Addressing the "lost in hyperspace" problem in hypertext

September 1997

Abstract

Addressing the "lost in hyperspace" problem in hypertext

Theng Yin Leng

End-users tend to lose their way in the maze of information within hypertext. Much work done to address the "lost in hyperspace" problem is *reactive*, that is, doing remedial work to correct the deficiencies within hypertexts because they are (or were) poorly designed and built. What if solutions are sought to avoid the problem? What if we do things well from the start?

This thesis reviews the "lost in hyperspace" problem, and suggests a framework to understand the design and usability issues. These issues cannot be seen as purely psychological or purely computing, they are multi-disciplinary. The *proactive, multi-disciplinary* approach undertaken in this thesis is drawn from current technologies in sub-disciplines of hypertext, human-computer interaction, cognitive psychology and software engineering, and they include investigations into:

- good design principles and guidelines for the building of hypertexts in the first place;
- task analysis and user modelling techniques for end-users' tasks to be carefully defined;
- effective hypertext structures to ensure usability of hypertexts; and
- designer tools to create hypertexts with a reduced tendency for the "lost in hyperspace" problem to arise.

To demonstrate these ideas, this thesis presents HyperAT, a hypertext research authoring tool, developed to help designers build usable web documents on the World Wide Web. The World Wide Web is chosen as a hypertext example since it is the most widely used and largest hypertext system ever. The approach taken in HyperAT is novel: multi-disciplinary approaches are integrated and implemented with a practical authoring tool.

Evaluation of HyperAT indicates that it does address the "lost in hyperspace" problem, particularly for the World Wide Web by: (i) helping designers manage the complexity of the design and validation processes; and (ii) helping hypertext end-users navigate hypertexts produced by HyperAT without getting "lost".

Preface

This thesis is based on research undertaken between December 1994 and September 1997. The research has been funded by and carried out in the School of Computing Science at the Faculty of Technology, under the supervision of Professor Harold Thimbleby and Dr Matthew Jones.

Because the "lost in hyperspace" problem is a complex one, this thesis carried out research into many "strands" of investigations involving various disciplines. Consequently, this productive problem-solving approach generated many new ideas that would otherwise be missed out if this thesis had focused on just one aspect of the problem. The intention of this thesis is to carry out a "breadth-first investigation" so as to obtain a good understanding of the design and usability issues surrounding the "lost in hyperspace" problem in hypertext. Therefore, much of the experimental work designed was kept small and focused.

The work presented in this thesis is composed and carried out by the author. Part of this work has been published in a form similar to that which appears here (for example, Theng, 1995; Theng, Thimbleby & Jones, 1995a; Theng, Thimbleby & Jones, 1995b; Theng, Thimbleby & Jones, 1996; Theng, Rigny, Thimbleby & Jones, 1996a; Theng, Rigny, Thimbleby & Jones, 1996b; Thimbleby, Jones & Theng, 1997; Theng, Rigny, Thimbleby & Jones, 1997). These papers had been presented at international (HCI'95 – '97, APCHI'96 and Hypertext'97) and national (IEE'95) conferences, and have received encouraging and favourable comments. A total of nine papers have been published during the period this research was undertaken, and they are reproduced in appendix J at the back of the thesis.

Part of the work submitted to APCHI'96, HCI'96 and HCI'97 has been carried out in collaboration with other authors. The use of executable cognitive user models in understanding end-users' behaviour and navigation issues in the design of hypertext systems has been jointly developed with Cécile Rigny at Middlesex University. The author's contributions are: (i) providing inputs for the drawing up of task diagrams on common tasks such as browsing, information search, seeking references, and recall; (ii) the design of a prototype hypertext implemented using HyperCard (version 2.01); (iii) a prototype hypertext authoring tool called HyperAT written in Macintosh Common Lisp (version 3.9) to help hypertext designers manage the complexity of the design and validation process without getting "lost". Cécile Rigny's contribution is in the design of CUM-DesTool, also a Macintosh Common Lisp program (version 3.9) that can run cognitive user models.

Acknowledgements

This work has been the support and encouragement of many people, who have contributed directly or indirectly to the completion of this thesis.

I wish to thank my Director of Studies, Professor Harold Thimbleby for his support, time, interest and valuable feedback he has given in my work. He has prompted many new ideas and created many new avenues of interest for me. Thanks must go to my Supervisor, Dr Matthew Jones for his constructive feedback and support provided in helping to shape this thesis. Special thanks and wishes go to Cécile Rigny who not only has been a wonderful friend, but has inspired me greatly in Lisp programming. I enjoyed the collaborative work we did together, and have learned much from her in the area of executable user modelling. Many thanks must go to Dr Mark Addison for his generous help, in the initial period of this work.

Thanks must also be given to many for their frank and generous comments in helping to make the thesis more coherent and cohesive. The list includes Dr Diana Bental, Dr Ann Blandford, Dr Kay Dudman and Dr Gordon Rugg. In particular, my heartfelt thanks go to Professor Alan Dix, Dr Ian Benest, Linda Hardman, Dr Gi Kwong Chung, Dr Rod Nicholson, Emeritus Professor Brian Shackel, Professor Ian Witten and Dr Patricia Wright for their valuable and insightful comments on early developments of the work. I am also grateful to Anthony Basiel, Uncle John David, Dr Gareth Palmer and Serengul Smith for giving me useful comments on early drafts of the thesis. The positive feedback provided by anonymous referees who reviewed the papers submitted to conferences has been of great encouragement to me in affirming the work being carried out. Special thanks also go to friends and students who have unselfishly given up time to take part in the experiments, which include Dr Mark Addison, A. Albers, Yin Shun Chao, Karin Guthe, Joanne Hyde, Gil Marsden, Dr Gareth Palmer, Doreen Ng, Dr Matt Smith, Serengul Smith, William Thimbleby, Lichun Wang, C.W. Wong, and students from COM0001 and the Masters Educational Multimedia courses.

Joanne Hyde, Serengul Smith, together with Cécile Rigny have been great room-mates. We share many fun moments discussing things outside our PhD work, but they have been greatly encouraging and supportive in times when the work is tough-going.

And certainly the greatest thanks go to my family, Natalie and James and close friends in Singapore whose love and support has sustained me so far in this research. I am also grateful to my church friends in London and Manchester who are always available to give me moral and spiritual support.

