting and less 'friendly'.

6.7 Accomplishment

There was little qualitative evidence to support the existence of the *Accomplishment* factor found in the Principle Components Analysis. However the concept is similar to Shackel's (1986) idea of satisfaction so it is likely that *Accomplishment* is a meaningful factor in the learning and use of HAPs. It is posited here that *Accomplishment* is a function of *Transparency of Operation, Transparency of Purpose* and *Accommodation* (see figure 6-3). It is not possible to say whether the use of ***SHAPE***© or TBM resulted in a greater sense of *Accomplishment*, it is only possible to say that subjects found ***SHAPE***© more transparent and more accommodating. It could be that neither HAP was sufficiently transparent or accommodating or that the task was not something subjects would get much sense of *Accomplishment* from achieving.

6.8 Ease of Use

Subjects were asked after each session to consider what they thought made a HAP easy to use. The rationale for doing this was that after using each HAP any thoughts

they had on the matter would be affected by their recent experiences. Therefore their comments would constitute their indirect evaluations of the HAPs. Almost half of the subjects made no comment. The analysis shown in table 6-3 suggests that after using TBM simplicity was an important ease of use factor and indeed, referring to section 6.3, TBM was considered to be very 'noisy'. After using *SHAPE*© the idea of playability or enjoyment features strongly in the subject's responses, although there is more diversity in the comments made. Whether this means that they found *SHAPE*© enjoyable or the converse is impossible to say but they clearly felt that HAPs should possess this property.

Table 6-3 Comments made about EU after the one-to-one training sessions.

TBM Comments on Ease of Use	Count of Subjects making comment	*SHAPE*© Comments on Ease of Use	Count of Subjects making comment
Holism	1	*Logic of Operation*	1
Simplicity	5	Physical Ergonomics of Interface	1
Windows consistent	3	Consistency	2
Feedback	1	Playable	2
		Enjoyable	1
		Simplicity	2
		Use of Graphics	1
No comments	6	**No comments**	7

6.9 Other Issues

Other phenomena arose whilst examining the data bits that seemed to interact with the main factors of *Transparency,* these were the process of linking and issues surrounding the structuring of knowledge/information, the desire to use paper for 'thinking' and finally the desire for support and practice.

6.9.1 Linking and Structuring

Appendix O contains examples of comments made about linking. When using *SHAPE*© the focus seemed to be on the semantics of the links with subjects expressing the feeling that it made them to think harder about the structure that they were creating, i.e. it requires more intellectual effort. When using TBM, comments about linking focused on the mechanics of the link rather than the semantics - subjects were frustrated because they could not see the links on the screen. Other miscellaneous comments related to the desire to link concepts at different levels, multiple linking between the same concepts and the problem of clutter on the screen (when a large number of concepts are being linked). It can be concluded that since *SHAPE*© forces its users to focus on the semantics of the linking (rather than the mechanics) visualisation of links is desirable. The desire to visualise is strong and many expressed the wish to 'draw' their structure on paper first even when they were using *SHAPE*©. This is an important issue, one of the objectives of this study was to make hypermedia authoring as easy as word-processing. When lecturing staff create text they will in most cases compose at the keyboard because it is easier than writing by hand and then entering it into the word-processor. The aim is to get staff to compose their hypermedia at the computer not on paper first so the desire to use paper has to be addressed. One frequent request made was for an 'overviewer' so that subjects could see the whole picture and this is partly the reason why there was a need to use paper.

6.9.2 Need for Practice or Support

One of the objectives of the study was to discover the factors that would lead to reducing the cognitive overhead of learning a new HAP. However, there was a number of staff who expressed the desire, at various points during the question session, to have someone alongside to help with learning and using the HAPs and also that they would need to practice using the packages otherwise they would forget. One of the aims of the study was to see how, if expert computer users can learn by themselves, the same skill can be imbued in more novice users.

6.10 Interaction of Factors

Figure 6-3 below illustrates how it is suggested the factors defined in the proceeding sections interact. What is being postulated as the dynamic of interaction with the HAPs is that the sub-factors of *Transparency of Operation* and *Purpose* contribute to the emergence of these holistic properties. In turn, *Transparency of Operation* and *Purpose* then contribute to creating a sense of *Accommodation*. *Transparency of Operation* and *Purpose* and *Accommodation* lead to a final sense of *Accomplishment*. The sub-factors are properties of the HAPs, *Transparency of Operation* and *Purpose*, *Accommodation* and *Accomplishment* are emergent properties, however there may well be functions or attributes of the HAPs that give rise to these emergent factors directly and further research is required. For example, implementing a particularly helpful and unthreatening feedback mechanism in an HAP may increase the feeling of *Accommodation* in a user with lower computer skills. To improve on the design of HAPs in the sense of the dynamics of these factors, it is necessary to must ensure all the sub-factors on the left-hand side of figure 6-3 are addressed and then to a large extent the emergent factors should improve correspondingly. There is a loose

similarity between the dynamics of the factors illustrated in figure 6-3 and Eason's

(1984) causal framework of usability.
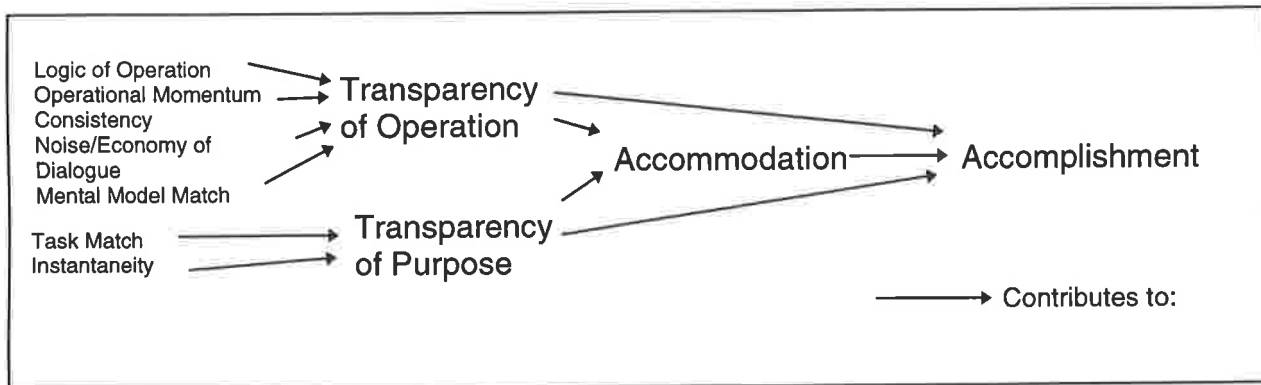
Figure 6-3 Interaction of factors in a model of learnability

6.11 Learning - *Transparency, Accommodation* and *Accomplishment*

There was little, if any, qualitative evidence to support the similar set of factors,

identified in the PCA of chapter five. The learning factors mirror the usability factors

above and may be are the same factors. Roberts and Moran (1983) and Whiteside *et*

al. (1985) found that ease of use and ease of learning are strongly related and even

congruent. Davis (1989) suggests that learning and using are not separate, disjoint

activities. The question is whether there is anything about a HAP that makes it easy to

learn but which does not necessary contribute to its usability? Common-sense dictates

that something that is easy to learn is also easy to use. When subjects were asked the

question 'What makes a HAP easy to learn?' after using each HAP their responses

were as described in table 6-4. Interestingly several subjects did make the comment

that something that is easy to use is also easy to learn. The probable congruence of

the learnability factors and the usability factors lends weight to the existence of the

usability factors in the first place.

Table 6-4 Comments on EL for both programs

Comments on Ease of Learning TBM	Count of Subjects making comment	Comments on Ease of Learning *SHAPE*©	Count of Subjects making comment
External Consistency	2	*Logic of Operation*	5
Level of Support	2	Support/Help Facility	5
Logic of Operation	2	Economy of Dialogue	2
Economy of Dialogue	2	*Operational Momentum*	1
Feedback	1		
No comments	2	**No comments**	1

What is interesting about table 6-4 is that most of the listed factors of importance are usability factors with similar comments for both HAPs. The only additional factor is 'level of support' and the 'help facility' . These findings seem to support the identity of ease of learning and ease of use. *External Consistency* is seen as important which is not surprising - mastering a new HAP is easier if its users have a good idea of where common items will be found like the File menu.

The differences between the learnability of *SHAPE*© over TBM comes down to the differences between the *Transparency of Operation and Purpose* and *Accommodation and Accomplishment* of the two HAPs. It can be argued that *SHAPE*© was perceived to have more *Transparency of Operation and Purpose* and more *Accommodation* and therefore a higher learnability and usability.

6.12 Result of Output Content Analysis

Concomitant with the findings of the analysis of chapter four, TBM and *SHAPE*© produced similar outputs, i.e. equivalence of task-match. Both HAPs produced outputs that were in keeping with the standard model with an estimated match of about 70-80% with the standard model. However, the outputs from *SHAPE*© were richer with more levels of hierarchy and more cross-links than the TBM outputs compared to the standard model. In a strict sense, TBM output matched the standard model better than the *SHAPE*© output because of the extra richness of the *SHAPE*© output. The outputs from both HAPs produced some idiosyncrasies and often the same idiosyncrasy appeared in the outputs from both HAPs. A typical idiosyncrasy would be an emphasis on a particular area like the role of research in the faculty or school depending on the particular individual idiosyncrasies of the subjects.

There were occasions where the output from *SHAPE*© contained idiosyncrasies which were not present in the TBM output. This tended to happen for subjects who used the TBM HAP first, followed by the *SHAPE*© HAP. This could be because *SHAPE*© is a better tool for exploring ideas with a greater degree of *Operational Momentum*, as emerged from the analysis described in section 6.3, and leading users on to expand their ideas. When subjects used *SHAPE*© first their ideas were probably explored more thoroughly to begin with and then these ideas were carried over into their TBM output hence the opposite phenomenon did not occur.

There were some other interesting phenomena which are worth commenting on. The names of the links in *SHAPE*© are often idiosyncratic, e.g. one subject linked a

concept 'institute management' with his 'school' with the label 'they know best' other labels were more clumsy and did not seem to relay the apparent, obvious connection. Section 6.9 above discussed the challenge some subjects faced when naming the relationships between concepts in *SHAPE*©, however this was not a problem with TBM because the link mechanism did not make subjects articulate the name of the relationship between pages. Deciding how to structure the concept maps caused some concern with several subjects separating the personnel-structure of the faculty from the rest of the details of the faculty particularly when using *SHAPE*© and then became entangled with trying to resolve where certain concepts should be placed in the hierarchy. Some subjects wanted to connect concepts at different levels or recreated the identical concepts at different levels to accommodate the problem, typified in the example shown in Figure 6-4 (a) and 6-4(b) below. One subject came up against the problem of many-to-many linking as in Entity-Relationship Diagrams as illustrated in Figure 6-5 below, which lead to a confusing concept map. Other subjects tried to represent ideas in the way the map was constructed. Figure 6-6 shows the output from *SHAPE*© which is meant to exemplify the idea of team working between members of a particular school. Figure 6-7 shows the output from *SHAPE*© where subjects wanted to create a kind of 'table of contents'.
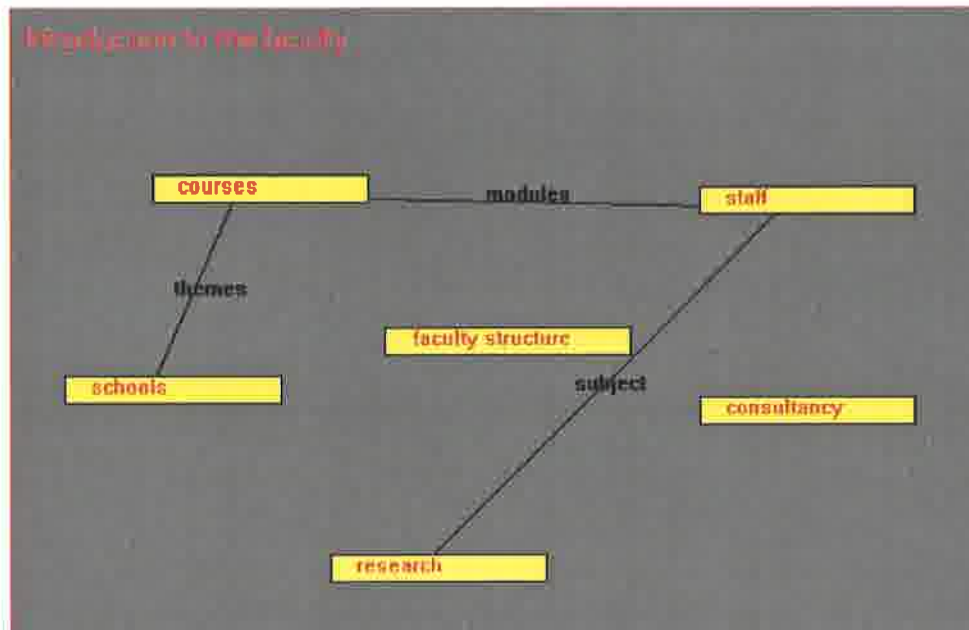
(a)
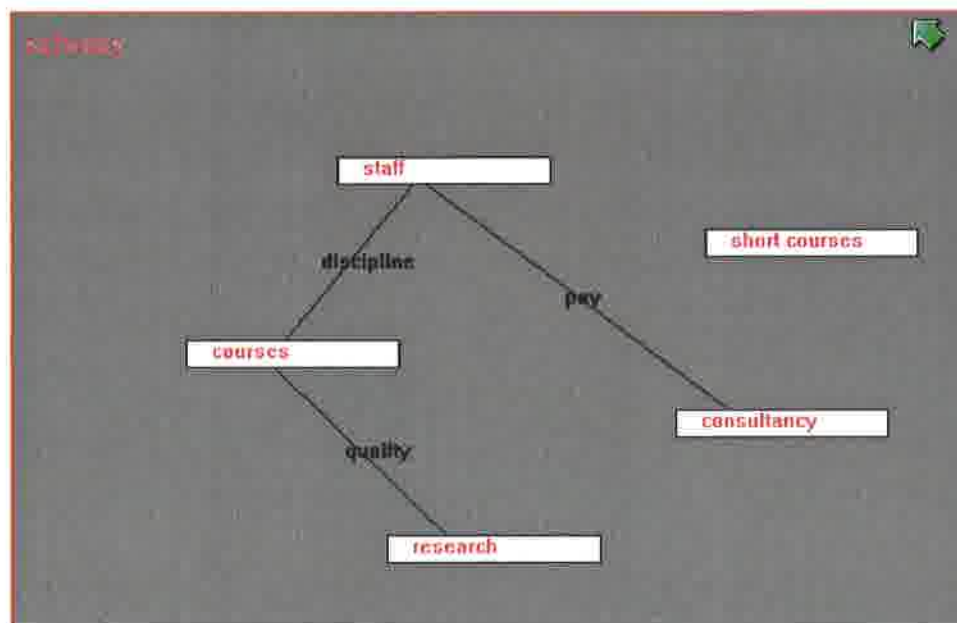
(b)

Figure 6-4 (a) and (b). Example of two levels in the same concept map where the subject recreated the same concepts.
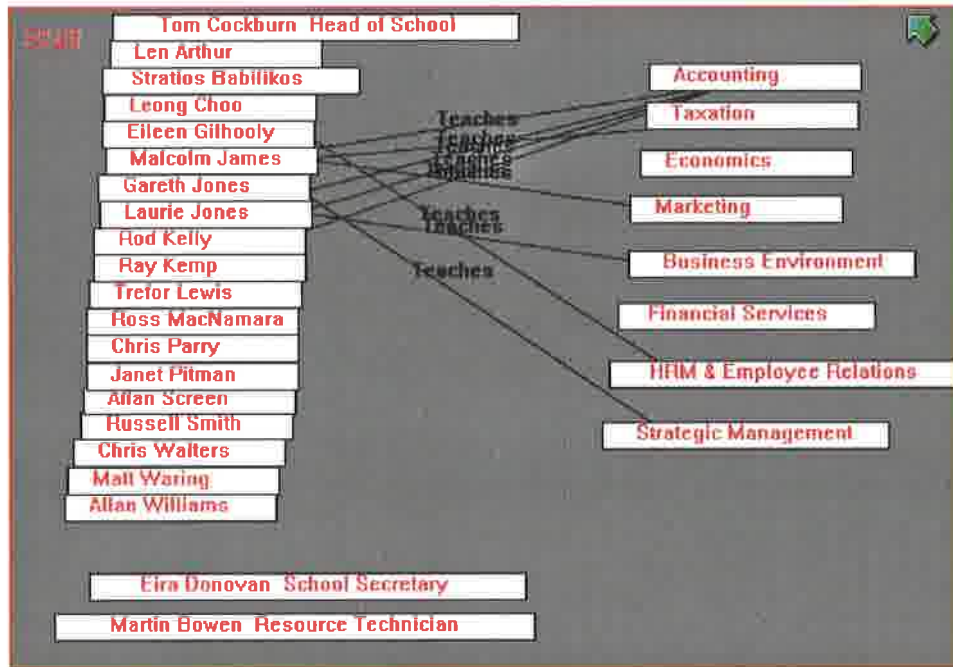
Figure 6-5 Example of a subject's concept map illustrating the problems of 'many-to-many' relationships .
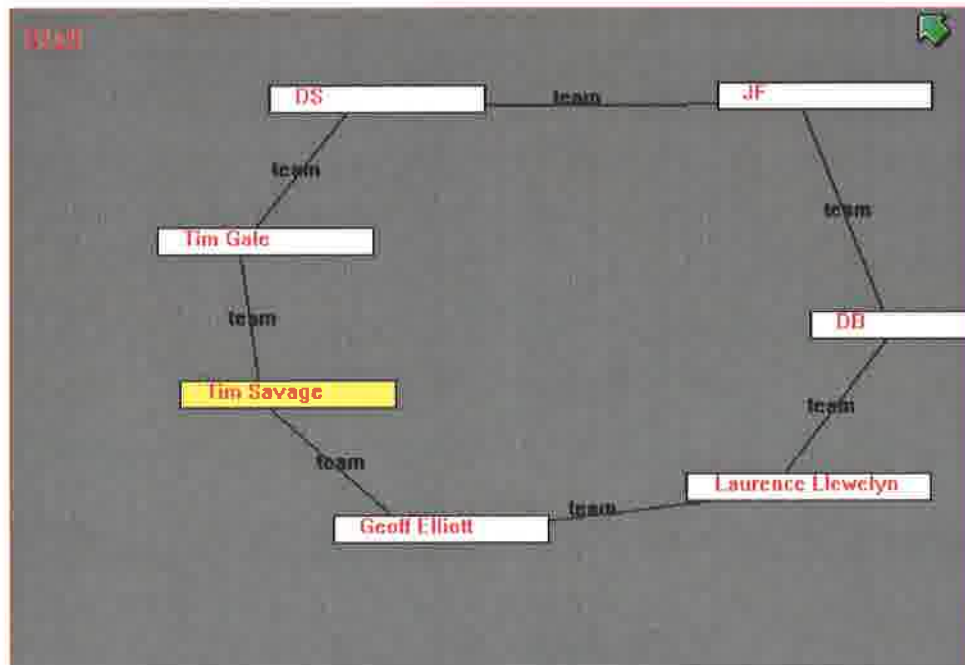
Figure 6-6 Example of the layout of a subject's concept map being used to signify 'Team-Working'
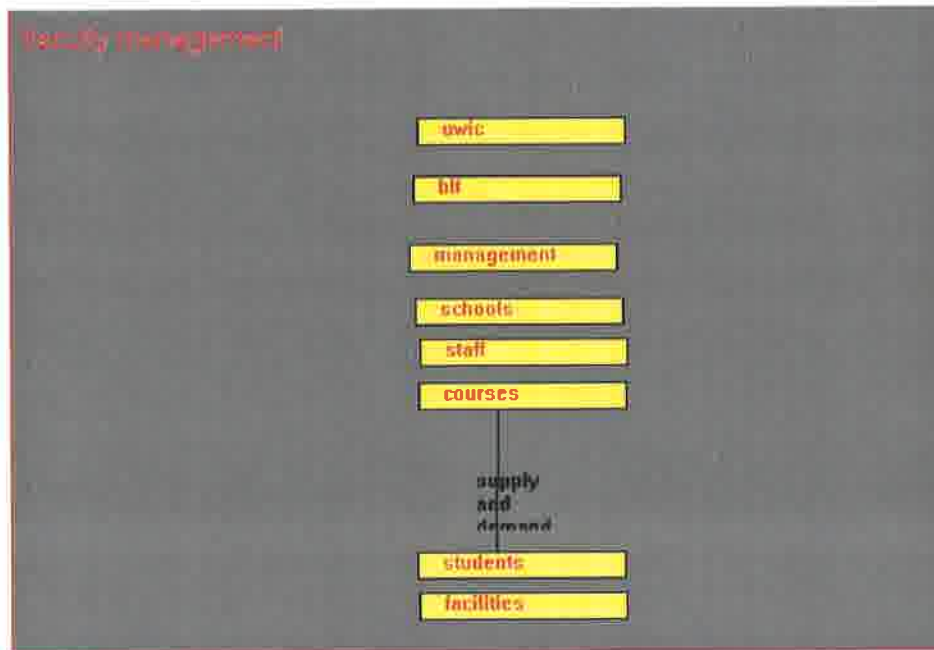
Figure 6-7 Example of the layout of a subject's concept map being used to represent a 'List of Contents'.

Chapter Seven

Comparison with other HAPs

7. Comparison with other HAPs

This chapter describes the results of a desk-based analysis of the more common HAPs and an analysis of the transcripts of the focus group. It is beyond the scope of this study to analyse every extant and emerging HAP so a selection was made (as described in the methodology - chapter three). The details of the methodology for the focus groups is described in chapter three (section 3.6). Using NUD.IST, the transcripts were searched and salient data bits were indexed to either existing categories defined during the analysis of the one-to-one interviews, or to new categories that emerged from the data.

7.1 Overview of other HAPs

Many newer HAPs have been developed in HE research projects usually for producing Web applications, e.g. W3Lessonware (Siviter, 1997), the 'exercise authoring tool' (McKeever, 1997), Webcosm, a Web-based version of Microcosm (Davis *et al.*, 1993), Hyperwave (Maurer, 1997). Some of these HAPs are similar in nature to FrontPage© (e.g. W3lessonware and the 'exercise authoring tool'), some are Web versions of open hypermedia systems (like Hyperwave and Webcosm). It was decided to investigate these other Web-based HAPs through the literature and evaluation copies. It was felt that their inclusion in the one-to-one training sessions or the desk-based analysis or the focus group study would not contribute much extra, since the major of design metaphors used in HAPs were already included (including a Web-based HAP) (see section 3.6 for the chosen HAPs).

7.2 Desk-based Comparison

Chapter one discussed in general terms the functionality of common HAPs. This chapter attempts to provide a more detailed functional analysis of the common HAPs based on the factors derived from the results of the qualitative analysis of chapter six. Riley's (1994) review of hypermedia authoring programs tended to focus on the technical aspects of the HAPs and paid little attention to the HCI and cognitive issues of authoring. Koegel and Heines (1993) analysed visual programming languages for multimedia authoring but not from the perspective of hypermedia.

7.2.1 Analysis of Operation of HAPs

In this section the way each HAP operates (Modus Operandi) is analysed together with the mechanisms used to link nodes to produce hypermedia.

ToolBook©

Modus Operandi

ToolBook© is used extensively in HE for the construction of educational hypermedia. The basis of use is the construction of pages into which media content is added. Pages are hyperlinked together. Screen objects can receive and transmit messages and a scripting language allows code to be attached to any object to respond to or generate messages like 'mouse button down'.

Hyperlinking

Hyperlinks can be created in two ways in ToolBook©, the simplest method uses a dialogue box which constructs a link from one page to another. The alternative method for hyperlinking uses ToolBook©'s 'Openscript' scripting language which is attached to screen objects like text or images, etc. in the hypermedia.

Authorware©

Modus Operandi

Authorware© is used extensively in computer-based training because of its ability to enable any interactivity created can be explicitly modelled. The basis of use is the construction of a control flow diagram where a number of parametric icons provide the interactivity and branching mechanisms.

Hyperlinking

There is no easy support for hyperlinking in Authorware©. Links are created using the interaction icon which when used results in changes in the display although in a strict sense these links are not hyperlinks. The frame icon loosely imitates the book metaphor and hence supports a means to create hyperlinks through dialogue boxes. The frame icon also provides navigational hyperlinks automatically.

SHAPE©

Modus Operandi

Described in chapter four.

Hyperlinking

The semantic links of the concept map become hyperlinks in the hypermedia. and are essentially structural, i.e. they provide 'semantic' structure to the hypermedia. The semantic links can also be construed to be navigational by the author without the need for 'typing'. Further links can be provided by ToolBook© which lies beneath **SHAPE©**.

Webmapper

Modus Operandi

Webmapper is an experimental HAP based on the concept map metaphor which creates Web pages. The basis of use is the construction of a concept map which is transposed into a skeletal set of Web documents. Unfortunately the semantic links of the concept map do not become hyperlinks in the Web documents, these have to be recreated using another function.

Hyperlinking

Although similar in nature to *SHAPE*© the link mechanism in Webmapper is different. The semantic links of the concept map have no relationship to the hyperlinks created in the final application.

FrontPage©

Modus Operandi

FrontPage© is Microsoft's flagship authoring tool for Web documents. FrontPage©, like a number of emerging Web authoring tools, works within the limitations of HTML The basis of use is the construction of a directory structure of Web documents. Hyperlinks can then be created between documents in the directory structure. Content is added using a tool that looks like a wordprocessor. FrontPage© provides a visual representation of the hyperlinked documents but this can not be edited.

Hyperlinking

Hyperlinking is achieved either through using a dialogue box mechanism or straight coding in HTML. After creation, hyperlinks can be viewed visually but can not be edited or created visually.

7.2.2 Knowledge Structuring

Following on from the discussion of knowledge in section 2.3.4 (chapter two), knowledge structuring is defined here as the process of modelling the underlying knowledge of a domain. The emphasis is on the extent to which the knowledge is explicitly represented rather than implied from the content. Knowledge representation

can take many forms as described in the literature on knowledge engineering and is discussed in chapter two. In this study the emphasis is on visualisation of knowledge since the thrust of this study is based on the experimental hypothesis that visual representations of knowledge, i.e. concept maps are more accessible to staff with lower computing skills. Tables 7-1 and 7-2 summarise the analysis of the knowledge structuring aspects of the HAPs under study.

Table 7-1 Summary of the knowledge structuring aspects of the HAPs in the study

Program	Knowledge[7]		
	Visibility of knowledge	Knowledge Types Represented	Content Construction Method
	None (N) Partial (P) High (H) Switchable (S) Zoom, (Z)	Declarative (D) Procedural (P) None (N) Structural (S)	Dialogue boxes (Di) Code (C) Diagrammatic (Dm) Templates (T)
Toolbook[©]	N	N	Di, C (Openscript), T
Authorware[©]	P, S, Z	D,	Dm
SHAPE[©]	H, S	D, S	Dm, Di, C (Openscript)
Webmapper	H, S (but not efficiently)	D, S	Di, C (HTML)
FrontPage[©]	P, S	D, S	Di, C (HTML), T

---

[7] 'Partial' means that the HAP doesn't allow an explicit representation of the knowledge but could be construed to be a visual representation. 'Switchable' means that the user can switch between a content view and a knowledge view. 'Zoom' means that the user can zoom in or out to the level of detail required. Knowledge can be roughly defined as being either 'declarative' or 'procedural' although there are many other types of knowledge. 'Structural' knowledge (Jonassen, 1993) is primarily declarative but refers to larger grain relationships (McAleese, 1990). 'Templates' or wizards refers to additions to HAPs that support some standard functions like multiple choice questions.

Table 7-2 Summary of the hyperlinking aspects of the HAPs in the study.

Program	Hyperlinking	
	Hyperlink Types[8]	Linking Mechanisms
	Navigational (N) Text to Text only (T) Semantic (S) Media to Media (M) Interactive (I)	Dialogue boxes (Di) Code (C) Diagrammatic (Dm)
Toolbook©	N, I, M	Di, C
Authorware©	N, I, M	Di
SHAPE©	N, S, I (through ToolBook©, T, (but M through ToolBook©	Dm
Webmapper	N, T (but M through HTML)	Di
FrontPage©	N, M	Di, C

With the exception of *SHAPE*© and Webmapper, none of the other HAPs really allows users to explicitly represent knowledge. ToolBook© just enables users to create content-pages and hyperlink the pages together, it does not have any facility for describing the underlying knowledge that the hypermedia corpus represents. FrontPage© provides the facility to visualise the hyperlinks between Web objects but does not allow users to create links visually. Authorware© enables users to model the interaction and sequencing of material visually but this is a long way from being able to model the underlying knowledge. Creating hypermedia in Authorware© is not particularly straightforward and becomes complex very quickly. *SHAPE*© as already stated allows users the opportunity to visually structure the knowledge using concept maps. The key to *SHAPE*©'s ability to allow users to model their knowledge is the assumption that a semantic relationship corresponds to a useful hyperlink. The content analysis conducted in chapter six suggested that the corresponding hyperlinks are useful. The limitations of concept maps as a knowledge representation scheme

---

[8] 'Semantic' hyperlinks reflect relationships in the underlying knowledge. 'Interactive' links are part of an interactive sequences. 'Text to text' is a subset of 'media to media'

were considered in chapter four. Webmapper also allows users to structure their knowledge with concept maps but the correspondence to hyperlinks is missing - they must be created separately. All the HAPs in this study were deficient in one way or another in terms of their ability to represent knowledge.

This analysis raises two questions: how can HE teaching subjects be adequately and easily modelled and represented; and how best can the knowledge representations be 'converted' into hypermedia?

7.2.3 Output

ToolBook© and Authorware© create applications which require freely available auxiliary files when distributed. Like the suppliers of many other HAPs, ToolBook© and Authorware© applications can be made to run over the Web with some extra free software and some additional processing of the application. Webmapper and FrontPage© are HTML applications and can therefore run over the Web or be distributed with browsers. *SHAPE©* produces ToolBook© applications so whatever applies to ToolBook© applies to *SHAPE©*.

7.2.4 Summary

It is clear that commercially-available HAPs do not have well-developed facilities for structuring knowledge. The reasons for this maybe that the developers do not think it is necessary - it may be only relevant in HE or it was simply not considered. In terms of the output, all HAPs could produce equitable material although the ease with which

hyperlinks could be added varied considerable and is considered in the section below.

Based on the above analysis it was felt that HAPs should possess the following properties:

- An ability to model the underlying knowledge

- An ability to switch between views

- An ability to zoom in and out

- An ability to represent different knowledge types and links

- An ability to provide several mechanisms for creating content and hyperlinks

- An ability to allow the definition of various hyperlink types

It is important however to note that any feature is optional when implemented

7.3 Construction of an Application

Excluding **SHAPE**©, and possibly Webmapper, all the other HAPs have been used to build educational hypermedia and used in teaching-learning scenarios. In order to evaluate the effectiveness of **SHAPE**© it was seen important to construct a real educational application as noted in the methodology (section 3.4). In order to ensure the subject who developed the hypermedia applications could make intelligent comments he was asked to construct a similar application using Microsoft FrontPage© The subject concerned - Greg Dainty, - was a participant in the focus group so his experiences of developing applications are analysed indirectly through the focus group results given below.

The undergraduate module that was chosen for the creating the hypermedia package was a compulsory module on basic IT knowledge and skills that has around 240 students. Delivering this module has always been a challenge in terms of ensuring students are adequately supported and have access to appropriate learning materials. An underpinning philosophy for the hypermedia application (later called **CLICK-IT**©) was developed and is given in figure 7-1 below. Essentially **CLICK-IT**© was designed to be a learning support package.
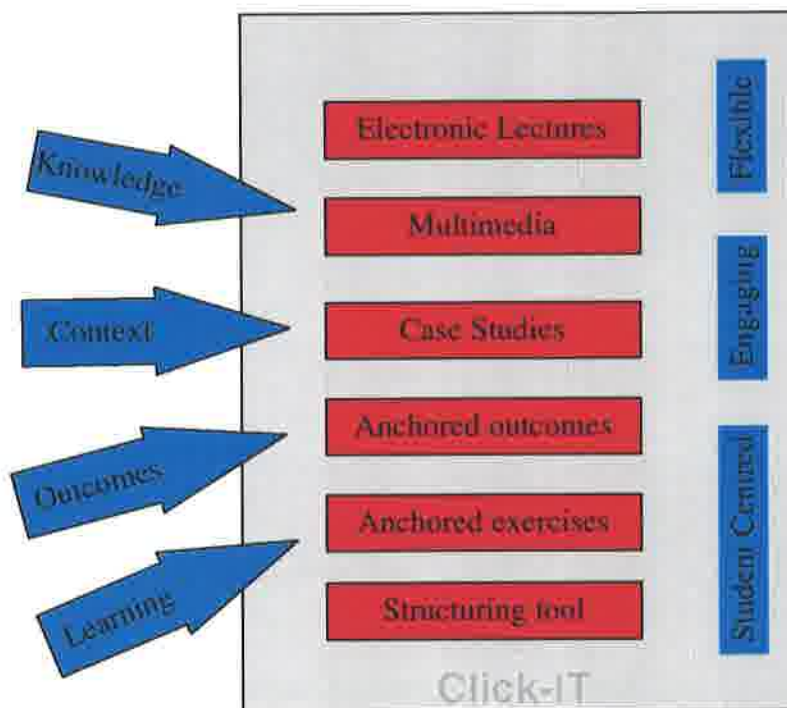
Figure 7-1   Diagrammatic representation of the underpinning philosophy of CLICK-IT©.

It was found that *SHAPE*© could be used to produce the structure of the application very quickly. This involved generally only two levels of concept map with one branch descending to a third level (see appendix S for the complete set of concept maps of **CLICK-IT**©). Figure 4-8 and 4-9 (in chapter four) are examples of content pages

taken from **CLICK-IT**©. The first interesting point to note in the development of **CLICK-IT**© was the blurring between concepts and links that represented the knowledge domain and those that helped to structure the material for the students. In the top-level concept map shown in appendix S the concept 'Introduction' appears which is clearly not a concept in the knowledge domain of IT but a clue for students about where to start - it seems the book metaphor is very pervasive! This issue is considered a problem by some, for example Reader (1995) who believe that tools should be constraining. However, in the case of *SHAPE*©, it has proved to be useful in allowing both kinds of knowledge to be represented. In the concept maps at the second level (see appendix S), the developer (Greg Dainty) dispensed with using meaningful labels for the links and relied on the default label 'related-to' on the basis that they were unnecessary since the hyperlinks in the resultant hypermedia would exist whether they were labelled in the concept map or not. It is believed that this possibly reflects the difficulty sometimes of labelling semantic relationships meaningfully (noted in the analysis of the one-to-one training sessions section 6.9) rather than as a basis for a good modus operandi. The rapid development of the structure of the material was motivating. However the content and functionality outlined in figure 7-1 still had to be introduced into the content pages using ToolBook©. ToolBook© is not diagram-based and suffers from all the problems that this study has attempted to resolve. *SHAPE*© allows for the linking of discrete chunks, i.e. concept to concept, but links from within content-objects (like text or pictures) to other content-objects again had to be constructed using ToolBook©. Exactly how a diagram-based design metaphor might be developed to allow this is difficult to imagine but remains an important goal.

7.4 Learnability/Usability Comparison - Focus Group

7.4.1 Participants

The participants were chosen because of their involvement in IT staff training and expertise in developing computer-based learning materials. It was felt that these people would be best able to comment on the learnability, usability and utility of the HAPs under review. The participants were:

Eleri Jones

Eleri Jones is Dean of Resources and Special Projects in the Faculty of Business Leisure and Food and is responsible for the development of a faculty teaching and learning strategy that attempts to emphasise the use of computers. She has been involved in a number of projects in the faculty that utilise computers for the delivery of educational objectives. Eleri has also used *SHAPE*© to develop a hypermedia application for the teaching of Biochemistry.

Mike Snelgrove

Mike Snelgrove teaches IT and information systems to undergraduates. Mike also develops materials for the Web and runs staff development workshops on the utilisation of the Web in teaching and learning. Mike is conducting research on information seeking strategies in intranets.

Kim Walsh

Kim Walsh is the college IT trainer for staff and has developed and delivered a range

of IT courses for staff. Kim Walsh has been developing Web-based materials for her

courses.

Greg Dainty

Greg Dainty is also involved in developing and running IT related courses for staff.

Greg Dainty has also built a hypermedia learning pack for helping to teach IT using

SHAPE© and FrontPage© (this work is discussed below in section 7.3). Greg Dainty

is currently pursuing a PhD on the educational effectiveness of hypermedia.

7.4.2 Analysis of the Results of the Focus Group

The details of the procedure for conducting and analysing the focus group session are

given in chapter three. The transcripts of the focus group were added to the NUD.IST

database constructed during the analysis of the one-to-one interviews. There was a

separate transcript of Greg Dainty's experience of creating a full application with

SHAPE© and FrontPage©. As the transcripts were analysed, data bits were indexed

to pre-existing categories or if it was felt appropriate, new categories were created in

NUD.IST. Data bits were often indexed to more than one category. Once the

transcript had been indexed the data bits indexed to each category were printed then

searched for the presence of any emergent phenomena. It is important to note that the

objective of the focus group was to capture emerging ideas/notions from the

discussions of all the HAPs and not specifically the evaluation of any individual

HAPs. The discussions were generally oriented around the use of the HAPs by

academics with relatively low Computer Skills who perhaps could word process but nothing else. The term 'notion' used here defines a statement or 'handle' that attempts to encapsulate the comments made by the participants. It is important to note that notions are not the words of the participants but the synthesis of a number of their comments into an embracing statement. Appendix P details the notions/ideas that emerged for the category that were defined in NUD.IST. Appendix Q contains the output data bits indexed to learnability to demonstrate the principle behind the analysis. The table in appendix P was successively analysed, refined, restructured, reduced, ideas realigned and categories regrouped, in line with the interaction of factors model presented in figure 6-3. The final results are discussed below with two new subfactors defined. The discussion is laid out with the key 'notions' associated with the factor and them a discussion of the factor. In listing the notions related to a factor a key is used to indicate from which HAP it originated. (A) is Authorware$^{©}$ (S) is **SHAPE$^{©}$**, (F) is Frontpage, (W) is Webmapper, (T) is ToolBook$^{©}$ (D) is the post demonstration discussion, (GD) identifies specific comments made by Greg Dainty.

Learnability

Notions

Long learn time is undesirable(A)

The need for memorability (S)

Some people will always need some instruction (S)

Content can affect learnability (S)

Iteration helps learnability (S)

Motivation to learn driven by accomplishment/utility - Computer Skill dependent (A)

Motivation and accomplishment are interrelated (A)

Analysis

The notions above suggest ways in which learnability can be improved. The shorter the learn-time, the better the memorability of the program. Memorability is affected by the mental model of users of the HAP as discussed below under mental model match. There is a dynamic between learnability, motivation and *Accomplishment*. The better the sense of *Accomplishment*, the more motivated the learner, and hence the more learnable the HAP will appear. If the initial target application is difficult, i.e. too complex or ill structured, the learnability of the HAP is likely to be reduced. No matter how learnable a HAP appears to be intrinsically or otherwise there will be some users who will want some human-to-human help.