Initially, I thought it might be inappropriate to acknowledge God in a scientific piece of work such as a PhD, and I should share my gratitude to God with my christian friends only. However, to do so would deny the very source of my motivation, perseverance and strength in this often lonely work. Without God sustaining me through those painful and stressful moments, finishing this PhD on time would have been difficult.

Table of Contents

Section IV

Section V

Section I

Introduction, literature survey and area of study

Chapter 1

Introduction

Hypertext technology still continues to excite many, ever since it became popular in the late 1980s. Many people then were sceptical that the technology might just be a passing fad. However, there are some who think that the promise behind hypertext is too fundamental to disappear quickly (for example, Berk and Devlin, 1991; *etc.*), and there are reasons to believe that hypertext technology promises something special. Early visionary thinkers such as Bush (1945), Engelbart (1963) or Nelson (1987) in their own ways believed that hypertext and its underlying philosophy would enable man to locate, retrieve and use easily the store of human knowledge that lies in books, journals and associated materials in libraries all over the world. In hypertext, end-users not only benefit from the information they read but also from the richness of associations supported by the network of nodes and links. Hypertext has affected us directly, or indirectly in almost every facet of our lives, ranging from scientific work to business and education needs, to our general way of life. Take for example, the most widely used hypertext system in recent years is the World Wide Web (or simply the web) on the Internet. Helped by an avalanche of servers, documents and hyperlinks, today the web is used by a burgeoning user population ranging from computer scientists to businessmen to the general public. Thousands of websites are created to publish all kinds of activities ranging from cutting-edge research discoveries and financial analyses to cost-saving shopping tips.

1.1 Motivation and background

Though hypertext promises a whole new paradigm and freedom of reading and information retrieval, they are not without problems. Ironically, it is precisely this freedom and power with which hypertext equips end-users that give rise to many of these problems. Yet it is now more than *fifty* years - a long time! - since Vannevar Bush (1945) envisioned his hypertext "memex", a machine that would allow an individual to browse and create a multimedia library of technical documents including personal notes, photographs and sketches using microfilm and photocell technologies. Although many hypertexts have been developed since then, none has yet claimed to have completely and successfully addressed these problems (Pittas, 1995). There is certainly room for further research into addressing these problems.

One of the problems that seems inherent within hypertext is that end-users tend to lose their way in the maze of information within hypertext. This is commonly referred to as the "lost in hyperspace" problem abbreviated as LIH in this thesis.

§1.2 explains the aims of the investigation and lists the objectives and deliverables of this thesis. §1.3 compares and contrasts the approach taken in this thesis with other research efforts. §1.4 gives a brief history of the work carried out. §1.5 outlines the structure for the remainder chapters of the thesis.

1.2 Objectives and deliverables

Much work has been done to address the LIH problem which involves the use of graphical browsers and query/search mechanisms (for example, Conklin, 1987; Parunak, 1989; Edwards and Hardman, 1993; Nielsen, 1995b; Dvorak, 1995; Pittas, 1995; *etc.*). These examples represent a *reactive* approach, that is, doing remedial work to correct the deficiencies within hypertexts because they are (or were) poorly designed and built.

What if such solutions are being sought for an avoidable problem? What if things are done right from the start? The LIH problem can perhaps be *dissolved* by ensuring that good hypertext design principles and guidelines are incorporated into the building of hypertexts in the first place. (If so, then what are these principles and guidelines?) Perhaps designer tools to create hypertext documents with a reduced tendency for the LIH problem to arise should be investigated. Perhaps the task and structure of hypertexts should be carefully defined taking into consideration end-users' needs.

To address the LIH problem in hypertext, this thesis defines the following objectives:

Objective 1: Identify reasons why end-users are "lost" while navigating hypertext.

Objective 2: Address the LIH problem in hypertext by doing things well from the start:

- identify good design principles and guidelines in hypertext
- define end-users' tasks using task analysis and user modelling techniques
- identify effective hypertext structures and useful heuristics to ensure usability of hypertexts

Objective 3: Examine authoring tools for standalone hypertexts with a reduced tendency for the LIH problem to arise.

Objective 4: Iteratively design and build a hypertext prototype incorporating the findings obtained from objective 2.

Objective 5: A further line of investigation this thesis will do is to explore the LIH problem on a wider scope, that is, on the web. By integrating the approaches described above into a single system, end-users may be helped to navigate around the web without getting "lost". A prototype hypertext authoring tool is iteratively designed and developed to aid authors build usable hypertexts on the web.

1.3 Contribution to hypertext research

What this thesis hopes to achieve is to eliminate or at least, reduce the impact of the LIH problem in hypertext. It examines the LIH problem by taking a different stance compared to other research efforts. It identifies *proactive* approaches to address the LIH problem with the emphasis of doing things well from the start. These approaches draw upon and integrate knowledge and findings in seemingly diverse disciplines such as hypertext, human-computer interaction, cognitive psychology and software engineering. By integrating and developing the approaches proposed to address the LIH problem, a methodology and an authoring tool for the design and building of usable hypertexts are described. This has significant implications. No longer need research efforts be focused on a particular discipline to solve a problem. It opens up possible collaborative efforts involving many disciplines that may be necessary to solve complex problems, of which the LIH is an example. The benefits can be tremendous: savings in terms of money and time; contributions in the design of better hypertexts which can be tested, further built upon by other researchers.

1.4 A brief history of the work

Initial work on the content of this thesis began in December 1994 examining contemporary issues in interface design of interactive systems. By February 1995, an interest in hypertext and the associated LIH problem emerged. The idea of employing multi-disciplinary approaches to address the LIH problem in hypertext took shape in September 1995.

Work carried out to investigate the various approaches started from October 1995 to July 1996. Literature surveys were carried out, and a couple of experiments were conducted to investigate design issues. A prototype hypertext was developed on a Macintosh PowerPC using HyperCard (version 2.01). A transfer report was then written in July 1996 detailing work done, and outlining the remaining work that needed to be completed to make the thesis more coherent and cohesive.

Development of HyperAT, a prototype research authoring tool to help designers build usable hypertexts, began in August 1996. HyperAT was also developed on a Macintosh PowerPC using Common Lisp (version 3.9). By December 1996, the majority of the development of HyperAT was complete, although enhancements were made in response to feedback from colleagues and reviewers of papers submitted to conferences. Development on HyperAT as presented in this thesis was completed in August 1997.

Writing up of the thesis began in January 1997 and completed in its present form in September 1997.

1.5 Structure of thesis

This thesis is divided into major sections: (I) introduction, literature survey and area of study; (II) multi-disciplinary approaches to address the LIH problem; (III) applying hypermedia research on the LIH problem onto the web; (IV) concluding remarks, and (V) supporting materials. Section I introduces the work undertaken in this thesis and provides an overview of hypertexts and the LIH problem. Section II and III present the main area of study proposing concrete ideas and implementing these ideas. Section IV makes concluding remarks about the thesis. Section V contains supplementary materials to provide further explanation of the work carried out in this thesis.

Section I
Introduction, literature survey and area of study

- Chapter 1 provides the background, motivations and objectives of this thesis.

- Chapter 2 starts off by describing what hypertext is and the reasons for using hypertext technology. The problems associated to hypertext are identified and the motivation for solving the LIH problem is discussed. It re-visits the LIH problem by perceiving it within the end-users' mind, in conventional hypertexts and on the web. It also sets out the reasons for the *proactive, multi-disciplinary* approaches and methodology taken in this thesis. This thesis believes that if the concepts of disciplines such as hypertext, human-computer interaction, cognitive psychology and software engineering are synthesised, more usable hypertexts may be built. The experimental strategies taken to test the hypotheses put forward in this thesis are explained. This chapter concludes by clarifying the terminology used.

Section II
Multi-disciplinary approaches to address the "lost in hyperspace" problem

- Chapter 3 postulates a framework, a systematic approach to define principles and guidelines for hypertexts. It begins with a discussion on general design principles for interactive systems. To adapt general design principles and guidelines to hypertext systems, this thesis examined them from a psychological perspective, focusing specifically on principles relating to interface design of hypertexts. A list of principles essential for hypertext design was drawn up, incorporating sound cognitive principles. The emphasis is that designers should do things well from the start, so that remedial work need not be done to correct the deficiencies within hypertexts because they are (or were) poorly designed and built. Based on these design principles proposed, a prototype hypertext was built. Assimilating findings from usability research, design guidelines for hypertext were formulated. Experiments were carried out to investigate these objectives: (i) whether design principles are embedded in the design of commercially-produced hypertexts; (ii) the relationship between implementation of design principles and end-users' attitudes; and (iii) whether hypertext prototype built using the design principles is "good enough". The analysis of these experiments draws interesting insights regarding design principles and guidelines for hypertexts.

- Chapter 4 proposes a systematic, engineering task-based approach to understand end-users' behaviour and the common tasks they perform when navigating hypertexts. Using the task browsing as a concrete example, this chapter describes a task-based approach on how common tasks can be broken down into simpler subtasks, taking into consideration end-users' reasoning and learning. The results of this cognitive task-based approach are used to guide the design of a prototype hypertext. To gather information about the usability or potential usability of hypertext, evaluation methods and techniques are investigated. Real user testing is recommended for collecting and analysing real data, providing designers with feedback from potential users. Non-user testing methods were also examined. The potential of non-human users as a cost-effective means to rapidly iterate and test design is also explored. The chapter concludes by discussing how the Conceptual Design stage in the iterative, conventional lifecycle can be improved by incorporating cognitive task diagrams and executable user models.

- Chapter 5 begins by examining the need to build well-structured hypertexts from a human factors perspective. One proposition to help designers is through improved authoring to create more comprehensible structures. An empirical study was conducted to understand the relationships LIH has with hypertext structure and user representation of the hypertext structure. The results of the investigation suggest that it is not sufficient for designers to ensure that a hypertext is well-structured, designers must also communicate that design structure clearly to end-users. It also investigates how hypertext might be structured to maximise usability by: (i) identifying ways to structure information in hypertext; and (ii) conducting a study into the types of structures designers commonly used to represent hypertexts. Mathematical insight into hypertext structural design is examined to enable designers to ask analytical questions that relate to getting "lost" and other usability problems. By identifying useful heuristics to analyse the structure of hypertext, designers are in a better position to guarantee that well-structured hypertexts are built.

- Chapter 6 begins with identifying authoring demands faced by hypertext designers. The findings of the investigation into end-users' and designers' satisfaction and expectations with popular authoring tools indicate a need for better authoring tools. From these findings, a list of helpful authoring features is identified as essential in helping designers cope with the demands of authoring. Combining with new ideas generated from the multi-disciplinary approaches described in chapters 3 – 5, this chapter proposes the building of a hypertext research authoring tool called HyperAT. To demonstrate the ideas, the World Wide Web was selected as a hypertext example since it is the largest and most widely used hypertext. This thesis acknowledges that although the ideas proposed in the four approaches are crucial in addressing the LIH problem, and ignoring them would minimise the usability of hypertexts. In practice, however, the pressures of time, money and resources often leave designers having to come to a compromise between implementing quick-fix solution as opposed to searching for the best solution.

Section III
Applying hypermedia research on the "lost in hyperspace" onto the World Wide Web

- **Chapter 7** begins by introducing the web, reviewing the LIH problem and surveying solutions to the problem. Because the LIH problem is a complex one, a framework is required to understand the design and usability issues and that these issues cannot be seen as purely psychological or computing: they are multi-disciplinary. HyperAT, a prototype web authoring tool, was developed to demonstrate these ideas. This chapter describes the design and implementation of HyperAT. The approach taken in HyperAT is novel in that established HCI elements were integrated and implemented to ensure proper structuring and presentation of hyperdocuments, as well as to provide different modes of usability evaluation of the hyperdocuments. As a specific computer science contribution, HyperAT aims to help hypertext designers build usable hyperdocuments on the web without themselves getting "lost", and hypertext end-users in navigating hypertexts produced by HyperAT without getting "lost".

- **Chapter 8** gathers qualitative results and impressions on the usefulness and usability of HyperAT as a web authoring tool. The activities carried out include: (i) using HyperAT to re-construct an existing web document; (ii) comparing the web document created by HyperAT with a similar web document (in content) produced by a standard HTML editor; (iii) getting feedback from hypertext designers; and (iv) comparing HyperAT with some web authoring and management tools. These informal evaluation activities of HyperAT indicate that HyperAT does provide the appropriate framework to help designers cope with the complexity of the design and validation processes without feeling "lost".