Separating Development Phases

Notions

Separating the phases increases Transparency of Operation (D)

The need to separate the phases of authoring (F)

Separate phases can increase utility (D)

Analysis

Learnability and other usability factors can be improved if the authoring process is divided into phases, provided that transparency of purpose can be maintained. The likely phases would be a knowledge definition phase and then function supplementation phase.

Knowledge Representation

Notions

Concept maps don't match entirely to the task (D)

Inadequacy of representational scheme to model what is required (S)

How explicit can you make the knowledge? (A)

How much of the knowledge should be visible? (general comment)

The need for clarity of representation (S)

What constitutes a representation of knowledge? (F)

Knowledge needs to be visible to help increase functionality (GD)

How to represent procedural knowledge? (S)

The need for zoom in and out of knowledge representation (W)

The need to see the whole 'structure' (D)

Analysis

These notions raise a number of issues but offer little in the way of solutions. There is a problem of adequately describing the required knowledge of a domain (shared by knowledge engineers). What constitutes a description of knowledge is an interesting question. The notion suggested by Greg Dainty above is that a visible representation of knowledge can improve functionality; presumably, once a domain has been made 'visible' it becomes easier to add functionality. The facility to zoom in and out of the representation was seen as being important.

Linking

Notions

Semantic links are useful but are not the only way to create hyperlinks (S)

Intuitiveness of hyperlinking is dependent on prior knowledge (F)

Intuitiveness of linking is dependent on the domain (S)

The importance of naming semantic links is not necessarily important to creating hypermedia. (GD)

Things can be linked with 'frames' instead of 'lines' i.e. embracing similar things (S)

What is the purpose of linking visually? (W)

Analysis

A related problem to that of knowledge representation in hypermedia is the creating of links. If the links are semantic then the ease with which they can be created is partially dependent on the domain, ill-structured or complex domains will not have obvious semantic relationships. An interesting phenomena described by two of the participants was that the labels of some of the semantic relations in *SHAPE*© were

either not easy to articulate or were considered to be unnecessary. Defining semantic relations by other means than lines was also discussed and having frames around concepts was considered the same as linking them.

Operational Momentum (part of Transparency of Operation)

Notions

Variety of prompt types helps users to proceed (S)

Some programs do have a sense of momentum (A)

Providing cues increases momentum (D)

Feeling that one is on a trip is synonymous with operational momentum (D)

Making decisions for the learner can increase operational momentum (S)

Eliminate activities that don't contribute to the end result - Transparency of purpose (W)

Analysis

The notions derived here provide some suggestions on improving *Operational Momentum* and also explicit reference to its existence as a property of a HAP.

Hidden structure (part of Transparency of Operation)

Notions

Hidden limitations reduces Task-Match but only after discovery (F)

The danger of hidden information (F)

Hidden conditions (F)

Analysis

HAPs should address assumptions about knowledge of the system. The notions are all related to FrontPage©and the limitations of HTML. The HAP should somehow ensure that users have the necessary prerequisite knowledge of the paradigm. The second notion refers to FrontPage©'s attempt to mitigate a lack of understanding by users by hiding the details of the documents it generates. This concealment can cause its own problems when users try to edit or modify their applications once they become more proficient.

Complexity (part of Transparency of Operation)

Notions

Usability versus functional power (A)

Functional power versus usability (D)

Functional power can increase operational momentum (W)

Complexity reduces momentum (A)

The need for a complexity gradient for learners (T)

It is difficult to make complex applications transparent (D)

Analysis

The problem of functional power and expressiveness versus complexity was addressed in chapter four in the design of *SHAPE©*. Increasing functional power can increase operational momentum but runs the risk of increasing complexity and hence reducing operational momentum. A solution is suggested which is to have a complexity gradient for learners something some application producers have attempted to do.

Mental Model Match (part of Transparency of Operation)

Notions

Finding a handle - looking for an equivalent idea to understand the new HAP - mental model match. (in reference to hypermedia generally)

Users seek for a 'handle' to understand (A)

Accessibility of design metaphor (F)

Usefulness of design metaphor to end product (T)

The need to reduce cognitive overhead (F)

The need to provide a handle (F)

The concept maps are intuitive not the program (S)

Logical consistency of the HAP may not correspond to the users' sense of logic (W)

Analysis

The issue of mental models was discussed in chapter two (section 2.4.1) and its development as a factor in the qualitative analysis in chapter six (section 6.5). The above notions suggest a number of issues that need to be taken into account. Users seek to understand a new HAP by 'fitting' it into a pre-existing mental structure. With this in mind the design metaphor used in the HAP should be capable of addressing as broad a sample of the potential users as possible and perhaps be adjustable to the mental models that users may have. Complications occur when the design metaphor does not have a clear correspondence to the end application.

The Effect of Prior Knowledge (part of Mental Model Match)

Notions

Required wider knowledge increases cognitive overhead (F)

The need for prior knowledge (F)

Prior knowledge can lead to misconceptions or transparency (F)

Extent of use of the metaphor or prior knowledge (T)

Associating with prior ideas can mislead and cause misconceptions (F)

Usability affected by prior knowledge (F)

Preconditioning dictates usage (W)

The need to avoid misconceptions (F)

Prior knowledge can lead to misconceptions (F)

HAPs must respond to a wide range of prior knowledge (D)

Analysis

These notions are directly related to *Mental Model Match* and simply raise the problems of how a HAP can address all the complexity of users' mental model.

Transparency of Purpose

Notions

Visibility of final product when authoring (F)

The speed of visibility of result is important (D)

Any activity undertaken must reflect in the final product (W)

Speed of result improves transparency of purpose (S)

Analysis

These notions indicate the existence of *Transparency of Purpose* and the importance of seeing the form of the end result as early as possible which is the same concept as instantaneity developed in chapter six.

Utility (part of Transparency of Purpose)

Notions

Identity of HAP with final product is important (A)

Proficiency precedes an appreciation of Task-Match (A)

Proficiency precedes perceived usefulness (A)

Variety of final products expected of HAP. (GD)

Visual representations increase utility (D)

Connectedness to other tools (D)

Analysis

Utility is an accepted factor in human-computer interaction - see chapter two (section 2.4.3) - and these notions lend more weight to its importance in learnability and usability. This factor was not identified in the analysis of the one-to-one interviews but was identified by the experts in the focus group. This fact reflects the problems that novices have in seeing the value in a HAP, also reflected in the notions above where proficiency is needed before the *Utility* can be perceived. There are differences between *Utility* and *Task-Match* (as discussed in section 6.4, chapter six) but for the

sake of parsimony it seems reasonable to use one word for both. Since *Utility* is a more current term it will be used in place of *Task-Match*.

Accommodation

Notions

HAPs must induce confidence (D)

Confidence partially dependent on familiarity (D)

Analysis

Accommodation was recognised in the Principle Components Analysis of the one-to-one interviews given in chapter five (section 5.2) and expanded on in chapter six section (6.6). There was not a great deal of evidence here to reinforce its existence as a factor but the notion of the importance of imbuing confidence in users as a requisite for increased usability is recognised. The question is how can a HAP imbue confidence?

Accomplishment

Notions

Utility is a component of transparency of purpose (D)

Motivation to learn driven by accomplishment/utility - Computer Skill dependent (A)

Accomplishment derives from accommodation and is a property of users (A)

Motivation and accomplishment are interrelated (A)

Accomplishment and utility are not necessarily related (D)

Analysis

Accomplishment as a factor is supported in the notions above and indicates the dynamics between factors that result in users feeling a sense of *Accomplishment.* *Accomplishment* is driven by the utility of the HAP, the motivation of users - which in turn is affected by the perceived utility of the HAP. The last notion makes the important point that a HAP might have a low utility but a user could still gets a sense of *Accomplishment* because of the effort they have put in.

7.4.3 Learnability and Usability Ratings of the HAPs

The participants were also asked to rate each HAP on a number of items designed to measure learnability and usability. The items were derived from the qualitative analysis of the one-to-one training sessions. The results are clearly not statistically analysable but calculating the mode values for each item is legitimate and provides a more concrete picture of the opinions of the focus group participants. The results are summarised in tables 7-3, 7-4 and 7-5 below and discussed in the next chapter.

Table 7-3 Ratings of intuitiveness of hyperlinking based on mode values.
T=ToolBook© S=*SHAPE*© F=FrontPage© W=Webmapper, A=Authorware©

Item		Very intuitive:	Reasonably intuitive:	Un-intuitive	Very un-intuitive
1	How intuitive is hyperlinking?	T	S, F	W, A	
2	How intuitive is hyperlinking for a learner?		S	W, F, T	A

Table 7-4 Ratings of *Transparency of Purpose* for each HAP based on mode values.

Item		High	Medium	Low
3	To what extent do you get a sense of how the final product will look like while you use this program?	S, A, T, F		W
4	To what extent do you think learners would get a sense of how the final product will look like while they use this program?	S, A, T, F	W	
5	How well does this program create a product that matches the task?	S, T	W, A, F	
6	How easily does this program create a product that matches the task for the learner?	S	T, F	W, A

Table 7-5 Ratings of Learnability and Usability of each HAP based on mode values.

Item		High	Medium	Low
7	How obvious are the operations of this program?	T	S, A	W, F
8	How obvious are the operations of this program for a learner?		S	W, A, T, F
9	How well does the program 'carry' you through?		A, T	S, W, F
10	How well do you think the program will 'carry' the learner through?			S, W, A, T, F
11	How much do you feel the program is making decisions for you?		A, T	S, W, F
12	How much do you feel the program would make decisions for the learner?		A, T	S, W, F
13	How 'logical' is the overall operation of this program?	S, A, T		W, F
14	How 'logical' do you think the overall operation of this program will be for the learner?	S, A	T	W, F
15	How well does this program match how you would expect it to operate?	A, T	S, F	W
16	How well do you think this program will match how the learner would expect it to operate?	S, T	A	W, F
17	How consistent are the operations of this program?	S, A	T, F	W
18	How consistent do you think the learner will find the operations of this program?	S, A	T, F	
19	How much unnecessary detail is there on the screen for the learner?		A, T, F	S, W
20	How at ease are you with this program?	S, T	W, A	F
21	How at ease would users be with this program?		S, T, A	W, F
22	Do you think this program gives you a sense of accomplishment?	S, A, T	W, F	
23	Do you think this program will give the learner a sense of accomplishment?	S, A	T, F	W

Interaction of Factors

The above analysis resulted in the model of learnability/usability developed in chapter

six now needing some amendments. These amendments are the addition of two

factors to *Transparency of Operation, Complexity and Hidden-Structure* and the

addition of *Utility to Transparency of Purpose. Motivation* is added as a factor,

contributed to by *Transparency of Purpose* and contributing to *Accomplishment.* The

original interaction of factors model is shown in figure 7-2(a) and the amended one is

shown in figure 7-2(b) below.
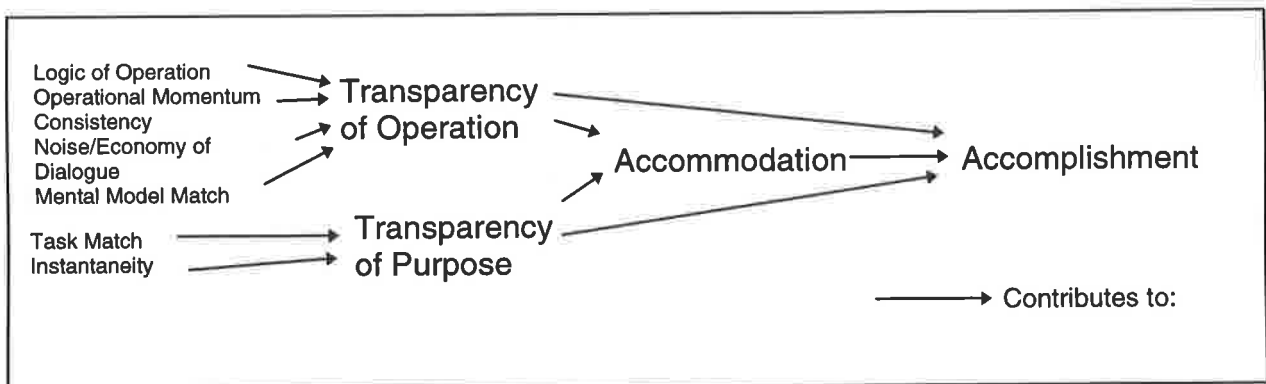
Figure 7-2(a) The model of learnability as developed from the analysis of the one-to-one training sessions
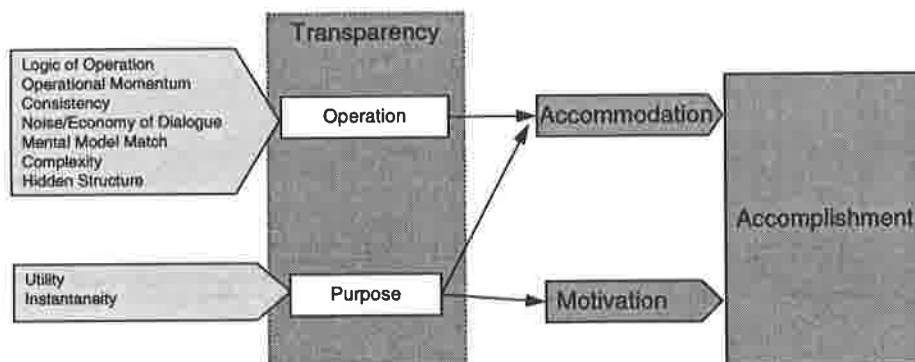
Figure 7-2(b) The model of learnability refined after the analysis of the focus group.

7.5 Summary

In order to evaluate the nature of HAPs in this study the following techniques were used:

- a desk-based examination of the modus operandi and functionality was conducted;

- the transcripts of the focus group (that considered the learnability and usability of the HAPs) was analysed using NUD.IST; and

- the transcripts of the experiences of someone developing an HE teaching application using **SHAPE**© and FrontPage© was analysed using NUD.IST.

The conclusions of the desk-based examination was that HAPs should allow users to express the underlying knowledge of the domain and for that to be mappable onto hypermedia. HAPs should have extra functionality that allows the easy expression of knowledge and also the ability the define and use different hyperlink types. There were many ideas that emerged from of the analysis of the transcripts of the focus group and the experiences of a developer. These are summarised below:

- There was more evidence to support the existence of the factors discovered during the analysis of the one-to-one interviews and Principle Components Analysis.

- In addition suggestions on how certain factors could be improved were suggested together with other issues not yet considered.

- A number of new factors were suggested which are concomitant with similar ones found in the literature, i.e. *Complexity, Utility* and *Hidden Structure*. These will be considered at greater length in the discussion of chapter eight.

- A number of other points to consider, i.e. separation of the design phase into a knowledge definition and functional definition phase and the consideration of the affect of prior knowledge on the mental models of users of the HAP.

Chapter Eight

Discussion

8. Discussion

8.1 Summary of Research Results

Before commencing the discussion a summary of the major findings are given. A list of the findings of this study are given in appendix R.

8.1.1 Summary of Quantitative Results

The results suggest subjects preferred *SHAPE*© to TBM in that they found it easier to understand, produced richer hypermedia, and it had a higher utility. Subjects were also more motivated to continue with *SHAPE*© than TBM. The perceived usability of *SHAPE*© was independent of Computer Skill but for TBM it was highly dependent. Subjects' Spatial Relations ability affected their ability to use *SHAPE*© in that they produced richer concept maps the higher their Spatial Relations ability but there was no similar correlation between subjects' output from TBM and their Spatial Relations ability. The effect of Order of Learning was mostly insignificant and did not detract from the main effects. Other results included the discovery of the significant positive correlation between Ease of Learning (EL) and Ease of Use (EU) and three underlying factors, *Transparency, Accommodation,* and *Accomplishment.*

8.1.2 Qualitative Results

Analysis of the one-to-one training sessions

Corroborating evidence was found for the factors discovered in the Principle Components Analysis and for the close relation between Ease of Learning and Ease of Using. The analysis pointed to a subdivision of *Transparency* into *Transparency of Purpose* and *Transparency of Operation.* Other subfactors were discovered and included *Operational Momentum, Logic of Operation, Noise/Economy of Dialogue, Mental Model Match, External and Internal Consistency.* There was evidence to suggest how the factors interacted as shown in figure 6-3, (chapter six).

In considering *SHAPE*© and TBM, subjects gave more precise definitions of *SHAPE*© and found it more enjoyable. What was also interesting was subjects' impression that *SHAPE*© was an 'ideas developer' and TBM a 'tool for creating information'. In addition when asked to comment on what factors contribute to Ease of Use, subjects suggested simplicity after using TBM and playable/enjoyment after using *SHAPE*©. and further comments on linking with *SHAPE*©. focused on the semantics but with TBM they focused on the mechanics of linking.

Desk-based analysis of HAPs

In terms of their knowledge construction/structuring functions, the desk-based analysis pointed to the following possible improvements in HAPs:

- Switch between views.

- Zoom in and out of detail.

- Represent different knowledge types.

- Create hyperlinks in a number of ways.

- Allow the definition of different knowledge types

Analysis of focus group and experiences of constructing a real application

The focus group mostly corroborated the factors discovered in the one-to-one interviews and the existence of a number of other factors, *Utility, Hidden Structure, Complexity* and *Motivation* were reiterated. The interaction of factors model of figure 6-3 was modified in figure 7-2(b). Other issues related to HAPs included:

- Inadequacy of the HAPs to represent what is required

- Facility to zoom in and out

- visual representations in general a good thing

- How necessary is it to 'see' knowledge

Most significantly the idea of separating authoring into a knowledge definition and a functionality-construction phase was considered as an improvement in HAPs.

8.2 An Evaluation of Diagram-Based Hypermedia Authoring (Objective One)

The results of the various parts of this study would suggest that concept maps can be beneficial in increasing the learnability/usability of HAPs. **SHAPE**© was found to be better than TBM in a number of respects as indicated in the results of the quantitative

analysis summarised above. Whether these results could be extrapolated to all types of diagrammatic representations remains to be seen.

8.2.1 Functionality

$SHAPE^{©}$ was shown to produce hypermedia that was equitable to that produced from other HAPs. $SHAPE^{©}$ does not have any facility for providing functionality but relies on the functionality of ToolBook$^{©}$. Ideally $SHAPE^{©}$ should also allow functionality to be added diagrammatically and is considered in section 8.5 below. $SHAPE^{©}$ does not address Halasz's concern, noted in chapter two, that HAPs need to be tailorable to produce various types of hypermedia as indicated in his taxonomy also (presented in chapter two). How a diagrammatic representation can achieve this goal in education is considered in section 8.5 below.

8.2.2 Learnability

The results clearly showed that $SHAPE^{©}$ was easier to learn than TBM but a number of factors do need to be considered. Spatial relations ability did affect the richness of the hypermedia that was produced however even subjects with lower Spatial Relations ability still produced better and richer hypermedia than they did with TBM. What contributed to the higher learnability of $SHAPE^{©}$ could be that it was more 'enjoyable' and subjects could see the results of their efforts sooner and more clearly (higher instantaneity). An important finding was the lack of relationship between Computer Skill and subjects' ratings of Ease of Learning of $SHAPE^{©}$, indicating that

design metaphors based on concept maps are advantageous for improving the learnability/usability for HE staff.

8.2.3 Usability

The study has underpinned the idea that learnability and usability although different are closely related. Subjects found *SHAPE*© easier to use but this was compounded by a subject's prior knowledge of concept maps and their Spatial Relations ability. Prior knowledge of concept maps is obviously a factor in a subject's mental model of the HAPs and is considered in section 8.4 below. It is argued here that diagrams (concept maps in the case of this study) can increase *Operational Momentum* and *Instantaneity* and therefore *Transparency of Operation*.