Section IV
Concluding remarks

- **Chapter 9** summarises the work done and opportunities arising from the work at the time of writing up this thesis. The chapter aims not only to highlight the technical aspects of the work but also to relate the author's experience in completing this PhD. By doing so, the author is able to share the complete experience of the work, which would be of use to anyone who may wish to continue the work or embark on a similar research project. Of course, the complete experience includes low level detail, but this is relevant to the overall results and may have more than incidental interest to researchers pursuing this work. This chapter reflects on how this thesis has met its objectives, the successes and the pitfalls encountered. Lessons learned from software experience practices in selecting a suitable programming language, design of prototype hypertext and HyperAT are recorded. It also describes the interdisciplinary working experience gathered while working on this thesis. It summarises the findings of this thesis and highlights limitations and the inevitable things not done at the time of writing up this thesis. This is followed by propositions on future work which consider relevant issues that are beyond the scope of current work, but merit further investigation. Several of these concern HyperAT, and its potential future development to correct its deficiencies as well as extend its functionality.

- **Chapter 10** presents final conclusion of this thesis.

Section V
Supporting materials

- The final part of this thesis contains a list of acronyms and abbreviations used, figures, tables, references, and appendices. Appendices include questionnaires, sample programs, data and papers published.

Chapter 2

Hypertext and the “lost in hyperspace” problem

Chapter objectives:

This chapter introduces hypertext and the LIH problem. Its aims are:

- to explain what a hypertext is
- to list the advantages of hypertext from authors' and end-users' perspectives
- to identify the uses of hypertext technology
- to classify hypertexts according to established dimensions
- to define the LIH problem
- to explain the motivation for solving the LIH problem
- to map out the methodology and the approach taken to address the LIH problem

2.1 Introduction to hypertext

When Vannevar Bush (1945) envisioned his hypertext "memex", he dreamed of a personal microfiche-based system that would help him tackle the problem of information overload at that time. His vision of the "memex" heralded the beginning of a search for a system that mimics the human mind to access information quickly and intuitively by reference. In 1965, Ted Nelson (1987) coined the term "hypertext" and presented it as a radical new way of structuring textual information into non-sequential format, a computer-based incarnation of Bush's dream "memex" (Berk and Devlin, 1991). Even though the technology in the "early" years of the hypertext history (for example, Conklin, 1987; Nielsen, 1995b, *etc.*) was not sophisticated enough for many of the ideas to be realised, hypertext pioneers staunchly believed that hypertext technology had something special to offer. The hypertext systems that have evolved over the past five decades give us some clues on the different perspectives taken by hypertext authors with regard to what they think hypertext is, what hypertext could be used for, and the then available technology that supported its implementation. This thesis will look at some of the more well-known hypertext systems (see figure 2.1) before defining what hypertext is, as generally agreed by the hypertext community:

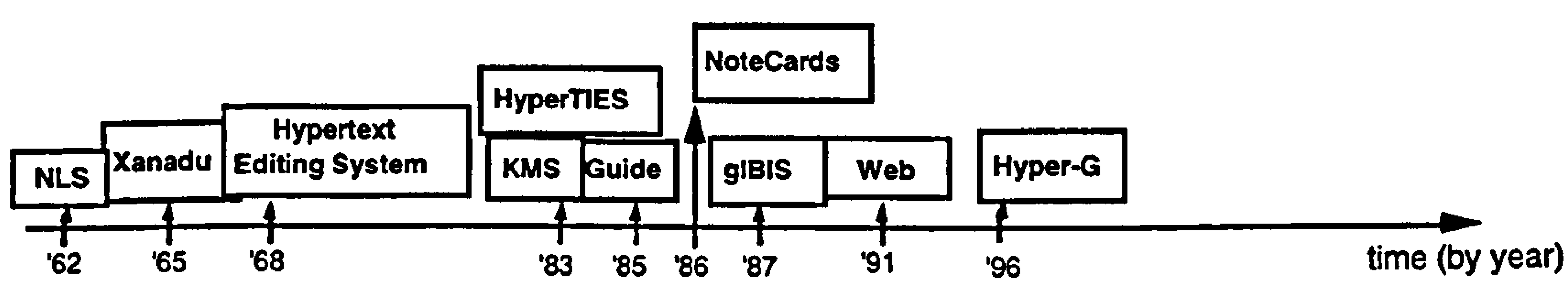


Figure 2.1. Some well-known hypertext systems developed over the years

- *Engelbart (1963)* built the *NLS* (oN-Line System) in 1962, which enabled researchers to store all their papers, reports, and memos in a shared facility where they could also include cross-references to other work. Although NLS was not developed as hypertext system, it had several hypertext features. Engelbart sees the potential in hypertext to link related items to augment human capabilities and productivity.

- *Nelson (1965)* had grand vision of what a hypertext was. His *Xanadu* vision consisted of developing a repository for everything that anybody had ever written, pushing the idea of a truly universal hypertext. This vision, however, has never been implemented.

- *Van Dam (1987)* developed the first working hypertext system on a mainframe, *Hypertext Editing System* in 1968, which was intended to produce printed documents nicely and to explore the hypertext concept. Van Dam sees hypertext as consisting of branches "holding" documents (text) together. These branches are uni-directional links that allow end-users to move within the documents.

- *Akscyn, McCracken and Yoder (1988)*, who developed the *KMS* (Knowledge Management System) in 1983, a commercial successor to ZOG, see hypertext as a potentially large database of small (screen-sized) segments called frames which can be viewed one at a time. These frames are generally hierarchically linked, though cross-referencing can be included.

- *Shneiderman and Kearsley (1989),* who supervised the development of the Interactive Encyclopaedia System *HyperTIES* in 1983, a tool for browsing in instructional databases, see hypertext as a database that has active cross-references and allows the reader to "jump" to other parts of the database as desired.

- *Brown (1987)* invented the first commercial hypertext authoring system in 1986 for a personal computer, *Guide,* which was developed to investigate computer display methods. Brown visualises hypertext as a set of documents, each of which is displayed in one scrollable window. Links in Guide either replace a piece of text with another, or import a piece of text from another document. As such, end-users are always left in the original document with information changing levels of detail.

- *Halasz (1987),* who developed *NoteCards* in 1986, thinks of a hypertext system as a hierarchical electronic generalisation of paper notecards with a title and some information associated with them. NoteCards' nodes can be more than one type and they are linked via link icons represented on the screen. These link icons are uni-directional and typed.

- *Conklin (1987),* who was involved in the development of *gIBIS* (graphical Issue Based Information System), a system that supports collaborative work among team members over a network during the whole design process, sees hypertext as a computer-based medium for thinking and communication. Thinking is non-linear, proceeding on several fronts at once, developing and rejecting ideas with each idea depending on and contributing to others. Nodes can be thought of as representing single concepts or ideas, internode links as representing semantic dependencies among these ideas. He also refers to hypertext as the "generalised footnote".

- *Jonassen (1989)* asserts that a hypertext structure mimics the associative networks of human memory and that associative network theories such as schema theory and active structural networks have provided the conceptual foundations for the development of contemporary hypertexts.

- *Knuth and Brush (1990)* find it easier to define hypertext system based on two general functions associated with the hypertext environment: browsing, the ability to retrieve information in large information spaces; and authoring, the ability to create non-sequential text.

- *Berners-Lee (1995),* who was instrumental in the development of the World Wide Web in 1991, sees hypertext as a distributed network of information that can be accessed over the Internet.

- *Maurer (1996),* who headed the team at the Graz University of Technology for the development of Hyper-G, and its commercial version, HyperWave, sees hypertext as an organised network of databases over the Internet.

The list of hypertext systems given above is by no means exhaustive. Not surprisingly, as the years roll by, hypertext systems grow more and more sophisticated and the computer is the technology that has enabled the concept of hypertext to be seen and not just heard. For example, computers have greater processing power than before to adequately support complex hypertext systems running into thousands of nodes. Technological advancements in microelectronics, *etc.* enable high-resolution screens to be produced, hence making reading easier and more pleasant than earlier character-based ones. The growing popularity of the Internet and advancements in networking see the birth of networked hypertext systems, such as the web.

2.1.1 Defining hypertext

Interestingly enough, even though many hypertext applications are finding their way out of the research laboratories and into the market, partly owing to the proliferation of more and more powerful computers at lower cost, there is as yet no general consensus on the definition of hypertext (Smith, 1994). However, it is irrefutably agreed in the hypertext community that hypertext consists of two nuclear elements namely, *nodes* and *links*, that form the building blocks of hypertext:

- *Nodes* represent the fundamental units of information which can be as small as a word or as large as a whole book. A node can contain a combination of text, graphics, audio, video and/or other forms of data. The traditional term "hypertext" refers to a system that deals with plain text only. "Hypermedia" is used by some to stress the multimedia aspects of "hypertext". In this thesis, the term "hypertext" is used since there does not seem to be any reason to reserve a special term for text-only systems and that the same usability issues discussed in this thesis apply to all of them.

- *Links* indicate the computer-assisted associations between two nodes. They are either uni-directional or bi-directional. Links can be categorised: explicit (hard-coded links) or implicit (keyword links or links that in one way or another emulate the human mind's associative mechanism); hierarchical (organisational links) or non-hierarchical (referential links).

Figure 2.2 shows a simple hyperdocument consisting of three nodes. Nodes are often referred to as pages on a computer screen. Node A has two links: a bi-directional, hierarchical link going to/from Node B and a uni-directional, referential link going to Node C. Nodes B and C are connected by a bi-directional, hierarchical link.