8.2.4 Hyperlinking

When first confronted with *SHAPE*© subjects were unable to see the purpose of adding semantic links but most felt comfortable with the idea of a semantic relationship. Some had difficulty in labelling the semantic relations and, as Greg Dainty[9] mentions in his experiences of using *SHAPE*© to create an application, once the purpose of the semantic link become clear, i.e. it was transposed into a hyperlink, labelling the relationships in the concept map became unimportant. The hyperlinks that were produced with *SHAPE*© as noted in the analysis of the one-to-one training sessions were utilitarian and comparable to links created purposely as in the case of the standard model to which the output was compared. Greg Dainty noted that he

---

[9]Greg Dainty developed **CLICK-IT**© using *SHAPE*©, see chapter seven, section 7.3 and 7.4

started to use the links in the concept maps purely for navigational purposes which raises the question of whether typing of links is necessary or whether it is up to the author to 'know' whether a link is semantic or navigational or whatever. Another question is whether it matters or whether all semantic links should be considered to be navigational links.

8.2.5 Utility

Despite some idiosyncrasies in the concept maps that subjects created when using *SHAPE*©, the resultant hypermedia was comparable to that produced by more traditional means. The idiosyncrasies were caused by the inability of concept maps to represent what subjects wanted to achieve. The educational-efficiency of the resultant hypermedia from *SHAPE*© is probably quite poor because it has low functionality, it is not much better than a collection of linked Web documents, a problem which is considered in section 8.5.3 below. The hypermedia produced by TBM was very similar to that produced by *SHAPE*© so it can be concluded that for these two HAPs at least, there is nothing unusual in the hypermedia produced by *SHAPE*©.

The methodology raised the question that if diagrams (in this case concept maps) do facilitate learning and use, what is it about them that has this property and what is it about users that finds them so? These questions are considered in the growing study of diagrammatic reasoning (Kulpa, 1994) mentioned in chapter two. Haber's (1970) startling discovery of the almost limitless capacity of human visual memory must be an important factor and is echoed by Buzan (1993), although the ability to recognise that one has seen an image is different to the ability to construct and reason with

images or diagrams. The choice of using concept maps was taken because they seemed to have more similarity with hypermedia and had more (semantic) information than the other techniques. It might be worthwhile exploring the use of the other diagramming techniques mentioned in chapter two although, as noted there, many of the techniques are essentially the same.

8.3 Evaluation of HAPs (Objective Two)

As noted in the literature review, in chapter two there has been little or no work done on the usability of HAPs, the effort has been on the usability of the hypermedia itself. This was one of the main reasons for its instigation as an objective of this study.

8.3.1 Functionality

The HAPs used in this study were selected on the basis of being a reasonable sample of HAPs designed for creating discrete hypermedia so no examples of open hypermedia HAPs (e.g. Hyperwave) were included. The conclusions of the desk-based evaluation of HAPs described in chapter seven have been summarised above in section 8.1. Essentially the conclusions point to the need for HAPs to address the knowledge construction aspects of authoring, discussed in section 8.4.5 below. One of the conclusions of the focus group study was that HAPs should interconnect as seamlessly as possible with other content tools rather than try to embody all necessary content provision. The principle of interrelated programs is in line with the process PC-based programs are generally following using Microsoft's Dynamic Link Library (DLL) and Object Linking and Embedding (OLE) architectures.

8.3.2 Learnability and Usability

As pointed out in several places in this thesis a variety of design metaphors are used in the operation of HAPs. The HAPs used in the focus group study represent the main design metaphors plus the inclusion of concept maps. As might be expected the focus group participants found the book metaphor of ToolBook© the most obvious (item 7 in table 7-5) - emphasising the main reason why the book metaphor was used as a bench mark for *SHAPE©*. What is also interesting is that the participants thought that none of the design metaphors were very obvious for learners, including the book metaphor (item 8 table in 7-5). Participants rated *SHAPE©* highest for obviousness for learners but this must be tempered with the fact that the other map-based design metaphor was rated low, the conclusion being that it is not just the design metaphor that needs to be considered in the usability of a HAP. The participants felt that the HAPs did not carry them through (*Operational Momentum* item 9) and they considered the situation even worse for learners - rating all HAPs low (item 10). Participants rated *SHAPE©* and Authorware© highly logical (items 13 and 14) with similar ratings on logicality given for learners, supporting the experimental hypothesis that diagram-based HAPs have a higher *Transparency of Operation*. The two Web based-HAPs (Webmapper and FrontPage©) were rated low, which might be indicative of the extra knowledge required to understand Web-based documents. ToolBook was rated high (items 15 and 16) on how one would expect it to operate which again is probably due to the familiarity of the book metaphor. *SHAPE©* scored high on how a learner would expect a HAP to operate although again the other map-based HAP

Webmapper scored low indicating that the design metaphor is not the only consideration. The participants found *SHAPE*© and Authorware© - the two diagram-based HAPs to be highly consistent (items 17 and 18). Participants felt that the two map-based HAPs had the lowest amount of unnecessary detail on the screen (item 19). *SHAPE*© and ToolBook scored high on the level of *Accommodation* of HAP (items 20 and 21) but all HAPs scored lower for *Accommodation* for learners. *SHAPE*© and Authorware© again scored high on *Accomplishment* (items 22 and 23).

In terms of *Transparency of Purpose*, *SHAPE*© and ToolBook scored high on items 3-6, followed by Authorware© These results seem to follow on from the results of learnability and usability above since ToolBook uses the most accessible design metaphor and *SHAPE*© and Authorware© are both diagram-based. An interesting discrepancy is the rather low ratings achieved by Webmapper. There could be several reasons for this, not necessarily related to the design metaphor of Webmapper. It could be that Webmapper ran on an Apple Macintosh whereas all the participants were PC users, it could be because the demonstrator was less familiar with Webmapper and this had an adverse affect on its ratings. On the other hand Webmapper may possess some undesirable properties.

8.3.3 Hyperlinking

Table 7-5 in chapter seven illustrates the participants opinions on the intuitiveness of hyperlinking with the HAPs under study. Understandably, ToolBook© scores highest for the participants probably because the idea of linking between pages is familiar

whereas linking in the other design metaphors is not so accessible, i.e. there is no mental 'handle' to which to refer. For learners the situation is different and only *SHAPE*© scores 'reasonably intuitive'. Hyperlinking in Authorware© is rated un-intuitive and very un-intuitive for learners, possibly highlighting a disparity between the control flow design metaphor and hypermedia.

8.3.4 Utility

In this study both *SHAPE*© and TBM produced similarly-structured hypermedia, only that produced by *SHAPE*© was richer and more complete. The similarity in the output might be because a deliberate attempt was made to make TBM as simple to use as *SHAPE*© so that it was only the metaphor that differed. An alternative conclusion might be that subjects would structure material in a similar fashion whatever the design metaphor. This conclusion is corroborated by the two applications built with *SHAPE*© and FrontPage© which were undoubtedly different in appearance but with similar hyperlink structure. The question that therefore needs to be addressed is whether the design metaphor makes a difference to the hypermedia that is produced?

One of the underlying objectives of this study was to see how expressive a design metaphor could be without having to resort to programming. The hypermedia produced by *SHAPE*© was very basic and with little or no interactivity, any sophisticated interactivity would have to be done with ToolBook© and its programming language 'OpenScript'. The question is how can a diagram metaphor (like concept maps) be extended so that programming can be avoided while at the

same time ensuring the learnability and usability? An associated issue is how flexible an HAP is at producing a variety of applications? This question relates to Halasz's task-specificity dimension in his taxonomy of hypermedia, discussed in chapter two. The more flexible an HAP becomes, the more complex is its the functionality, and correspondingly, the lower its learnability and usability. For example, Authorware© and ToolBook© are the most functionally complex but are also rated (item 8, table 7.5) low in terms of their obviousness or *Transparency of Operation*. The question raised in the design of **SHAPE**© in chapter four remains - how is it possible to construct an HAP that is expressive and functionally-rich but which is also easy to learn and use?

8.4 Cognitive Level Issues (Objective Three)

One of the main objectives of this study was to elicit the cognitive issues for learning and using HAPs by HE staff. To achieve this objective a number of activities were instigated: (a) a Principle Components Analysis of the Ease of Learning and use scales was conducted; (b) the transcripts of the one-to-one training sessions, the transcripts of the focus group and the transcript of the interview with Greg Dainty (see sections 7-4 and 7-5) were analysed using NUD.IST to discern any emergent patterns.

A number of factors and their interactions have emerged from this analysis and have evolved during the study as recorded in previous chapters. The relationships between the factors discovered in this study and those found in the literature were considered

in chapter six. In this section an attempt is made to construct a coherent picture of the emergent factors and discuss the design implications for HAPs.

8.4.1 Cognitive Factors

Figure 6-1 in chapter six showed a mapping of Green's cognitive dimensions on to Shackel's criteria for usability. Figure 8-1 below attempts to map the factors found in this study with Green's cognitive dimensions. The closeness with which Green's dimensions map to the factors found in this study was found to be surprisingly poor and most of the correspondences shown in figure 8-1 are relatively tenuous. Only *Hidden Structure* and Green's Hidden Dependencies and *External and Internal Consistency* and Green's Consistency correspond reasonably well. It is probably fair to say that excluding *Hidden Structure* and *External* and *Internal Consistency* the factors found in this study are different to Green's cognitive dimensions. They are different in how they aggregate phenomena into factors, e.g. *Operational Momentum,* incorporates Green's Error-Proneness but embraces significantly more of the characteristics of a program. The factors are also different in that the phenomena they describe are different, e.g. Green's dimensions do not seem to have an equivalence for the more emergent properties of *Accommodation.* and although there is some correspondence between *Mental Model Match* and *Closeness of Mapping* they address mostly different phenomena. The factors described by Dix *et al.* (1998) correspond better to the factors in this study but as noted in the methodology (chapter three) Dix *et al.* provide no empirical evidence supporting the existence of their factors. Jordan *et al.* (1991) refer to guessability which roughly corresponds to *Operational Momentum* but in Jordan *et al.*'s definition it refers to the time taken to 'get going

with the system' which is vague. Seely Brown (1986) refers to system opacity, which

is the opposite of *Transparency*, but does not deconstruct it into any greater detail.

Shneiderman (1998) refers to the 'transparency principle' as:

'The user is able to apply intellect directly to the task; the tool itself seems to disappear'
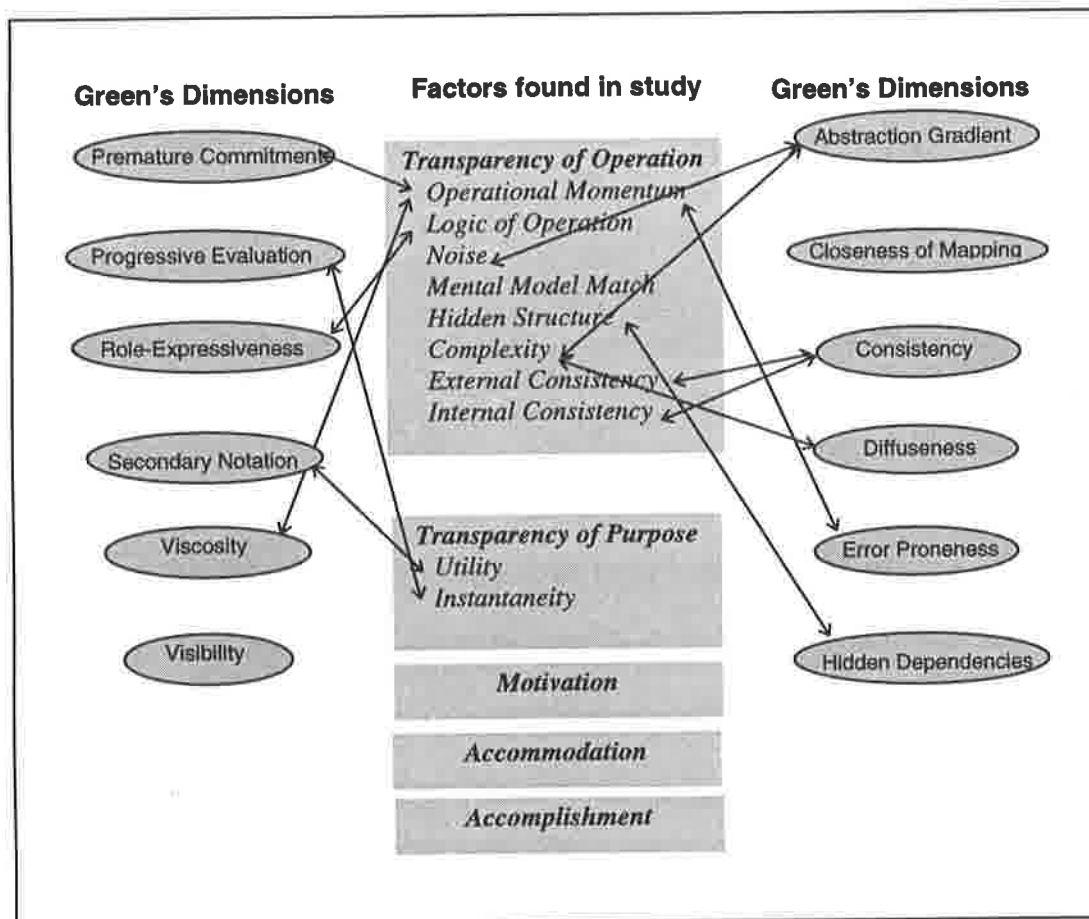
Figure 8-1 Mapping Green's (1989) cognitive dimensions to factors found in study

Thus the transparency principle has much in common with *Transparency of Operation* found in this study. Shneiderman's 'transparency' is defined holistically with no reference to empirical data. The *Transparency of Operation*, found in this

study, emerged more or less from a bottom-up analysis of the data and is therefore significantly more substantive.

There is also some resonance between the factors discovered in this study and Nielson's 'System Acceptability Model' shown in figure 2-2. However his model does not have the same level of detail and a more intelligent comparison is made with Green's cognitive dimensions (figure 8-1).

A key researcher in the field of the learning of computer programs, Carroll (1998) has recently reviewed the work contained in his influential book on minimalist design of instructions for computer programs 'the Nurnberg Funnel' (Carroll 1990). Carroll's work is primarily concerned with the principles of the design of instructions that accompany computer programs rather than a study of the principles of usability. In spite of this his work resonates with the model of learnability. Carroll reflects that:

> *'Our interpretation of our subjects' struggles was that they were actually making rather systematic attempts to think and reason, to engage their prior knowledge and skills, to get something meaningful accomplished'*

Carroll concluded that systems[10] needed to evoke positive motivation toward meaningful activities and reoriented his focus to better supporting, self-initiated sense making by the user. Hence, encapsulated in his reflections are some of the concepts embodied in the model of learnability of *Transparency of Operation* and *Purpose, Accomplishment* and *Motivation* and the causal effect of instruction on successful outcome. However Carroll's work was manifest in an explicit

---

[10] System is used here to denote the computer program and the instructional arrangements for learning it.

model, rather it was a set of empirically derived principles of design, not a causal model.

8.4.2 The Model of Learnability and Learning Theories

Figure 7-2(b) presented the causal framework of the factors of the learnability of HAPs which was empirically derived from the research conducted in this study. The model has been compared to contemporary models of usability in section 8.4.1 above. In this section an attempt is made to relate this causal model to established learning theories in order to provide extra validity for its existence. The problem with considering learning theories is that there are many of them; there is no one accepted theory; and often theories 'borrow' ideas from each other so the distinction between classes of theory become blurred. No attempt will be made here to review the learning theories, since this would be a digression in this thesis and many reviews already exist. Contemporary reviews of learning theory, e.g. Bigge (1982), Slavin (1991) accept that theories can be ascribed to one of two paradigms, i.e. behavioural and cognitive theories of learning. Behavioural theories study the external changes in the behaviour of the learner in response to external stimuli, cognitive theories are more concerned with the internal cognitive processes that take place during learning. Texts also note that other theories, e.g. Bandura (1971) Gagne (1977) are eclectic in that they contain elements of both behaviourist and cognitive paradigms.

The model of learnability incorporates two aspects of the learning of a HAP, the characteristics of HAPs that affect their learnability (*Transparency of Operation and Purpose*) and the effects these characteristics have on the user's perceptions of the HAPs (*Accommodation and Accomplishment*). *Motivation*, which is a 'property' of

the user, is a confounding factor that is affected by the characteristics of HAPs and affects the sense of *Accomplishment*. Therefore the model of learnability is not a model of the learning of a HAP but the interaction of factors that affect its learnability. Another point to make here is that since it is only the intrinsic properties of a HAP that are considered, the model of learnability is independent of instructional strategy.

In reference to behaviourist theories, *Transparency of Operation* and *Transparency of Purpose* and the encompassed subfactors can be viewed as complex 'stimuli', *Accommodation* and *Accomplishment* can be viewed as 'responses'. However in terms of instrumental or operant conditioning (Skinner, 1938), there is no explicit reinforcement or feedback to the user, i.e. HAPs are not generally instructional. The other problem of considering Behaviourism is its focus on observable behaviour. When learning a HAP there is little tangible evidence to observe that learning is taking place since the kind of learning is conceptual and oriented towards problem solving. Hence behaviourist learning theory does not offer clear correspondences with the model of learnability.

When considering the cognitivist's information processing model of human learning (Lindsay and Norman, 1977), it is the success with which knowledge about something finds its way into long term memory that is important. Hence the *Transparency* of a HAP is critical to its 'memorability'. Gagne (1977) has extended the information processing model of learning and memory into a theory of the events of learning. Gagne's theory is quite pertinent to the model of learnability in that it includes the

external conditions necessary for learning to take place. In Gagne's model, the external phases of learning are paired with the cognitive processes taking place in the learners mind plus the external instructional events for each phase. Clearly the model of learnability does not incorporate deliberate instructional events and HAPs are not instructional programs. However there is no reason why HAPs should not be imbued with the characteristics that can promote instructional events. In this sense there are some direct correspondences with the model of learnability. Figure 8-2 shows the correspondences between Gagne's model and the model of learnability. There is some clear correspondences between *Transparency of Operation* and *Transparency of Purpose* and the first five phases and associated instructional events of Gagne's model. The diagram shows where the subfactors are most salient in each of the phases but there is overlap and the subfactors act over the other learning phases encompassed by each grey block. In the model of learnability, *Transparency of Purpose* promotes *Motivation* in the same way that Gagne suggests 'activating motivation' is an instructional event that promotes the motivation learning phase. There is no direct correspondence with *Accommodation* however it is the emergent property of *Transparency of Operation* and *Transparency of Purpose* and therefore acts over the first five learning phases. *Accomplishment* corresponds to the performance phase and could be construed as supporting the instructional event 'eliciting performance'. The comparison with Gagne's model indicates that there a significant correspondence with the model of learnability. It is interesting to note that Carroll (1998), whose work on minimalist instruction was considered above, was highly critical of the systematic approach to instructional design espoused by Gagne. In relation to the model of learnability this is not important, since the model does not

present a framework or strategy for learning computer programs, it presents the causal

model of the characteristic factors of HAPs.
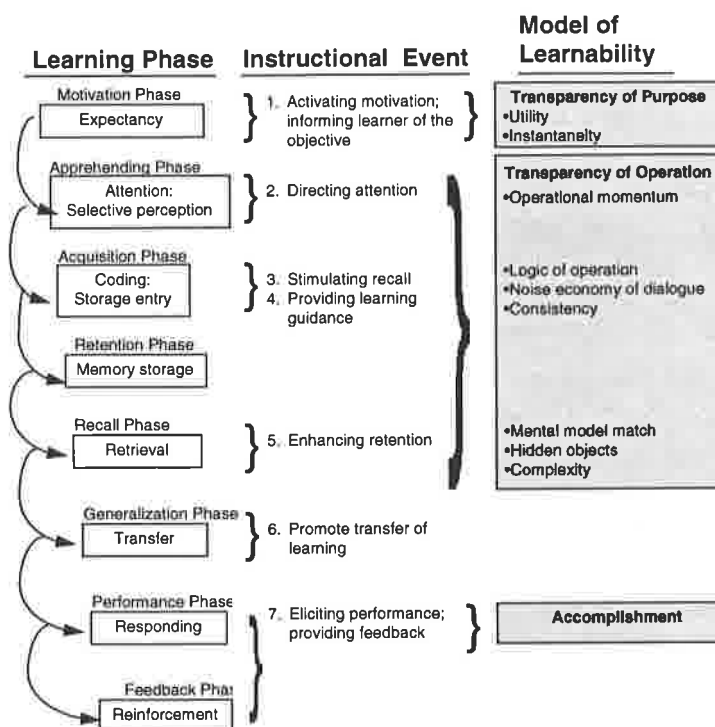
Figure 8-2 Relation between Gagne's phases of learning and the model of learnability

In addition to Carroll (1990), exponents of the constructivist of learning theory

(described in chapter two) have criticised the 'planned' approach implied by Gagne's

theory of learning for being prescriptive and not representative of the actual way

knowledge and skills are acquired. Constructivism is pertinent to the model of

learnability because it advocates the design of the learning settings but not the precise

steps by which learning will take place. This approach is reflected in the model of

learnability in that it defines the characteristics of the HAP that promote learning.