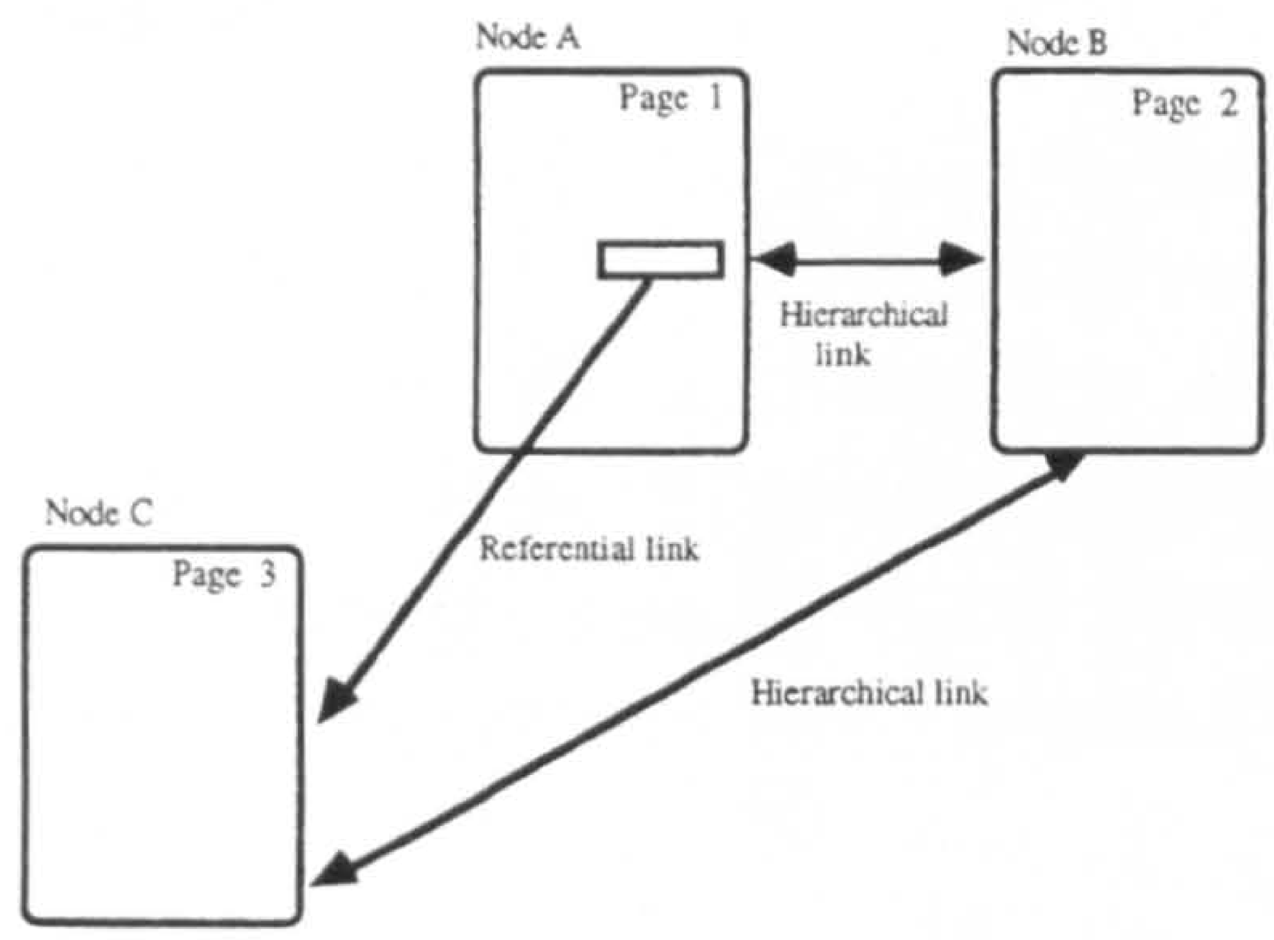


Figure 2.2. A simple hyperdocument showing how three nodes are linked

Of course, hyperdocuments vary in size, in terms of the number of nodes and links present. For very large hyperdocuments, they can run into thousands of nodes and links thus producing an intricate, interconnected network of nodes and links. More precisely, hypertext is an approach to information management in which data is stored in a network of nodes and links (Shneiderman and Kearsley, 1989). Hypertext technology has opened up new possibilities of extending the traditional notion of "flat" text files to more complex organisation of materials.

In this thesis, the term "hyperdocument" or "hypertext" is used to denote one specific hypertext document that is made up of interlinked pages or nodes. Whereas the term "hypertext system" refers to a set of software tools used to create a hypertext. Examples of hypertext systems are: NoteCards; HyperCard; Intermedia; *etc.* It is to be noted that a single hyperdocument in different versions, can be created using different hypertext systems.

2.1.2 Advantages of hypertext

The essential advantage of hypertext is the ability to organise text in different ways depending on differing views. This is where the power of hypertext lies and this thesis will now look at the operational advantages hypertext has over traditional linear text seen from both authors' as well as end-users' perspectives: how authors are empowered in the authoring process; and how end-users can assume ownership of their learning process (see figure 2.3).

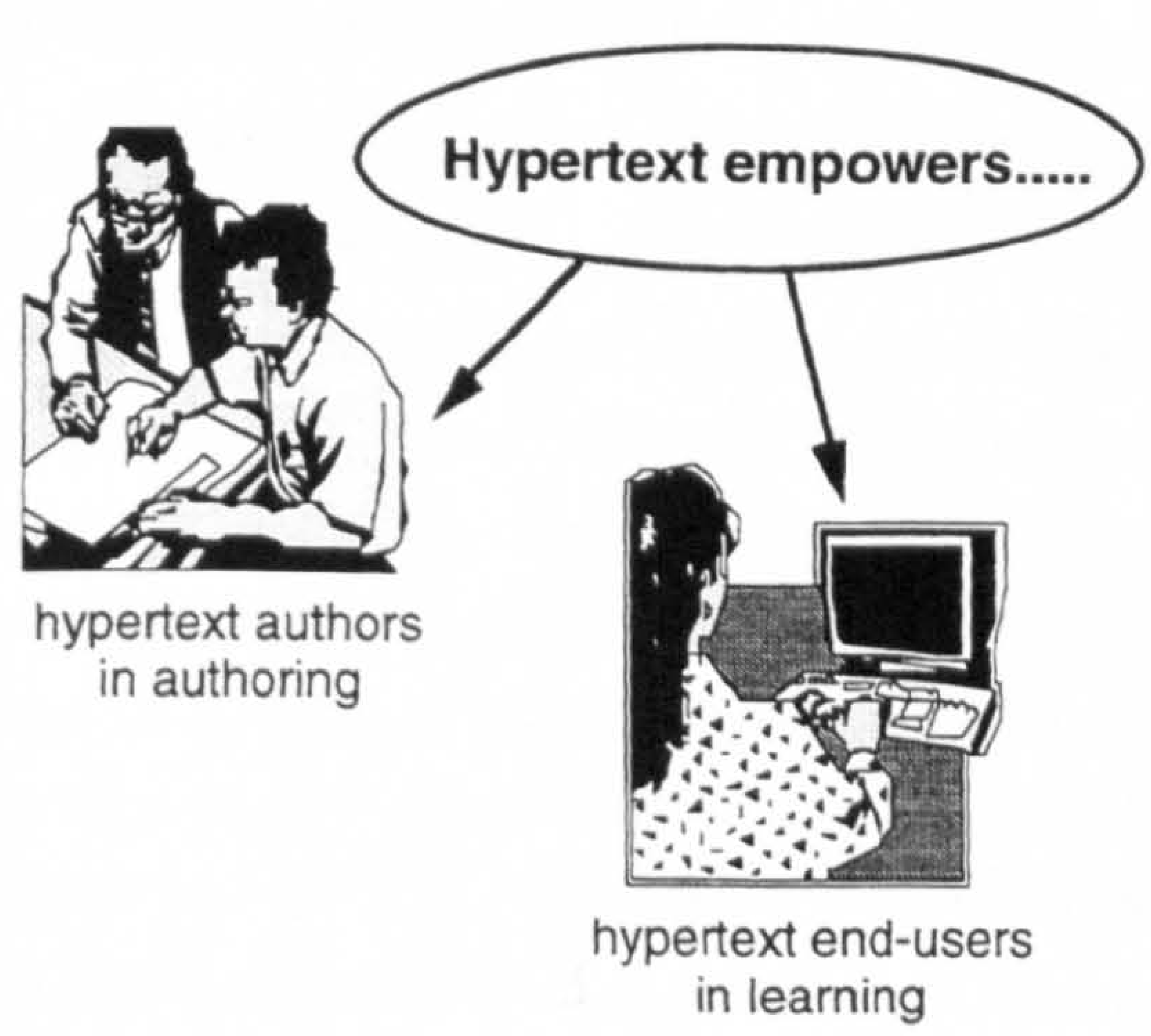


Figure 2.3. Advantages of hypertext

2.1.2.1 Power given to authors

Unlike books where there is only one way of organising the materials, that is, in a linear order, information structuring in hypertext can be more creatively done. It can be structured both hierarchically and non-hierarchically to provide an easier and more effective organisation of materials. Because the links are machine-supported, all references can be equally easy to follow. Customised hyperdocuments can be created since text segments (for example, a page, a card, *etc.*) can be threaded together in many ways, allowing the same document to serve multiple functions.

Because the same text segment can be referenced from several places, information can be modularised so that there will be less overlap and duplication. The consistency of information can be maintained. For example, references are embedded in the text, and if the text is moved, the link information still provides direct access to the reference. With good authoring tools, authors can manage the design, development and maintenance of hypertexts more effectively than say, an author of a book with his word-processor.

2.1.2.2 Power given to end-users

In conventional books, information is organised in a linear fashion and is broken down into chapters, sections, paragraphs and sentences. Authors have considerable control of what they want the readers to read. For example, any material coming from a particular point in the text can be presumed to have been read before. This means that authors can build upon previously mentioned ideas. For readers to fully understand what is written in a book, they have to start from the beginning and read sequentially until they reach the end. In contrast, hypertext authors cannot presume that material before the current point in the authors' context has been read. However, end-users in hypertext are more flexible. They are able to jump from one area of the text to another, via hypertext links, by a simple click of a button. This also means that end-users can grow their own networks, or simply annotate someone else's document with a comment. End-users are given more autonomy to map their own learning experiences, thus making them more independent learners.

2.1.3 Uses of hypertext technology

As hypertext researchers appreciate more and more of the potential in hypertexts and the benefits authors and end-users derive, it is no wonder that since its inception in the 60s, a plethora of hypertext systems has appeared ranging from scientific work to business and education needs to "general way of life". It has influenced us directly, or indirectly, in almost every facet of our lives, for example, teletext kids' stories, questionnaires, Bible cross-references, training programs, *etc.* In scientific research, hypertext technology has provided useful tools and conventions for managing and retrieving information from huge multimedia databases. It is also instrumental in integrating three technologies and industries that have been separate until recently: publishing, computing and broadcasting in the form of television and film (Nielsen, 1995b). In hypertext, we are provided with a powerful flexible mechanism to browse non-sequentially via a network of richly connected nodes and links, and extract that part of the information that is relevant to us. With so much information bombarding us daily, perhaps hypertext is a viable solution to help us cope with the problem of information overload.

Computer scientists are excited by the potential inherent in hypertext. They think that soon we may be able to have computer-based filtering systems with hypertext structure that can filter off and select only those information that are of interest to us. Hypertext is a potential solution to problems that involve voluminous, densely cross-referenced databases that must be searched by many different people attempting to retrieve highly divergent types of information (Berk and Devlin, 1991). Therefore, large technical manuals intended to provide quick solutions to very specific problems are prime candidates for hypertext. Using node associations for information, it is predicted that hypertext readers are able to assimilate information much easier than they would have with traditional texts (Beeman, Anderson, Bader, Larkin, McClard, McQuillan and Shields, 1987). Because hypertext works in collaboration with end-users, who have the intelligence to understand the semantic contents of the various nodes and determine which of its outgoing links to follow, end-users can even go on to develop new methods of learning. Educational, multimedia packages are cashing in on this benefit and flooding the marketplaces, enticing eager learners to embark on this new, exciting way of learning. Organisations such as schools, professional bodies and businesses recognise the benefits are also turning to these packages to enhance their own teaching and training programmes.

Finally, hypertext provides the opportunity to publish information structures to the general public in much the same way as books and newspapers are currently published. The web is a good example of the amount of information exchange, co-operation and collaboration that can take place among a group of end-users separated geographically across the world. With the promise of the "Information Superhighway", our lives will be greatly changed: shopping, education, entertainment, *etc.* will take on different forms.

2.1.4 Classifying hypertexts

Because there is generally no accepted definition of hypertext, it is difficult to categorise hypertexts based just on necessary and sufficient features (Knuth and Brush, 1990). Figure 2.4 illustrates some dimensions that have been used to classify hypertexts: control authors have; kinds of support provided; who it is designed for and by whom; and size of the hypertext. Other dimensions can also be considered, for example, how hypertext is structured, how the information content is managed, *etc.*