However the model of learnability does not address the important principle of the

social context of learning in constructivist learning theory (Vygoskii, 1962; Bandura,

1986). The lack of a social factor in the model of learnability is not seen as a problem since the aim of the model is on developing the characteristics of HAPs for learning regardless of the social context. The social context is clearly important and influences *Motivation, Accommodation and Accomplishment* and in time the model should perhaps be developed to take account of the social context. A complementary idea to the constructivist paradigm is the concept of scaffolding first described by Bruner (1975) and developed by others e.g. Jonassen (1996) and Jackson *et al* (1996). Scaffolding is defined by Jonassen (1996) as

> *a cognitive apprenticeship technique in which the instructor performs or supports the performance of parts of the task that the learner is not yet able to perform. When the students are stuck, suggesting a solution or performing that step for them before they give up will help them complete the task and gain confidence to try other problems.*

There are clear parallels here to the model of learnability. The factors of *Operational Momentum, Instantaneity, Hidden Objects* all support performance of the task and lead the user to a better sense of *Accommodation*. In the definition above the instructor can be replaced by the HAP and the student by the user.

Other learning theorists present ideas on the cognitive processes of knowledge acquisition - Bruner and Anglin's (1974) simultaneous processes of acquisition, transformation, check of the pertinence of knowledge corresponds with Lindsay and Norman's (1977) processes of accretion, restructuring and tuning of knowledge. Accretion is the process of filling existing mental schemas, restructuring is the process of creating new schemas to accommodate new knowledge and tuning is the process of making minor adaptions to schemas. There are no direct correspondences between these cognitive processes and the factors in the model of learnability. However, in

order to facilitate these processes in the learning of a HAP, the details of the HAP must fit as well as possible with the existing schemas of the user (*Mental Model Match*). Where restructuring is necessary, i.e. the user has no pre-existing schemas with which to understand the HAP, then the HAP (a novice for instance) must be have a high level of 'intrinsic' *Transparency* so that the user can develop a new schema easily.

8.4.3 Learnability versus Usability

The resonance between learnability and usability has been encountered in several places during this study and the question remains as to whether there is a difference. The differences are perhaps in the relative strength of particular factors. For example *Logic of Operation* is important to learning a HAP but is probably more important when the HAP is being used in earnest. *Operational Momentum* is probably more important when learning an HAP than when using it, once the HAP is well understood. *Consistency* is probably equally important but *Noise/Economy of Dialogue* is more important for learners. Likewise *Mental Model Match* is more important when learning whereas *Complexity* and *Hidden Structure* will remain a problem once an HAP has been learned. It is important in motivating learners that they perceive the HAP to have a high *Utility* but once they have learned an HAP it becomes less important. Actually getting learners to perceive *Utility* is, as already noted in previous chapters, more difficult. *Instantaneity* is an important learnability and usability property, learners and users alike must be able to continuously see that their effort results in something concrete.

Hence the holistic properties of *Transparency of Operation* and *Transparency of Purpose* will have different emphasis within them depending on whether the focus is learnability or usability. The level of *Accommodation* of an HAP will depend on whether it is being used by a learner or regular user. Hence the relative importance of the subfactors of *Transparency of Operation* and *Transparency of Purpose*.

8.4.4 What is Required for Learnability and Usability of HAPs?

The relationship between the factors emerging from this study and extant factors presented in the literature have been discussed in chapters six and seven. In this section the implications of the factors for learnable and usable HAPs is discussed. The modified interaction of factors model shown in figure 7-1 (chapter seven) indicates that *Accommodation* and *Accomplishment* are dependent on *Transparency of Operation, Transparency of Purpose* and *Motivation*. It is difficult to say whether excluding the other factors described here, there is anything extra about an HAP that increases a sense of *Accommodation* and *Accomplishment* in its users. *Accommodation* is the holistic effect of *Transparency of Operation* and *Transparency of Purpose* which are, in turn, the holistic effect of the other factors. *Accomplishment* is the holistic effect of *Accommodation* and *Motivation*.

It is interesting that *Motivation* emerges as a factor (in chapter seven) since it was a dependent and independent variable in the analysis of the evaluation of *SHAPE*[©] presented in chapter five. The only significant finding in the evaluation of *SHAPE*[©] was that *SHAPE*[©] had a higher Motivation to Continue than TBM. There were no significant correlations with either the intrinsic or extrinsic motivation ratings of the

subjects and their motivation-to-continue with either of the HAPs. The re-emphasising of the importance of *Motivation* by the participants in the focus group indicates that perhaps its study in the learnability and usability of HAPs needs to be examined further.

General research on human motivation is extensive and well beyond the scope of this study. The interest in *Motivation* here is on whether the intrinsic properties of a HAP can motivate users to develop applications and not on external factors related to its users. From figure 7-2 it is clear that a HAP which embodies good values of the factors will result in higher *Motivation* although some are more important in imbuing *Motivation* than others such as *Utility* and *Instantaneity*.

The emergent properties discussed above are therefore largely dependent on the more intrinsic properties of the HAPs. Table 8-1 below is therefore an attempt to consider what can be done to HAPs to improve these factors of HAPs for learnability and usability.

Table 8-1 Possible improvements to HAPs to increase learnability and usability based on the factors found in the study.

Factor	Possible improvements in HAP
Logic of Operation:	It is vital that learners understand the *Logic of Operation* of an HAP prior to actually using it. This point was reinforced with the issue of training still being seen as important regardless of the HAP. Understanding the *Logic of Operation* is dependent to some extent on the familiarity of the design metaphor to the user, i.e. *Mental Model Match.*
Operational Momentum:	Possible improvements include, 'signposting' or providing 'scent trials' or even including a degree of intelligence into an HAP which can anticipate what a user is likely to do next.
Consistency:	There are already guidelines available on maintaining *Consistency.* (e.g. Dix *et al.,* 1998; Shneiderman, 1998). In this study a distinction was made between *Internal* and *External Consistency.* How far can one make HAPs externally consistent is debatable. Most Windows users will be familiar with the standard Windows 'WIMP' environment and will be able to search the menus and toolbars for whatever function they are looking for, but that is probably as far as *External Consistency* impinges on the learnability of the HAP. *Internal Consistency* is more involved and means

	making sure all functions adhere to the design metaphor. It is important to balance full implementation of a metaphor for usability but avoiding overuse and over complication of the interface.
Noise:	One of the reasons for the positive evaluation of *SHAPE*© was the use of a very sparse interface compared to the other implementation of the concept map design metaphor Webmapper which received a rather poor evaluation. A large number of GUI-based programs allow users to switch off various parts of the interface but for learners this should be the default. Another feature of modern GUI programs is the tendency for having a multitude of icons in toolbars crowded around the main workspace. For learners it is important to keep transactions to a minimum. A complexity gradient is required which could be either user driven or via an intelligence-based HAP.
Mental Model Match:	Mental models in HCI have been considered important for a long time (Jih and Reeves, 1992) and has resulted in numerous factors that could be taken into account in a user's mental model of a HAP. In the case of the findings of this study HAPs must be able to 'accommodate' HE academics with relatively low Computer Skill and for diagram-based HAPs, users with a range of Spatial Relations ability. The term loose model was introduced in chapter six to indicate the various facets that go to make up a user's mental model. The design metaphor will be a key issue and it is interesting

	to speculate on what design metaphors HE academics would find amenable. HE academics are used to structuring their ideas/knowledge so it would seem sensible for HAPs to respond to generic characteristics of HE academics.
Hidden Structure:	Green refers to 'hidden dependencies' in his cognitive dimensions framework. Hidden dependencies seems to be contained within the idea of *Hidden Structure*. The *Hidden Structure* might actually be external to the HAP, e.g. knowledge of the WWW in the case of FrontPage© and Webmapper, in which case there is a requirement that the HAP conveys the need for additional understanding to learners. It is difficult to provide learners with a complexity gradient without at the same time hiding some of the HAP's structure. The issue is balance; where possible the structure of a HAP should be made as transparent as possible but only when it creates confusion. This difficult requirement seems to emphasise a need for some degree of intelligence in the HAP.
Utility:	All commercial HAPs are generalist in nature in order to accommodate a large range of possible uses. This will mean that the *Utility* of HAPs for educational purposes will be lower than is possible. What is required are HAPs which respond specifically to the needs of HE academics and their teaching courses. Some initiatives in HE, like the TLTP programme, have resulted in tools

	(not necessarily HAPs) designed specifically for HE purposes like CASTLE (CASTLE, 1998). However, as noted in chapter two, it is important that HAPs do not become prescriptive when designed for a particular area of application. Section 8.5.3 below discusses the functionality required for HE below. It is more important for users that the HAP ensures that they can see that their efforts correspond to something they think they want - as noted above, i.e. *Transparency of Purpose* is probably more important than knowing the HAP can do what they want.
Instantaneity:	Intermingled with *Utility* and *Transparency of Purpose* is *Instantaneity* which simply means that all actions of user which result in a changes in the end application should be visible to that user immediately. Another facet of this property is the speed with which a HAP can make visible what an end application will look like. Subjects were motivated with *SHAPE*© because they could see 'results' immediately. Participants of the focus group concluded the same about Authorware© suggesting that instantaneity can be improved with a diagrammatic design metaphor as noted above.

8.4.5 Knowledge, Knowledge Construction and Representation

The issue of knowledge has been threaded throughout this study but particular attention has been paid in sections 2.3.4 (chapter two), section 4.3 (chapter four) and section 7.2.2 (chapter seven). The questions raised in chapter two and three can be listed as follows:

- How can different types of knowledge be modelled, particularly procedural and declarative for different HE domains?

- How useful is the visualisation of knowledge from the perspective of hypermedia?

- How is knowledge best represented for higher educational purposes bearing in mind the significance of newer ideas on learning like constructivism?

- How effective are concept maps in addressing the above questions?

The quantitative analysis of the comparison of *SHAPE*$^{©}$ and TBM showed that subjects could produce richer hypermedia with *SHAPE*$^{©}$, it also showed that the Ease of Use rating of *SHAPE*$^{©}$ was largely independent of Computer Skill and was significantly better than the rating for TBM. The emphasis in using *SHAPE*$^{©}$ was on the semantics of the relationships with less distraction from the linking mechanism which was more of a problem with TBM. These benefits have to be tempered by the correlation of the richness of the hypermedia produced with *SHAPE*$^{©}$ and the Spatial Relations ability of the user. The implication of this correlation is that for some subjects, spatial representations of knowledge may not be the most appropriate or effective metaphor for knowledge construction. The problem of determining which metaphor is best for each staff member is difficult to solve and the answer may be that

a number of ways of constructing the representation of a knowledge domain must be incorporated into a single HAP. In the light of this, the movement towards spatial representations of knowledge, including three-dimensional representations, e.g. (Fowler *et al.*, 1997; Foley 1995; Chen and Czerwinski 1998), needs to be tempered. The analysis of the focus group and the desk-top study of the selected HAPs each highlighted the inadequacy of HAPs (whether diagram-based or not) to represent the required knowledge. The HAPs were also found to have little functionality to allow users to 'explore' the knowledge by zooming in and out, switching between views, etc. Another problem already mentioned was the inability of the HAPs to be able to represent different knowledge types like procedural, declarative, etc. The content analysis also exposed the problem of the idiosyncratic representations (where subjects were trying to represent their own personal perspective), a phenomena noted by McAleese (1987) in his evaluation of KIM (a concept mapping tool for knowledge construction). Linking in **SHAPE**$^{\copyright}$ was also sometimes idiosyncratic. Even when subjects knew there was a clear relationship between two concepts, (e.g. between faculty management and staff), the choice of label for this relationship was surprisingly diverse. However, from the perspective of the resultant hypermedia, it did not seem to matter what the label was since the resultant hypermedia produced by the subjects was very much in accordance with the standard model of the hypermedia.

8.5 The Design of Future HAPs (Objective Four)

In this section an attempt is made to consider what a next generation HAP should be like for use by HE academics, based on the findings of this study. The section will consider the design metaphor, the best way for the HAP to be used to provide the required functionality for use in HE and then will conclude with an outline requirements specification. Consideration of the design of HAPs that address the factors raised in this study has already been made in section 8.4.4 above and will therefore not be reiterated here.

8.5.1 Drawability in the Design Metaphor

Section 2.7 in chapter two debated the merits of various paper-based techniques for representing ideas/knowledge. Concept maps were chosen over the other techniques simply because of their similarity to the structure of hypermedia (Jonassen, 1990a; McAleese, 1987). These studies have shown that concept maps have some benefits over the other design metaphors already used in HAPs. There was, however, a disparity in the ratings of the two map-based metaphors, **SHAPE**© and Webmapper, the conclusion being that it is not just the metaphor that needs to be taken into account - the factors that emerged from the analysis also need to be addressed. The reason the use of various paper-based techniques like Mind Maps (Buzan, 1993) were considered was because they are believed to be easy to draw, relatively unconstrained and have minimum cognitive overhead associated with their use. The design metaphor of a new HAP should maintain and improve upon this principle of 'drawability'. However Dillon *et al.* (1993) warn of the fallacy of thinking of hypermedia as a physical space

not a semantic space, a point borne out in this research where a variety of names were given to the identical link and where some subjects found difficulties in representing some of their 'semantic space' in a physical dimension. Another issue that Reader (1995) found in his work was the need for some constraint in the syntax of concept mapping tools. Having said this, physical representations, although falling short of being true representations of human knowledge are at least pragmatic in enabling a usable product to be authored. The results of an evaluation of their DOLPHIN system for group work demonstrated the advantages of hypermedia to collaborative work (Mark *et al.*, 1995). Unfortunately, they do not appear to comment on the user interface of DOLPHIN which allows users to 'scribble' their ideas on to the screen, and presumably this aspect of DOLPHIN must have played some part in its positive evaluation. A more complete account of a computer system that 'recognises' user scribbles and turns them into a set of objects from a database is given by Gross (1994) but both systems point to one of the possible ways forward for diagram-based authoring. What is called for is the development of a diagrammatic language that combines domain visualisation and the provision of functionality for use in the design metaphor for a HAP design for use by HE teaching staff. The design metaphor needs to be enhanced to allow:

- Free-hand scribbling of ideas.

- The defining of types of concept and links.

- Linking at different levels in a hierarchy of diagrams

- Enable functionality to be associated with the concept map knowledge representation.

- Enable other knowledge types like procedural and temporal knowledge to be represented.

8.5.2 The Need for Two-Phase HAPs

It was clear from the analysis of the focus group session that there needs to be two distinct phases for creating hypermedia for use in HE: (i) knowledge and content construction; and (ii) education-functionality construction. These two phases are likely to be iterative but they need to be identified in the HAP separately. This separation of phases is consistent with the now widely-accepted Dexter model of hypermedia architecture (Hardman *et al.*, 1994) and extends the 'Aworld and 'Kworld' idea of McAleese (1987). The design metaphor must be expressive enough in syntax and symbology to support both knowledge and functionality construction but also maintain high learnability and usability. Figure 8-3 gives an illustration of how this new HAP would relate to existing HAPs. *SHAPE*© has some ability to express a knowledge domain but by itself, it is low on functional power, whereas conventional HAPs have reasonable functional power but very little means of expressing a knowledge domain. The new HAP would have the ability to both express the domain and have high functional power.
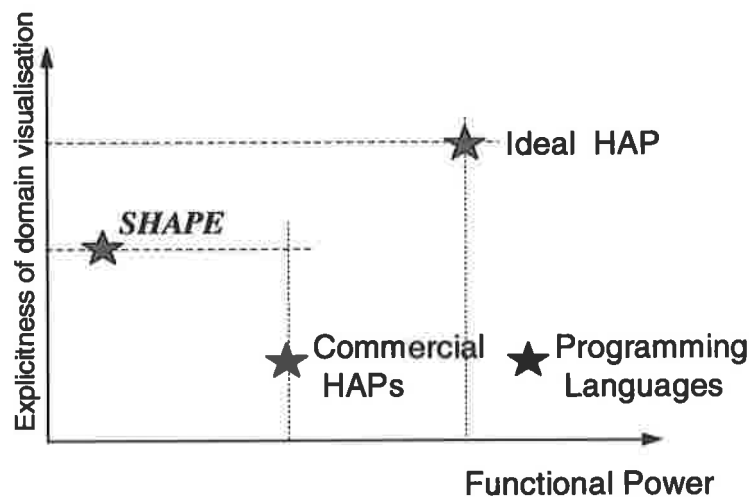
Figure 8-3 Taxonomy of HAPs based on their ability to express the knowledge of a domain and their functional power.

How a knowledge representation schema should be integrated with a functional representation schema raises some interesting research questions. For example, it would be necessary to consider how closely the knowledge representation and the functional representation should be coupled. The problems that need to be resolved are the design of the symbology and syntax of the diagrammatic language, and the mechanisms for integration knowledge representation and functionality.

8.5.3 Functionality Required for Higher Educational Use

Although educational efficacy is the paramount goal of hypermedia produced by HAPs (this was considered in detail in chapter two), it has been largely omitted in this study because the focus was on the cognitive level issues of hypermedia authoring by HE teaching staff. However, two real applications were built using two of the HAPs in this study and they were actually used in the teaching-learning process although

evaluation of their educational efficacy is part of another study (Dainty, 1997). The main questions raised on hypermedia in education in the literature review (chapter two) is how should hypermedia be used in teaching-learning processes and what functionality should it possess? Russell (1990) makes a plea for campus-wide, integrated, highly functional hypermedia that provides learners with rich learning environments. Furthermore, as noted in chapter two, in their taxonomy of educational hypermedia, Hutchings *et al.* (1992) recommend that educational hypermedia needs to move from passive to engaging and from an essentially presentational medium to a creational one. The stance this study takes is slightly different, there is a place for all types of hypermedia - passive hypermedia has its own educational value, provided it is used in teaching-learning processes effectively and there is much scope for creating engaging hypermedia which allows students to supplement or create their own perspectives. The majority of hypermedia which is used in HE at present is Web-based and passive but nevertheless still has reasonable educational value. There are only a few examples of hypermedia that engage learners and allow them to supplement or create their own perspectives in HE in the UK (e.g. CLASS, 1994). Constructivism seems to provide an underpinning philosophy to the moves in HE towards increased student-centred learning, whether it is taking place for economic or ideological reasons. Jonassen *et al.* (1994) presented their 'manifesto' for a constructivist approach to using technology in HE and suggest a number of ways in which technology can be used in HE. This was covered in the literature review, (chapter three). What is pertinent in the manifesto to the design of a new HAP is the nature of the functionality that should appear in the end product. Jonassen *et al.* (1994) suggest that technology should:

- Provide environments that represent multiple realities.

- Provide real world, case-based tasks and environments.

- Provide collaborative knowledge construction environments.

- Provide tools that support active learning and observational learning.

These requirements are broadly in line with the list of features that a HAP should support as suggested by Elliott *et al.* (1995a) and given in chapter two. The key research question is how the above functionality could be provided in a diagrammatic language. The functionality must in some way be encapsulated as symbols and then the interaction between the knowledge representation and these functional symbols or objects needs to be resolved and defined. The danger will be, as always, maintaining expressiveness, learnability and usability and avoiding over prescription in the design of the functional objects.

The development of **CLICK-IT©** raised an interesting phenomena in the links in the concept maps. Some of the links were attempts at modelling the knowledge but some were really structural links for the benefit of students which seemed to be a good attribute of *SHAPE©*. Constraining the syntax of diagram-based HAPs might limit the flexibility to represent knowledge and provide structural cues for students.

8.6 Critique of the Study

As noted in the introduction although there has been extensive work done on the usability of hypermedia (e.g. Nielson, 1990c; Garzzotto and Matera, 1997; IJHCS,

1997), there has been little or no work done on the usability of HAPs (Nielson, 1990) . This study has therefore experimented with new ideas for the evaluation the usability/learnability of HAPs. In the process of designing and executing this study important findings on the methodological approaches to HAP usability measurement have been made. Nielson (1990a) makes the point that evaluating the usability of HAPs has probably been avoided because it is harder than evaluating the usability of hypermedia.

8.6.1 Scope of the Study

As noted above, a decision was made at the outset that the educational efficacy of the hypermedia produced by the HAPs in this study would be left out because the emphasis was on the cognitive and usability issues of hypermedia authoring. This standpoint could be criticised as an indication that the study was incomplete. However, it was decided that an evaluation of the educational efficacy of the hypermedia produced by the HAPs would not contribute much to the investigation of the cognitive and usability issues, perhaps only in terms of the perceived *Utility* of the HAPs. Since the study focused on hypermedia in HE the sample group were all HE academics. The only problem with this is that the results could not then be generalised to the whole potential user population of HAPs, however it was decided this was not the concern of this study.

8.6.2 Methodology Choices

The decision was made to utilise a number of approaches to achieve the objectives of this study. The criticism could be made that this was simply a 'shot gun' approach

used in an attempt to ensure some useful data was gathered. However, it is argued that with a number of objectives and because it was anticipated that the evaluation of the usability of HAPs would be difficult, a multimodal approach was desirable to ensure all aspects of the learning and using of HAPs were captured. The design of the methodology was progressive, in that each activity in the methodology informed the next activity and this proved to be very fruitful in discovering important principles related to learning and using HAPs.

Approaches to Learnability and Usability Analysis

There are extensive references in the literature to methodologies for usability analysis and engineering (e.g. Nielson, 1993b; Gazotto and Matera, 1997; Bailey 1993) and, as noted in the design of the methodology, they cover the whole gamut of possible research approaches. There is a distinction between qualitative and quantitative techniques with some researchers favouring one over the other. It was decided that there were benefits in utilising both approaches and it is believed that this has been productive. The quantitative results from the one-to-one training sessions gave a sound basis upon which to structure and examine the qualitative data. The case against qualitative approaches is the lack of legitimacy and the case against the 'scientific' approach is the artificiality and inability to capture the human dimension. Therefore the recommendation from this study is that both approaches should be used in usability studies.

The factors that have been used in the evaluation of the map-based design metaphor and the other HAPs emerged from the study itself rather than from factors cited in the

literature. This was an important stance because the factors are not manufactured but truly relate to the study, following the ideas of Kolb's experiential learning cycle and of grounded theory and Gunn's idea (1996) of 'situated evaluation'. There were two other reasons for this:

a) some of the factors described in the literature, for example Dix *et al.* (1998) have no empirical evidence to support their existence; and

b) new factors might emerge which could be integrated with the existing body of factors cited in the literature.

One variable seen as important by others (Jordan *et al*, 1991) is the measurement of time, - that is, the time to learn something or to undertake an activity. In this study time was considered to be unimportant. The time scale for learning or using a HAP is probably weeks not seconds or hours or even days and will vary between users so time analysis would probably be meaningless and difficult to measure either legitimately or accurately. Time measurement could however be important in measuring certain features of the HAP that increase *Operational Momentum* or *Instantaneity* for example, and would form part of future studies.

It was found that subjects could reflect on their experiences after the one-to-one training sessions (critical incident) but subjects reflected little during the sessions (think aloud). The conclusion from this study is that think aloud is an inefficient means of collecting qualitative data for HCI studies because users cannot reflect and work at the same time.

Analysis of Qualitative Data

Although there are many references to qualitative approaches to usability studies there is almost no reference to how the qualitative data is analysed. Often qualitative data is used merely to find a few 'convenient quotes' that support preferred theories. This study adopted the 'hierarchical categories' technique developed by Richards and Richards (1994) embodied in NUD.IST and used extensively in sociological studies for the rigorous analysis of qualitative data. This approach has been fruitful and has enabled a greater degree of validity to be placed on the findings of the analysis. The other benefit of hierarchical categories is that data from any part of the study can be integrated into the analysis. Hence, the discussions from the one-to-one training sessions could be correlated with the building of the two applications and the focus group work, even quantitative findings can be integrated with the choice of categories in the analysis. Therefore, the recommendation from this study is that the use of hierarchical categories in usability analysis be encouraged.

One-to-One Training Sessions

The execution of the one-to-one training sessions was successful with a number of significant findings. Possible criticisms of the design and execution of the sessions could be:

- Lack of a control group
- Variation in location of the sessions

- Impartiality relationship of the experimenter to the subjects

- Size and constitution of sample

- Authenticity of the task

One further consideration was the decision to separate Ease of Learning from Ease of Use and whether it was realistic to expect subjects to distinguish between them since they only received one training session. The results also pointed to this problem in that Ease of Learning and Ease of Use were highly correlated. Perhaps the Ease of Use metric should be re-applied when subjects have used the HAPs for some time. It was impossible to design the experiment with a control group because none of the relevant variables had both a pre- and post-test value. As noted in chapter three, the two HAPs acted as controls for each other in that the Order of Use was also a variable. A legitimate argument could be made for the variation of the location of the sessions. In many rigorous psychological experiments the environment is kept constant. It is argued, however, that one-to-one training sessions were located in the work-place of each subject and hence was effectively constant. It is argued that environmental changes in this study were insignificant. The one-to-one training sessions were conducted by the researcher who also built the two HAPs and worked with all the subjects. There was obviously a danger that the researcher and the subjects would be influenced by this. However, prior knowledge of these facts ensured the researcher made every effort to remain objective and also the same script was used for each session. There are no real guidelines on the sample size other than it should be as large as is practical. If the sample size had been greater than 16 it is unlikely there would be any other significant findings. In his study of the use of concept maps by learners, Reader (1996) used samples of 11. The constitution of the sample was

uneven in gender terms, hence no differences for gender could be evaluated legitimately. Further studies with a better gender balance are recommended. The subjects were all from one HE institution; hence, it would be desirable to extend the study to other colleges although again it is believed that this would not change the results. Perhaps the study could be extended to different user groups however this is not within the remit of this current study. Increasing the sample size would also make the logistics of qualitative analysis very difficult. The choice of task was inevitable in one sense because of the need to control for it and also allow an analysis of the output from each subject. Perhaps a new study (based on a new HAP) in which subjects built hypermedia applications pertinent to their teaching would be the best place to implement a study of this nature.

Measurement of variables

A number of scales were developed and used in this study to measure the dependent and independent variables. In the light of the findings the scale for Ease of Learning and Use should be modified to take account of the factors like *Operational Momentum*, etc. The Computer Skills Metric for Windows (Elliott *et al.*, 1997) - details reproduced in appendix A - needs to be updated to included Windows 95 or 98. Whether a generic measure of Computer Skill could be developed is another interesting research question. The measurement of Spatial Relations ability was satisfactory and seems to be more sophisticated than in other studies, for example Chen and Czerwinski (1998), who used simple paper folding to measure spatial ability. The Spatial Relations ability battery of the Technical Skills battery

(Psychological Corporation, 1995) required reasoning with images which is more in keeping with the tasks required in this study.

Use of Focus Groups

The use of focus groups as noted in the methodology is a relatively new development in usability analysis although expert evaluation is not. The analysis of the focus group data in this study has proven to be fruitful although the validity of the findings possibly need to be corroborated with further studies for completeness. The focus group in this study consisted of experts, i.e. IT trainers of staff and developers of educational hypermedia. It might be fruitful to conduct a similar focus group of ordinary HE teaching staff. The problem with using non-experts in the evaluation of HAPs, as discovered in the analysis of the one-to-one training sessions, is that subjects felt 'unqualified' or simply unable to say anything about the value of the HAPs for creating hypermedia. This study recommends the use of focus groups but care must be taken in the choice of participants and focus groups should be used in conjunction with other studies in order to triangulate and validate results.

8.7 Research Questions

This study has evaluated HAPs and their use by a selected group of HE academics. The analysis of the results of the evaluation have raised a number of new research questions which can be categorised under the headings of factors and diagrammatic language.

8.7.1 Factors

The factors discovered in this study lead to a number of new investigations:

- an examination of possible interface design characteristics and functionality of HAPs that could improve the 'value' of each of the factors;

- constructing the design metaphor of HAPs to be flexible enough to respond to a number of different types of user;

- development of a taxonomy of the characteristics of HE users of HAPs that are important in the design of HAPs; and

- development of a model of the interaction between the design of HAPs and user motivations based on the one presented in this study.

8.7.2 Development of a Diagrammatic Language

This study recommends the development of an HAP based on a new diagrammatic language The problems that need to be resolved are:

- reconciling richness of expression with simplicity and understandability;

- mechanisms for integrating knowledge and functionality;

- development of primitives that provide the basis for a constructivist learning approach in HE; and

- the learnability and usability of this new HAP in HE.

Chapter Nine

Conclusions

9. Conclusions

Suddenly Christopher Robin began to laugh ... and he laughed ... and he laughed ... and he laughed. And while he was still laughing - Crash went the Heffalump's head against the tree-root, Smash went the jar, and out came Pooh's head again...

Winnie the Pooh, A. A. Milne.

In this chapter the evaluation of each of the objectives of the study is addressed in turn. The problems and issues of hypermedia authoring in HE that were discussed in chapter two are reconsidered in the light of the evaluation of the objectives of the study. The chapter concludes with suggestions for further work and research.

9.1 Objectives

The objectives of this study arose as a result of an analysis of the problems that beset hypermedia authoring in HE.

9.1.1 Diagram-based Hypermedia Authoring - Objective One

Objective one arose as a potential solution to the problems of hypermedia authoring in HE, namely, the of lack of time to learn a HAP and then develop suitable hypermedia. The solution considered was the use of diagramming techniques like concept maps, fish bone diagrams, mind maps, etc. as the design metaphor for HAPs with a decision to use concept maps. The evaluation of *SHAPE*© alongside TBM, the building of an application and then the comparison with other HAPs in the focus group has shown that concept maps offer advantages over other design metaphors for HAPs. Using

concept maps in a HAP has the potential to create usable hypermedia. Concept maps could form the basis of a diagrammatic approach to hypermedia authoring that can be easier to learn and use than other design metaphors used. Concept maps on their own are perhaps not sufficiently expressive but could be supplemented and modified to enable the easy creation of educational hypermedia.

9.1.2 Evaluation of HAPs - Objective Two

Objective two extended objective one to other industry standard HAPs and were primarily addressed through the focus group and desk-based study. The design metaphor used in HAPs is important in facilitating learnability, usability and utility. The design metaphor must be expressive enough to allow the knowledge construction aspects of authoring. Diagram-based metaphors seem to have advantages in terms of their learnability and usability although the accessibility of the book metaphor cannot be ignored. Concept map-based design metaphors seem to offer advantages over other metaphors although the precise implementation is also important. None of the metaphors evaluated in this study possessed a high degree of *Operational Momentum* so further work is necessary to examine ways of improving it. Some metaphors are unsuitable to hypermedia, for example the control flow metaphor of Authorware©. Further work is required to examine ways of improving the *Accommodation* of the HAPs

9.1.3 Cognitive Level Issues - Objective Three

Objective three arose from considerations of hypermedia authoring and human-computer interaction issues and has been addressed through all parts of the study. The

factors of learnability and usability that have emerged have some correspondence with ones found in the literature particularly Green's cognitive dimensions. The relative importance of the factors varies depending on whether the HAP is being used by a novice or experienced user. *Transparency, Accomplishment* and *Accommodation* are holistic factors whose value is dependent on the relative values of the sub-factors from which they derive. None of the HAPs in the study were well-designed in all the factors, however the concept map-based design metaphor of **SHAPE**© seemed to be better than the other HAPs examined. This fact however must be tempered by the poor rating of the other concept map-based design metaphor Webmapper. A number of improvements can be made to HAPs addressing these factors to improve their learnability and usability. It would be desirable to have objective measures of learnability and usability based on the factors uncovered in this study.

Knowledge construction and representation can be made easier with diagram-based design metaphors for users with a wide range of computer skill but is affected by the spatial relations ability of users. Enabling better ways of construction knowledge can result in some idiosyncratic representations although they do not seem to have much of an adverse affect on the resulting hypermedia. Knowledge construction with existing HAPs is limited and do not allow a range of knowledge types to be represented, neither do they allow users to manipulate views of the knowledge.

9.1.4 The Design of Future HAPs - Objective Four

Objective four is effectively the conclusions of the previous three objectives. The results can be best summed up in terms of a requirements specification as indicated in the section below

Outline Requirements Specification

Based on the previous discussion of the results, it is now possible to draw up a requirements specification for a new HAP designed to increase usage by HE academics.

- The HAP should implement a diagram-based design metaphor. possibly an enhanced concept map metaphor. The enhancements should not jeopardise *Transparency*.

- The diagram metaphor should be as close as possible to 'drawing on paper' as is practical.

- The HAP should also partially implement other design metaphors to provide for variation in users, i.e. *Mental Model Match*.

- The design metaphor should enable the design process to be separated into two phases - a knowledge construction phase and a function construction phase.

- The syntax and symbology of the design metaphor should allow for a variety of knowledge types to be represented.

- The syntax and symbology of the design metaphor should allow for functional objects to be integrated in to knowledge representation.

- The functional objects should correspond to good instructional design principles that assume constructivism as a pedagogical basis.

- The HAP should wherever possible implement facilities like signposts that increase *Operational Momentum.*

- The interface design should be based on a complexity gradient and as a default the initial interface must have the bare minimum of *Complexity* and *Noise.*

- The HAP should warn first time users of the need for background understanding (like the Web) and provide it if necessary (*Hidden Structure*).

- Where ever possible the HAP should provide the user with knowledge of addition detail when they undertake an activity. This may be rely on some intelligence in the HAP(*Hidden Structure*).

- The HAP should at all times enable the users to see the effect of their actions on the end product (*Instantaneity*).

- The HAP should provide as rapidly as possible an indication of the end result (*Instantaneity*).

- The implementation of the design metaphor should mean that users will see that the operation of the HAP is *Logical.*

- There should be more than one means of creating links, e.g. diagrammatically, via a dialogue box, and code. The linking mechanism should be as easy as possible so that users can focus on the nature of the link and not on the mechanics of the link.

- Users should be able to define their own types of link if they wish, e.g. navigation, semantic, interactive.

- Users should be able to manipulate the viewing of the knowledge construction so that they can zoom in and out, construct their own views and have multiple views, etc.

- The HAP should adhere as closely to the Windows environment as possible and effort must be made to ensure that all facilities and functions operate in a similar fashion (*Consistency*).

- The HAP should have transparent interconnectivity to <u>any</u> standard content tools like word processors, video and sound editing tools, animation tools, etc. This is the model adopted in FrontPage$^{©}$ and is seems to becoming a standard means of construction Web documents.

9.2 Problems and Issues - Re-appraised

In chapter two the problems and issues surrounding hypermedia and hypermedia authoring in HE were laid out. In this section the relevance of the findings of this study to the problems and issues are considered.

9.2.1 Hypermedia

The problems and issues of hypermedia were encapsulated by Halasz's revised list of issues (1988, 1991) and categorised under technology and market-related issues. Educational issues of hypermedia have been considered by many researchers but typically Hall *et al.*'s (1992) taxonomy of educational hypermedia encompasses the direction in which it should develop. None of the objectives of this study were designed to address the fundamental nature of hypermedia.

9.2.2 Hypermedia in HE

There are many problems and issues surrounding the use of hypermedia in HE but the key problem (as noted in section 2.5) must be time - time to learn and time to develop hypermedia learning materials. This study has shown that a HAP based on concept maps can be easier to learn and use for HE teaching staff with a range of computer skill than other HAPs. The idea of separating the authoring process and also the HAP into two phases, knowledge construction and functionality supplementation, would reduce the cognitive overhead of using existing HAPs where these phases are done simultaneously. In addition to the problem of time is the need for educational hypermedia that is constructivist in nature, consistent with Hall *et al.'s* plea for hypermedia that is engaging, learner-controlled and supports learners in synthesising ideas and concepts. A new HAP based on the requirements specification developed in this study (given in section 9.1.4) would help to address the problem of the educational effectiveness of hypermedia. However this requires further work in the nature of the additional symbols and syntax that would result in educational hypermedia based on constructivist design principles.

9.2.3 Hypermedia Authoring

Section 2.5 summarised the problems and issues of hypermedia authoring in HE and are reconsidered in the paragraphs below.

The problem of expressiveness of HAPs versus learnability and usability could be considered to be intractable and the two goals mutually exclusive. In this study an attempt has been made to find a solution. Various design metaphors have been

evaluated and some useful conclusions have been drawn. Diagram-based design metaphors can be easier to learn and use than other design metaphors, even the 'book'. The concept map metaphor on its own is not expressiveness enough to produce educationally effective hypermedia and needs to have a richer syntax and symbology. The problem is the degree to which the concept map metaphor is enriched and it may be a case of compromise or having a 'complexity gradient' which moves from simple concept maps to highly syntactically enriched ones. This debate leads to another conclusion of this study that HAPs need to allow knowledge to be structured and explicitly represented and also allow functionality to be defined. This distinction is manifest in the difference in emphasis between the concept map metaphor and the iconic metaphor of Authorware©. Therefore, the conclusion is that authoring should consist of two distinct phases, knowledge construction and functionality construction and be should be incorporated in the design of new HAPs. Another tentative conclusion was that HAPs might implement a range of design metaphors to match the mental models of a range of users. The task of implementing such an impressive HAP might be economically un-viable. Concept maps allow the construction of structural or declarative knowledge, richer syntax and symbology should allow the construction of other knowledge types like procedural or even temporal knowledge.

9.2.4 HCI and Usability Issues of Hypermedia Authoring

Section 2.4.5 listed the problems and issues of the current state of the art in terms of learnability and usability. HCI is an extending and expansive field and this study has only addressed a small part of current work. The debate of whether studies should be

at the task level with its reductionist perspective or the holistic level which takes into account all human factors was presented in the methodology chapter three. This study has taken the stance that the holistic approach is preferable and the results of this study has shown it to be productive. Another issue raised in section 2.4 is that of a model of human learning of computer-based artefacts. The model of the interaction of factors that has evolved through this study and presented in figure 7-2(b) could perhaps form the foundation of a model that could be utilised and tested for its validity. One of the other issues listed in section 2.4.5 is that of the factors which need to be taken into account when designing for usability. In this study a number of factors have emerged from the studies which seem to be different from those of Green (1989) but similar to those of Dix *et al* (1998) but who provide no empirical evidence. It is argued here that some of these newly identified factors are new and where they are similar to others given in the literature they give evidential weight to their existence where there was none before. In terms of learnability *Operational Momentum* and *Instantaneity* have particular relevance.

Usability Evaluation

Nielson (1990a) noted the difficulty of evaluating the usability of HAPs and this was certainly found to be the case in this study. Section 2.4.5 listed the need for a framework for evaluating HAPs and it is believed this study has demonstrated the form that that framework should take. In brief the framework needs to:

- take into account the critical characteristics of the user group like computer skill;

- utilise more than one usability evaluation technique (which should include both quantitative and qualitative methods);

- focus on factors that emerge from the evaluation - 'situated evaluation' (Gunn, 1996) rather than pre-existing ones from the literature; and

- integrate findings using tools like NUD.IST.

9.3 Major Research Contribution

This study has attempted to address the problems of developing educational hypermedia in HE. Much of the work found in the literature on educational technology/computers in education focuses on the evaluating applications in teaching and learning processes from the student/user perspective. This study has examined an area which is relatively neglected in the literature - the HE teacher/academic perspective and the cognitive issues surrounding the learning and using of HAPs. The nature of the study means that it was multidisciplinary - computing, educational psychology and HCI. The study makes contributions to each of these areas, i.e. the recommendations on the design of HAPs for combining knowledge and functionality, the use and development of a diagrammatic language, the design and (partial) evaluation of hypermedia for learning incorporating hierarchical concept maps and the cognitive factors in the human-computer interaction of using HAPs. Another contribution of this study is the proposed solution to the problem of the skills and propensities of HE teachers/academics to develop educational hypermedia. The evaluations in this study point to the specification of a new HAP that would mitigate against the skills problem and the cognitive factors that need to be taken into account in its design for maximum learnability and usability.

However, the major research contribution of this thesis is in the field of usability engineering and the development of a causal model of learnability (figure 7-2b) derived from empirical observations. Contemporary models do not provide a theoretical framework for learnability, with most referring to learnability only as the time required to learn some aspect of an interaction with a computer (e.g. Preece, 1994; Schneiderman, 1998; Faulkner, 1998). Section 8.4.1 compared the model of learnability derived from this study with key contemporary models of usability.

As noted in 8.4.1 most contemporary models of usability are not causal, i.e. they simply list the factors that affect usability but not how or how the factors interact. Table 9-1 compares the models of usability, including the model of learnability derived here with other contemporary models. The model of learnability is causal and attempts show the interactions between the various factors and subfactors. The advantages of a causal model are that researchers can anticipate the effects of making changes to one of the subfactors/factors to other factors.

Table 9-1 A comparison of the models of usability with the model of learnability derived from this study.

Models of Learnability/Usability	Key advantages	Limitations
Nielson (1993b)	All inclusive model.	Too wide a scope. Does not address cognition factors in depth. Does not address learnability.
Green's Cognition Dimensions (1989).	Detailed focus on cognition.	Does not address learnability directly and is not a causal model.
Dix et al (1998).	Seems to address all phenomena.	No apparent empirical support, not causal.
Individual factors e.g. Jordan *et al* (1991) guessability Kellog (1987) consistency.	Detailed focus on specific cognitive issues.	No placement in a wider model.
learnability as time related phenomenon, Shneiderman (1998), Preece et al (1994) Faulkner (1998).	Easy to measure.	Provides no vehicle for investigating the underlying phenomena.
The model of learnability derived in this study.	Causal model. Detailed focus on learnability.	Needs to be placed in a wider framework, e.g. Nielson's model of system acceptability (1993b).

Section 8.4.2 considered the model of learnability in relation to contemporary theories of learning. This exercise provided corroborating theory for the model particularly in relation to Gagne's (1977) model of the learning process and the events of learning, constructivist learning theory and the concept of scaffolding.

9.4 Recommendations and Further Research

It is recommended that:

- New HAPs should take account of the requirements specification given above.

- Usability evaluation of HAPs should adopt the approach used in this study and summarised above. The approach could be extended to all computer-based tools.

Further research is needed:

- to design and evaluate a diagrammatic language that allows knowledge construction and functionality construction to be undertaken as two separate phases;

- find corroborating evidence for the existence of the factors discovered in this study and develop HCI strategies for improving their value or mitigating against their adverse effects;

- develop a symbology and syntax within a diagrammatic language that provides for hypermedia designed for constructivist approaches to learning; and

- develop the model of interaction of factors found in this study into a complete model for the learning of HAPs.

- A more focused study of the role of motivation in the learning of HAPs is also needed.

References

10. References

Amabile, T., Hill, K., Hennessey, B. and Tighe, E. (1994) The Work Preference Inventory: Assessing Intrinsic and Extrinsic Motivational Orientations, *Journal of Personality and Social Psychology*, 66, 5, 950-967.

Allwood, C. (1990) Learning and Using Text-Editors and Other Application Programs, In *Cognitive Ergonomics: Understanding, Learning and Designing Human-Computer Interaction*. London: Academic Press.

Anderson B., Smyth, M., Knott, R., Bergan, M., Bergan, J., Alty, J., (1994) Minimising Conceptual Baggage: Making Choices about Metaphor, *Proceedings of HCI 94*, Glasgow, Scotland, 179-194.

Apple (1984) *Apple Human Interface Guidelines: The Apple Desktop Interface*, Reading, MA: Addison-Wesley.

Atkinson, J. and Drew, P. (1979) *Order in Court: The Organisation of Verbal Interaction and Judicial Settings*, London: Macmillan.

Bailey, G. (1993) Iterative Methodology and Designer Training in Human-Computer Interface Design, *Proceedings of CHI' 93*, 198-205.

Bandura, A. (1971) *Social Learning Theory*. New York: General Learning Press.

Barenholz, H. and Tamir, P. (1992) A comprehensive use of concept mapping in design instruction and Assessment. *Research in Science and Technological Education*. 10, 1, 37-52

Barker, P. (1992) An object-oriented approach to hypermedia authoring. In M. Giardina (Ed.), *NATO Advanced Science Institute Series F. Interactive Multimedia Learning Environments*, Berlin: Springer-Verlag, 132-152.

Barker, P. (1993) *Exploring Hypermedia*. London: Kogan Page.

Barker, P. and Banerji, A. (1994) A Global Performance Support System for Students and Staff, *Educational and Training Technology International*, 32, 1, 35-44.

Barnard, P. (1991) The Contribution of Applied Cognitive Psychology to the Study of Human-Computer Interaction. In Shackel B. and Richardson S. (Eds.), *Human Factors for Informatics Usability*. Cambridge: Cambridge University Press.

Barnard, P., Wilson, M. and Maclean, A. (1986) The Elicitation of System Knowledge by Picture Probes. *Proceedings of CHI' 86 Human Factors in Computing Systems*, Boston.

Beeman, W. O. Anderson, K. T. Bader, G. Larkin, J. McClard, A. P. McQuilian, P. and Shields, M. (1987) Hypertext and pluralism: from lineal to non-lineal thinking. *Proceedings of Hypertext' 87*. University of North Carolina, Chapel Hill.

Beltran, T. (1993) Educational hypermedia: from theory to practice. *Proceedings of ED-MEDIA' 93 - World Conference on Educational Multimedia and Hypermedia*, Orlando, Florida, 68-75.

Bench-Capon, T., (1990) Knowledge Representation: An Approach to Artificial Intelligence, London: Academic Press.

Berelson, B. (1954) Content Analysis, In Lindsey G. (Ed.) *Handbook of Social Psychology: Theory and Method*, Volume 1, Cambridge MA: Addison-Wesley.

Bigge, M. (1982) *Learning Theories for Teachers*. New York: Harper and Row.

Boose, J. (1985) A Knowledge Acquisition Program for Expert Systems Based Upon Personal Construct Psychology. *International Journal of Man-Machine Studies, 5.*

Briggs, P. (1988) What We Know and What We Need to Know: The User Model Versus the User's Model in Human-Computer Interaction, *Behaviour and Information Technology*, 7, 4, 431-442.

Bruillard, E. and Weidenfeld, G. (1990) Some examples of hypertext applications. In D. Jonassen and H. Mandl (Eds.), *Designing Hypermedia for Learning*, Berlin: Springer-Verlag.

Bruner, J. and Anglin, J. (1974) *Beyond the Information Given: Studies in the Psychology of Knowing*. New York: Norton.

Bruner, J. (1975) The Ontogenesis of Speech Acts, *Journal of Child Language*, 2, 1-19.

Bryman, A. (1988) *Quantity and Quality in Social Research*, London: Unwin Hyman.

Bryman, A. and Cramer, D. (1997) *Quantitative Data Analysis with SPSS for Windows*. London: Routledge.

Buckingham Shum, S. and Hammond, N. (1994) Transferring HCI Modelling and Design Techniques to Practitioners: A Framework and Empirical Work, *People and Computers IX, Proceedings of HCI 94*, 21-36, Glasgow.

Buzan, T. (1993) *The Mind Map Book*, London: BBC Books.

Byerbach, B., and Smith, J. (1990) Using a computerized concept mapping program to assess preservice teachers' thinking about effective teaching. *Journal of Research in Science Teaching*, 7, 10, 339-347.

Carlson, P., and Slaven, M. (1992) Hypertext Tools for Knowledge Workers the Next Frontier: Tools that can Teach, *Information Systems Management,* Spring.

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

Carroll, J. (1990) *The Nurnberg Funnel*, Cambridge, MA, MIT Press.

Carroll, J. (1998) Reconstructing Minimalism. In Carroll J. (Ed), *Minimalism Beyond the Nurnberg Funnel*, 1-18, Cambridge, MA, MIT Press.

Carswell, L. (1998) Possible Versus Desirable in Instructional Systems: Who's Driving? *ALT-J,* 6, 1, 64-69.

CASTLE (1998) *Computer Assisted Teaching and Learning: The CASTLE Project,* www.le.ac.uk/cc/ltg/castle -accessed July 1998

CCTA (1990) *SSADM Version 4 Reference Manual*, NCC-Blackwell.

Chaloupka, M. and Koppi, T. (1998) A Vignette Model for Distributed Teaching and Learning, *ALT-J,* 6, 1, 41-49.

Chapanis, A. (1991) Evaluating Usability, In Shackel B and Richardson S (Eds.), *Human Factors for Informatics Usability,* Cambridge: Cambridge University Press.

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

Chen, C. and Czerwinski, M. (1997) Spatial ability and visual navigation: an empirical study, *The New Review of Hypermedia and Multimedia*, 3, 67-90.

Christie, A. (1990) The use of interactive videodisc in the teaching of orthopedics in physiotherapy. *Medical Teacher* 12, 175-179.

Clarke, J. (1982) *Resource-based Learning for Higher and Continuing Education,* London: Croom Helm.

CLASS (1998) *The CLASS Project,* www.icbl.hw.ac.uk/projects/CLASS.htm - accessed July 1998.

Coffey, A. and Atkinson, P. (1996) *Making Sense of Qualitative Data,* Thousand Oaks CA: Sage Publications.

Dainty, G. (1997) Ecclesiomophic Hypermedia, *Presented at ED-MEDIA' 96, World Conference on Educational Multimedia and Hypermedia,* Boston MA.

Darby, J (1993a) Executive Overview of the Computers in Teaching Initiative, *CTISS File,* 16, 3.

Darby, J. (1993b) Teaching and Learning Technology Programme: Phase I, *The CTISS File,* 15, 27.

Darby, J (1993c) Teaching and Learning Technology Programme: Phase II, *The CTISS File,* 16, 6-7.

Davis, F., Bagozzi, P. and Warshaw, P. (1989) User Acceptance of Computer Technology: A Comparison of Two Theoretical Models, *Management Science,* 35, 8, 982-1003.

Davis, H. Hutchings, G. and Hall, W. (1993) A Framework for Delivering Large-scale Hypermedia Learning Material, *Proceedings of ED-MEDIA' 93 World Conference on Educational Multimedia and Hypermedia,* 115-122.

Dearing Report (1997) *Higher Education in the Learning Society,* HMSO Books.

Derks, M., Bulthuis, W., Monaghan, P., Keane, M., Easterbrook, M., Niehuis, E. and Bloemberg, W., (1989) *Learning Material Development, Tools, and Environments. DELTA SAFE Project (D1014) Report OS-ID/1.* Philips, Eindhoven.

Dey, I. (1993) *Qualitative Data Analysis.* London: Routledge.

Diaper D. (Ed.) (1989) *Knowledge Elicitation. Principles, Techniques and Applications,* Chichester: Ellis Horwood.

Dillon, A., Knight, C. and Richardson, J. (1993) Space - The Final Chapter or Why Physical Representations are not Semantic Intentions. In Knight C Dillon A and Richardson J. (Eds.), *Hypertext a psychological perspective,* Chichester: Ellis Horwood.

Dix, A., Finlay, J., Abowd, G. and Beale, R. (1998) *Human-Computer Interaction,* Europe: Prentice Hall.

Dobson, M., Rada, R., Chen, C., Michailidis A. and Ulloa, A. (1993) Towards a Consolidated Model for a Collaborative Courseware Authoring System, *Journal of Computer Assisted Learning,* 9, 34-50.

Duffy, T. M. and Knuth, R. A. (1990) Hypermedia and instruction: Where is the match? In D. Jonassen and H. Mandl (Eds.), *Designing Hypermedia for Learning,* Berlin: Springer-Verlag.

Eason, K. (1984) Towards the Experimental Study of Usability, *Behaviour and Information Technology,* 3, 2, 133-143.

Eason, K. and Damordaran, L. (1981) The Needs of the Commercial User. In Coombs M. and Alty J. (Eds.), *Computer Skills and the User Interface.* London: Academic Press.

Edwards, D. and Hardman, L. (1993) 'Lost in Hyperspace': Cognitive Mapping and Navigation in a Hypertext Environment. In McAleese R. (Ed) *Hypertext, Theory into Practice*, London: Intellect, 90-106.

Elliott, G., Jones, E. and Barker, P. (1995a) Making sense: A review of hypermedia in Higher Education. *Proceedings of ED-MEDIA' 95-World Conference on Educational Multimedia and Hypermedia*, Graz, Austria, 205-210.

Elliott, G., Jones, E. and Barker, P. (1995b) The Application of Concept Mapping to Hypermedia Authoring. *Proceedings of AETT 95*, Plymouth, United Kingdom.

Elliott, G., Jones, E. and Barker, P. (1997) Development of the Computer Skills Metric for Windows. *Proceedings of ED-MEDIA97*, Calgary, Canada, 310-315.

Ericsson, K. and Simon, H. (1984) *Protocol Analysis: Verbal Reports as Data.* Cambridge MA: MIT Press.

Faulkner, C. (1998) *The Essence of Human-Computer Interaction.* London: Prentice Hall.

Fisher, K. (1992) SemNet: A Tool for Personal Knowledge Construction, *Cognitive Tools for Learning*, edited by Kommers P. Jonassen D. and Mayes T., 63-76, Berlin: Springer-Verlag.

Flanagan, J. (1954) The Critical Incident Technique, *Psychological Bulletin*, 51, 327-358.

Foley, J. (1995) Visualising the WWW with the Navigational View Builder, *Computer Networks and ISDN Systems* 27, 6, 1075-1087.

Follet Report, (1996) *Joint Funding Council's Libraries Review Group: Report*, Bristol: HEFCE.

Fowler, R., Fowler, W. and Williams, J. (1997) 3D Visualisation of WWW Semantic Content for Browsing and Query Formulation, *WEBNET' 97*, Toronto.

Freeman, H. and Ryan, S. (1995) Supporting the Process of Courseware Development from Concept to Delivery. *Proceedings of ED-MEDIA' 95-World Conference on Educational Multimedia and Hypermedia*, Graz, Austria, 223-228.

Freeman, H. and Ryan, S. (1997) Webmapper a Tool for Planning, Structuring and Delivering Courseware on the Internet. *Proceedings of ED-MEDIA' 97-World Conference on Educational Multimedia and Hypermedia*, Calgary, Canada, 372-377.

Gagne, R. (1977) *The Conditions of Learning*, New York: Holt Rhinehart and Winston.

Gaines, B. And Shaw, M. (1995) Concept Maps as Hypermedia Components, *International Journal of Human-Computer studies* 43, 323-361.

Garzotto, F. and Matera, M., (1997) A Systematic Method for Hypermedia Usability Evaluation. *The New Review of Hypermedia and Multimedia, 3*, 39-66.

Gilmore, D. (1991) Visibility: A Dimensional Analysis. In Diaper D and Hammond N (Eds.), *People and Computers VI,* Cambridge: Cambridge University Press.

Green, T. (1989) Cognitive Dimensions of Notations. In Sutcliffe A. and Macaulay L. (Eds.), *People and Computers V,* 443-460, Cambridge: Cambridge University Press.

Green, T. and Petre, M. (1996) Usability Analysis of Visual Programming Environments: A Cognitive Dimensions Framework, *Journal of Visual Languages and Computing, 7*, 131-174.

Gross, M. (1994) Recognising and Interpreting Diagrams in Design. *Proceedings of the Workshop on Advanced Visual Interfaces,* 88-94, Bari, Italy.

Gunn, C. (1996) *A Framework for Situated Evaluation of Learning in Computer Environments,* PhD Thesis, Heriot Watt University.

Haber, R. (1970) How We Remember What We See, *Scientific American,* 222, 5, 104-115.

Halasz, F. (1988) Reflections on Notecards: Seven Issues for the Next Generation of Hypermedia Systems. *Communications of the ACM,* 31, 836-852.

Halasz, F. (1991) Transcript of Seven Issues: Revisited Closing Plenary, *Hypertext' 91,* **www.parc.xerox.com/projects/halasz-keynote/transcript.html.** - -accessed July 1998.

Hammond, N. Gardner, N. Heath, S. Kibby, M. Mayes, T. McAleese, R. Mullings, C. and Trapp, A. (1992) Blocks to the Effective Use of Information Technology in Higher Education. *Computers in Education,* 18, 155-162.

Hardman, L., Bulterman, D. and Rossum, G. (1994) The Dexter Hypertext Reference Model, *Communications of the ACM,* 37, 2, 30-62.

Heeren, E. (1992) Flexibility of Expressiveness: A critical factor in the design of concept mapping tools for learning. In Kommers P. Jonassen D and Mayes T (Eds.), *Cognitive Tools for Learning.* Berlin: Springer-Verlag.

HEFCE (1997) *The Information Technology Assisted Teaching and Learning in Higher Education Report,* HEFCE Research Series.

Hoffman, R. (1987) The Problem of Extracting the Knowledge of Experts from the Perspective of Experimental Psychology, *AI Magazine,* Summer issue, 53-67.

Horney, M. (1993) Case studies of navigational patterns in constructive Hypertext. *Computers in Education,* 20, 257-270.

Hutchings, G., Hall, W., Briggs, J., Hammond, N., Kibby, M., McKnight, C. and Riley, D. (1992) Authoring and evaluation of hypermedia for education. *Computers in Education*, *18*, 171-177.

IJHCS (1997) *Special Issue on World Wide Web Usability*. International Journal of Human Computer Studies, 47, 1

Ishikawa, K. (1991) *Guide to Quality Control*, Asian Productivity Organisation.

Jackson, S., Stratford, S., Krajcik, J. and Soloway, E. (1996) A Learner-Centred Tool for Students Building Models, *Communications of the ACM*, 39, 4, 48-49.

Jih, J, and Reeves, T. (1992) Mental Models: A Research Focus for Interactive Learning Systems. *Educational Technology Research and Development*, 40, 3, 39-53.

JILR (1997) Special Issue on Concept Maps, *Journal of Interactive Learning Research*, 8, 3 and 4.

Jonassen, D. (Ed.) (1989) *Hypertext/Hypermedia*. New Jersey: Englewood Cliffs.

Jonassen, D. (1990) Semantic Network Elicitation: Tools for Structuring Hypertext. In McAleese R. and Green, C. (Eds.), *Hypertext: State of the Art,* Oxford: Intellect.

Jonassen, D. and Grabinger, S. (1990) Problems and Issues in Designing Hypertext/hypermedia for Learning. In Jonassen D. and Mandl H.(Eds.), *Designing Hypermedia for Learning,* Berlin: Springer-Verlag.

Jonassen, D. and Harris, D. (1991) Analysing and Selecting Instructional Strategies and Tactics. *Performance Improvement Quarterly*, 4.

Jonassen, D. (1992a) Hypertext as Cognitive Tools. In Kommers P. Jonassen D. and Mayes T. (Eds.), *Cognitive Tools for Learning,* Berlin: Springer-Verlag.

Jonassen, D. H, Ambruso, D. R, and Olesen, J. (1992b) Designing a Hypertext on Transfusion Medicine Using Cognitive Flexibility Theory. *Educational Multimedia and Hypermedia Journal*, 1, 309-322.

Jonassen, D. (1993) Effects of Semantically Structured Hypertext Knowledge Bases. In McKnight C., Dillon A. and Richardson J (Eds.), *Hypertext: a Psychological Perspective*. London: Ellis Horwood.

Jonassen, D., Mayes T. and McAleese, R. (1993) A Manifesto for a Constructivist Approach to Uses of Technology in Higher Education. In Duffy T. Lowyck J. and Jonassen D. (Eds.), *Designing Environments for Constructive Learning,* 231-247, Berlin: Springer-Verlag.

Jonassen, D. and Mara, R. (1994) Concept Mapping and Other Formalisms as Mindtools for Representing Knowledge. *ALT-J*, 2, 1, 50-56.

Jonassen, D (1996) *Computers in the Classroom*, New Jersey: Prentice Hall.

Jordan, P., Draper, S., MacFarlane, K. and McNulty, S. (1991) Guessability, Learnability and Experienced User Performance. In Diaper D. and Hammond N. (Eds.), *People and Computers VI*, 237-248, Cambridge: Cambridge University Press.

Junkala, J. (1991) Creating CAI Courseware for College-Level Instruction: Almost Anybody Can Do it. *Educational Technology*, 15-20.

Kaplan, N. and Moultrop, S., (1994) Where No Mind Has Gone Before: Ontological Design for Virtual Spaces. *Proceedings of the Fourth European Conference on Hypertext*, 206-216.

Kato, T. (1986) What 'Question-asking Protocols' Can Say About the User Interface. *International Journal of Man-Machine Studies*, 25, 659-673.

Kellog, W. (1987) Conceptual Consistency in the User Interface: Effects on User Performance. In Bullinger H. and Shackel B. (Eds.), *Proceedings of the Second IFIP Conference on Human-Computer Interaction - INTERACT' 87*, 389-394, Stuttgart, Germany.

Kelly, A. E. (1993) Designing Instructional Hypertext for Use in Lecture Note Review: Knowledge Engineering and Preliminary Testing. *Educational Multimedia and Hypermedia*, 2, 149-176.

Kelly, G. (1955) *The Psychology of Personal Constructs*. New York: Norton

Kibby, M. R. and Mayes, T. (1993) Towards Intelligent Hypertext. In R. McAleese (Ed.), *Hypertext: Theory into Practice*, 139-144, Oxford: Intellect.

Kitchens, L. (1996) *Exploring Statistics, A Modern Introduction to Data Analysis and Inference*, Pacific Grove, CA: Duxbury Press.

Koegel, J. and Heines, J. (1993) Improving Visual Programming Languages for Multimedia Authoring. *Proceedings of ED-MEDIA' 93 World Conference on Educational Multimedia and Hypermedia*, Orlando, Florida, 286-293.

Kolb, D. (1984) *Experiential Learning - Experience as the Source of Learning and Development*, New Jersey: Englewood Cliffs.

Kommers, P. (1992) TextVision and the Visualisation of Knowledge: School-based Evaluation and of its Acceptance at two Levels of Schooling, *Cognitive Tools for Learning*, edited by Kommers P. Jonassen D. and Mayes T., 33-62, Berlin: Springer-Verlag.

Kozma, R. (1992) Constructing Knowledge with Learning Tool, *Cognitive Tools for Learning*, edited by Kommers P. Jonassen D. and Mayes T., 23-32, Berlin: Springer-Verlag.

Kremmer, R. (1997) Multi-user Interactive maps for the Learning Web, *Proceedings of EdMedia97, World Conference on Educational Multimedia and Hypermedia,* Calgary, Canada

Kulpa, Z. (1994) Diagrammatic Representation and Reasoning, *Machine GRAPHICS AND VISION,* 3, 1 and 2.

Landow, G. (1990) Popular Fallacies about Hypertext. In Jonassen D. and Mandl H.(Eds.), *Designing Hypermedia for Learning,* Berlin: Springer-Verlag.

Laurillard, D. (1993) *Rethinking University Teaching* (2nd ed.) London: Routledge.

Laurillard, D. Swift, B. and Darby, J. (1993) Academics Use of Courseware Materials: a Survey. *ALT-J,* 1, 1, 4-14.

Lawton, D. and Smith, L., (1993) The Knowledge Weasel Hypermedia Annotation System, *Proceedings of Hypertext 93,* 106-117.

Lennon, J. and Maurer, H. (1994) Lecturing Technology: A Future with Hypermedia. *Educational Technology,* April, 5-14.

Levesque H. and Mylopoulos, J. (1979) A Procedural Semantics for Semantic Networks. In *Associative Networks,* London: Academic Press.

Lindsay, P. and Norman, D. (1977) *Human Information Processing: An Introduction to Psychology.* New York: Academic Press.

Linn, M. C. (1992) How Can Hypermedia Tools Help Teach Programming? *Learning and Instruction,* 2, 119-139.

Maas, S. (1983) Why Systems Transparency? In Green T., Payne, S. and Van de Veer G. (Eds.), *The Psychology of Computer Use,* London: Academic Press.

MacFarlane, A. G. J. (1990) Interactive Computing: A Revolutionary Medium for Teaching and Design. *Computing and Control Engineering,* July, 149-158.

Marchionini, G. (1988) Hypermedia and Learning: Freedom and Chaos. *Educational Technology,* 8-20.

Mark, G., Haake, J. and Sreitz, N. (1995) The Use of Hypermedia in Group Problem Solving: An Evaluation of the Dolphin Electronic Meeting Room Environment, *Proceedings of the European Conference on Computer-Supported Cooperative Work,* Stockholm.

Maurer, H. (1996) *Hyperwave: The Next Generation Web Solution.* Harlow: Addison-Wesley.

Mayer, S., Müldner, T. and Unger, C. (1993) NEAT: An Integrated Authoring Environment Based upon Toolbook. *Proceedings of ED-MEDIA' 93 - World Conference on Educational Multimedia and Hypermedia*, Orlando, Florida, 332-339.

McAleese, R., (1985) Some problems of knowledge representation in an authoring environment: exteriorisation, anomalous state meta-cognition and self-confrontation, *Programmed Learning and Education Technology*, 22, 4, 299-306.

McAleese, R., (1987) The Graphical Representation of Knowledge as an Interface to Knowledge Based Systems. In Bullinger H and Shackel B (Eds.), *Proceedings of the Second IFIP Conference on Human-Computer Interaction - INTERACT' 87*, 1089-1093, Stuttgart, Germany.

McAleese, R. (1990) Concepts as Hypertext Nodes: The Ability to Learn While Navigating Through Hypertext Nets. In Jonassen and H. Mandl (Eds.), *Designing Hypermedia for Learning*, 97-116, Berlin: Springer Verlag

McAleese, R. and Ching, L. (1993) An Instructional Design Advisor for Computer-based Learning: ISAAC. *Proceedings of ED-MEDIA' 93-World Conference on Educational Multimedia and Hypermedia*, Orlando, Florida, 340-358.

McAleese, R. (1994) A Theoretical View on Concept Mapping, *ALT-J*, 3, 1, 38-48.

McGrath, D. (1992) Hypertext, CAI, Paper or Program Control: Do Learners Benefit from Choice? *Research on Computing in Education*, 24., 513-532.

McKeever, S. McKeever, D. and Elder, J. (1997) An Authoring Tool for Constructing Interactive Exercises Based on Automatic Generation of Client Side Functionality. *Proceedings of WEBNET' 97*, Toronto.

Merrill, D. (1991) Constructivism and Instructional Design, *Educational Technology*, May, 45-53.

Midoro, V. Persico, D. Trentin, G. and Viarengo, V. (1992) Using a Database of Multimedia Learning in Courseware Authoring. *Educational Multimedia and Hypermedia* 1, 77-90.

Miller, P. (1995) CourseWorks: an extended concept mapping tool. *Active Learning* 3, 50-53.

Minsky, M. (1975) A Framework for Representing Knowledge. In Henry W. (Ed.) *The Psychology of Computer Vision*. McGraw-Hill.

Misanchuk, E. and Schweir, R. (1992) Representing Interactive Multimedia and Hypermedia Audit Trails. *Educational Multimedia and Hypermedia*, 1, 355-372.

Molnar, K. and Kletke, M. (1996) The Impacts on the User Performance and Satisfaction of a Voice-based Front-end Interface for a Standard Software Tool, *International Journal of Human-Computer Studies*. 45, 287-303.

Morse, A. and Reynolds, G. (1993) Overcoming Current Growth Limits in UI Development, *Communications of the ACM,* 36, 4, 73-81.

Motta, E., Eisenstadt, M., West M., Pitman, K. and Everstz, R. (1986) KEATS: the Knowledge Engineer's Assistant, Final Report, *Technical Report No. 20, Human Cognition Laboratory,* Milton Keynes: OU Press.

Mukherjea, S. and Foley, J. (1995) Visualising the WWW with the Navigational View Builder, *Computer Networks and ISDN Systems,* 27, 1075-1087.

Mülhauser, M. (1992) Hypermedia and Navigation as a Basis for Authoring/learning Environments. *Educational Multimedia and Hypermedia* 1, 51-64.

Myers, L. and Grossen, N (1978) *Behavioral Research, Theory, Procedure and Design,* San Francisco: W. H. Freeman and Company.

NATO ASI (1992) *Cognitive Tools for Learning, NATO ASI Series vol F81,* edited by Kommers P. Jonassen D. and Mayes T. Berlin: Springer-Verlag.

Napier, H., Lane, D. and Batsell, R. (1989) Impact of a Restricted Natural Language Interface on Ease of Learning and Production. *Communications of the ACM* 32, 10, 1190-1197.

Nelson, T. (1967) *Getting it Out of Our System, Information Retrieval a Critical Review.* Washington: Thompson Books.

Nelson, T. (1980) Replacing the Printed Word: A Complete Literary System. *Proceedings of IFIP,* 1013-1023.

Nelson, W. and Palumbo, D. (1992) Learning, Instruction and Hypermedia. *Educational Multimedia and Hypermedia,* 1, 287-299.

Newell, A. (1982) The Knowledge Level, *Artificial Intelligence,* 18, 87-127.

Nielson, J. (1990a) *Hypertext and Hypermedia,* London: Academic Press.

Nielson, J (1990b) Navigating through Hypertext, *Communications of the ACM,* 33, 3, 297-310.

Nielson, J. (1990c) Designing Hypertext Usability. In Jonassen D. and Mandl H. (Eds.), *Designing Hypermedia for Learning,* Berlin: Springer-Verlag.

Nielson, J. (1993a) Non command User Interfaces, *Communications of the ACM,* 36, 4.

Nielson, J (1993b) *Usability Engineering,* London: Academic Press.

Novak, J. and Gowin, D. (1984) *Learning How to Learn,* New York: Cambridge University Press

O'Donnell, P., Scobie, G. and Baxter, I. (1991) The Use of Focus Groups as an Evaluation Technique in HCI. In Diaper D. and Hammond N. (Eds.), *People and Computers VI,* Cambridge: Cambridge University Press.

Okebukola, P. (1992) Attitude of Teachers Towards Concept Mapping and Vee Diagramming as Metalearning Tools in Science and Mathematics. *Educational Research*, 34 (3).

Oliveira, A. (1992) Hypermedia and Multimedia. In. Oliveira A. (Ed.), *Hypermedia Courseware: Structures of Communication and Intelligent Help,* Berlin: Springer-Verlag.

Park, Ok-C. (1991) Hypermedia: Functional Features and Research Issues. *Educational Technology.*

Payne, S. (1985) Task-Action Grammar. *Human-Computer Interaction - INTERACT' 84.*

Piaget, J. (1977) *The Development of Thought: Equilibrium of Cognitive Structures.* New York: Viking Press.

Preece, J. (1994) *Human-Computer Interaction.* Harlow, England: Addison-Wesley

Psychological Corporation, (1996) Differential Aptitude Tests. The Psychological Corporation 5th edition. London: Harcore Brace.

Quillian, R. (1968) Semantic Memory. In Minsky M. (Ed.), *Semantic Information Processing.* Cambridge MA: MIT Press.

Rada, R. (1991) *Hypertext: from Text to Expertext,* Maidenhead: McGraw-Hill,.

Reader, W. (1995) *Structuring Knowledge: the Development and Evaluation of Tools to Support Learning.* PhD Thesis, York University.

Reader, W. and Hammond, N. (1994) Computer-based Tools to Support Learning from Hypertext: Concept mapping tools and beyond. *Computers in Education*, 22, 99-106

Reynolds, S. and Dansereau, D. (1990) The Knowledge Hypermap: An Alternative to Hypertext. *Computers in Education* 14, 409-416.

Richards, T. and Richards, L. (1991) The NUD.IST Qualitative Data Analysis System, *Qualitative Sociology*, 14.

Richards, T. and Richards, L. (1994) Using Hierarchical Categories in Qualitative Data Analysis. In Kelle U. (Ed.), *Computers and Qualitative Methodology.* Sage.

Riley, F. (1994) Understanding IT: *A Review of Hypermedia Authoring Packages.* Loughborough University.

Roberts, T. and Moran, T. (1983) The Evaluation of Text Editors: Methodology and Empirical Results, *Communications of the ACM,* 26, 4, 265-283.

Robertson. G., Card, S. and Mackinlay, J. (1993) Information Visualisation Using 3D Interactive Animation, *Communications of the ACM,* 36, 4, 57-71.

Rode, M. and Poirot, J. (1991) Authoring Systems-Are They Used? *Research on Computing in Education,* Winter, 191-198.

Ross, B. and Munby, H. (1991) Concept Mapping and Misconceptions: a Study of High-School Students' Understanding of Acids and Bases, *International Journal of Science Education.* 13 ,1, 11-23.

Ross, T. W. (1993) Bloom and Hypertext Parallel Taxonomies? *Educational Technology Review,* 11-16.

Russell, D., (1990) Alexandria: A Learning Resources Management Architecture. In *Designing Hypermedia for Learning.* In Jonassen D. and Mandl H. (Eds.), Berlin: Springer Verlag.

Savery, J. and Duffy, T. (1994) Problem Based Learning: An Instructional Model and its Constructivist Framework. *Educational Technology,* August.

Schvaneveldt, R. (Ed.) (1990) *Pathfinder Associative Networks: Studies in Knowledge Organisation,* New Jersey: Ablex Publishing Corporation.

Searle, J. (1981). Minds, Brains and Programs. In Hofstadter, D. and Dennett, D. (Eds.), *The Mind's I,* 353-372, Harmondsworth: Penguin.

Seely Brown, J. (1986) From Cognitive to Social Ergonomics and Beyond. In Norman D. and Draper S. (Eds.), *User Centred Systems Design,* London: Lawrence Erlbaum.

Shackel, B. (1986) Ergonomics in Design for Usability. In Harrison, M. Monk, A. (Eds.), *Proceedings of the British Computer Society Human-Computer Interaction Specialist Group,* 44-64.

Shackel, B. (1990) Human Factors and Usability. In Preece J. and Keller L. (Eds.), *Human Computer Interaction,* Hemel Hempstead: Prentice Hall.

Shneiderman, B. (1998) *Designing the User Interface,* Reading MA: Addison Wesley.

Siviter, D. and Brown, K. (1992) Hypercourseware. *Computers in Education* 18, 163-170.

Siviter, P. (1997) Authoring Tools for Courseware on WWW: the W3 Lessonware Project, *WEBNET' 97*, Toronto.

Skinner, B. (1938) *The Behaviour of Organisms*. New York: Appleton.

Slavin, R. (1991) *Educational Psychology: Theory into Practice*. NJ: Prentice Hall

Smith, A. (1994) Navigating Hyperspace: Assessing Usability, PhD thesis, Nottingham University.

Spiro, R., Feltovich, P., Jacobson, M. and Coulson, R. (1991) Cognitive Flexibility, Constructivism, and Hypertext: Random Access Instruction for Advanced Knowledge Acquisition in Ill Structured Domains. *Educational Technology*, May, 24-33.

SPSS Inc (1993) *SPSS Base System Syntax Reference Guide, Release 6.0,*: SPSS Inc Chicago.

SPSS Inc (1994) *SPSS Advanced Statistics, Release 6.1,*: SPSS Inc Chicago.

Steinberg, E. (1988) Cognition and Learner Control: A literature review, 1977-1988. *Computer-Based Instruction*, 16, 4.

Strauss, A. and Corbin, J. (1990) *Basics of Qualitative Data*, London: Sage.

Stubenrauch, R. Kappe, F. and Andrews, K. (1993) Large Hypermedia Systems: The End of the Authoring Era. *Proceedings of ED-MEDIA' 93-World Conference on Educational Multimedia and Hypermedia*, Orlando, Florida, 495-502.

Sullivan, K. (1996) The Windows 95 User Interface: A Case Study in Usability, *Proceedings of CHI' 96*, 473-480, Vancouver.

Sutcliffe, A. (1995) *Human-Computer Interface Design*, Basingstoke: MacMillan.

Thimbleby, H. (1992) Heuristics for Cognitive Tools. In. Kommers P. Jonassen D. and Mayes T. (Eds.), *Cognitive Tools for Learning*, 161-168, Berlin: Springer-Verlag.

Tomek, I. and Maurer, H. (1992) Helping the User Select a Link. *Hypermedia*, 4, 111-122.

Tullis, T. (1986) Optimising the Usability of Computer-Generated Displays. *Proceedings of the British Computer Society Human-Computer Interaction Specialist Group*, 604-614, York.

Von Glasenfeld (1984) Radical Constructivism. In Watzlawick P. (Ed.), *The Invented Reality*, Cambridge MA: Havard University Press.

Vygotskii, L (1962) *Thought and Language*, Cambridge MA: MIT Press.

Waern, Y. (1990) Human Learning of Human-computer Interaction: An Introduction, *In Cognitive Ergonomics: Understanding, Learning and Designing Human-Computer Interaction.* 60-83, London: Academic Press.

Weatherburn, P. and Project SIGMA (1992) Alcohol Use and Unsafe Sexual Behaviour: Any Connection? In Aggleton P. (Ed.) *AIDS: Rights, Risk and Reason,* London: Falmer Press.

Weiss, N. (1995) *Introductory Statistics*, Reading, MA: Addison-Wesley.

Whalley, P. (1990) Models of Hypertext Structure and Learning. In Jonassen D. and Mandl H.(Eds.), *Designing Hypermedia for Learning,* Berlin: Springer-Verlag.

Whiteside, J, Jones, S., Levy, P. and Wixon D. (1985) User Performance With Command, Menu, and Iconic Interfaces, *Proceedings of CHI' 85,* 185-191, San Francisco.

Wilson, M., Barnard, P. and MacLean, A. (1990), An Investigation of the Learning of a Computer System. In *Cognitive Ergonomics: Understanding, Learning and Designing Human-Computer Interaction.* 151-172, London: Academic Press.

Winograd, T. (1975) Towards a Procedural Understanding of Semantics, SAIL Memo AIM-292, Stanford University, Stanford.

Yin, R. (1994) *Case Study Research: Design and Methods*, California: Sage.

Zhao, Z. O'Shea, T. and Fung, P. (1993) Visualization of Semantic Relations in Hypertext Systems: A Set of Empirical Studies. *Proceedings of ED-MEDIA' 93 - World Conference on Educational Multimedia and Hypermedia,* 556-564, Orlando, Florida.