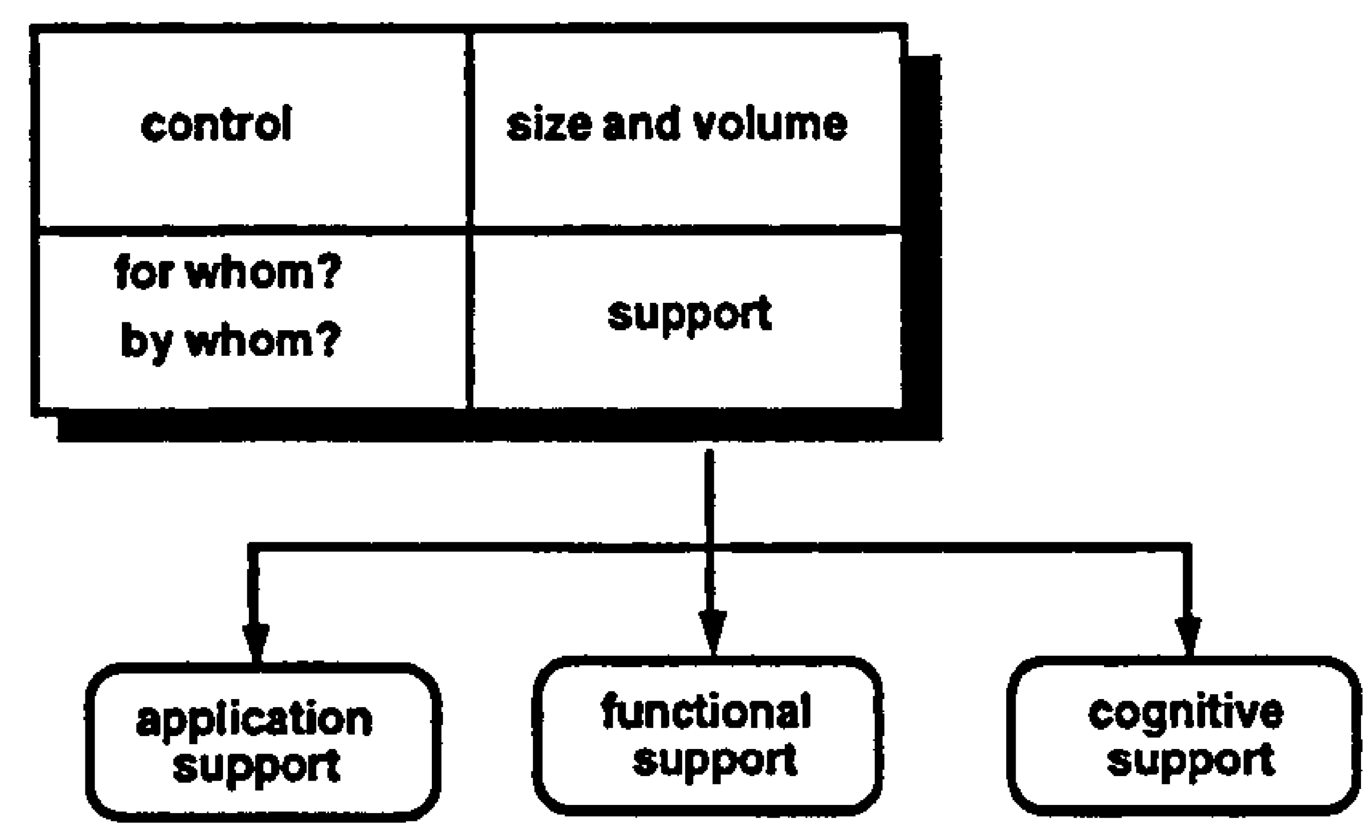


Figure 2.4. Dimensions to classify hypertexts

- *Control.* Maurer (1996) suggests that an important criterion to classify hypertext is to consider how much control over a hyperdocument's look and feel an author should be allowed to exercise. Traditionally hypertexts are frame-based, that is, documents must fit into a fixed-size frame (screen with or without scrolling) and authors have full control over content and presentation. Examples include KMS, HyperCard, Toolbook, *etc.* However, in window-based hypertexts, documents can be of any size, and each is displayed in its own (scrollable) window. Authors typically specify only the information content of documents, while style sheets control their presentation.

- *For whom? By whom?* Hypertexts can be classified based on who they are designed for and by whom. Hardman (1988) proposes these categories: (i) single end-user, single author; (ii) many end-users, single author; and (iii) many end-users, many authors. In (i), the author uses his own hypertext. In (ii) and (iii), the hypertext is produced for distributed usage. (iii) is the case where some of the end-users may be authors themselves.

- *Size and volume.* Rada and Murphy (1992) describe hypertext in terms of their size: traditionally, small volume hypertext supports browsing while large volume hypertext supports searching. Because hypertext differs in the size of the information base and the size of the user population, hypertext can be classified to reflect the differences in the underlying storage mechanisms, and user interfaces (Halasz, 1987).

- *Applications support.* Many diverse hypertexts have been designed and built to meet a variety of roles: tutorial and educational needs; support for collaboration among a number of individuals and production of on-line manuals. Conklin (1987) suggests grouping hypertexts under these four broad application areas: (i) *macro literary systems* which are large on-line libraries used to support publishing, literary criticism and collaborative working; (ii) *problem exploration tools* used to support early authoring and outlining, problem solving, programming and design; (iii) *browsing systems* used to support teaching, reference and public information; and (iv) *general hypertext technology* are general purpose systems for reading, writing and collaboration. Halasz (1987) suggests that because hypertext is usually designed with a target domain in mind, the features and capabilities of the system reflect the requirements of the target. For example, Intermedia

(Garrett, Smith and Meyrowitz, 1986) is designed for educational applications, and therefore emphasises interactive displays and annotation facilities.

- *Functional support.* Knuth and Brush (1990) advocate that it is impossible to find a set of features which are embodied by all hypertexts except browsing and authoring. Because browsing systems encourage end-users to browse, search for information and seek references, well-developed tools for information presentation and exploratory browsing are provided. Authoring systems, however, provide well-developed tools for the creation and modification of networks (Halasz, 1987).

- *Cognitive support.* Carlson (1990) distinguishes systems based on the type of cognitive activity supported such as reading, annotating, collaborating and learning. Recall, which is related to learning, can also be used.

2.1.4.1 The scope of the investigation in hypertext

Since it is impossible to investigate all aspects relating to hypertexts, this thesis focuses the investigation of the LIH problem on hypertexts with the following characteristics (see table 2.1).

Table 2.1. Types of hypertexts this thesis is concerned with

Application support	Functional support	Cognitive support	For whom? By whom?	Control
browsing systems used to support teaching	browsing authoring	reading learning recall	single author many users	frame-based window-based

This thesis will also extend the investigation onto the web, the most widely used hypertext system on the Internet[1] (Maurer, 1996), if not the world. §2.1.4.2 describes briefly what the web is and problems associated with it.

2.1.4.2 Hypertext on the Internet: Introducing the World Wide Web

Like hypertext, the web provides non-linear information spaces in which "chunks" of information are connected to other "chunks" of information via "links". When the web was developed in 1991, the intention was to link a select group of end-users such as physicists and engineers at different sites (Baecker, Grudin, Buxton and Greenberg, 1995b). Today, the web is used by millions of end-users all across the world. The web has changed the Internet to the extent that it has become almost synonymous with the modern use of the Internet (Maurer, 1996). The most common use of the web is for browsing, followed by entertainment and academic and business purposes as reported in the Graphic, Visualisation and Usability Center's 4th Web User Survey (Pitkow and Kehoe, 1995).

[1]Internet is the name for the interconnected set of computer networks around the world.

The web is here to stay and for a long time too! It forms part of the answer to the much-awaited "Information Superhighway" that will allow people to fulfil their information needs easily, quickly and cheaply. So, whenever hypertext is referred to in this thesis, it also includes the web since many of the usability issues with which this thesis is concerned, also apply to the web.

2.1.5 Problems with hypertext

Though hypertext is great, it is not a panacea to life's problems (Nielsen, 1995b). Associated with hypertext are two classes of problems (Conklin, 1987): (i) problems with current implementations, which include delays in the display of referenced materials, deficiencies in browsers, *etc.*; and (ii) problems that seem endemic to hypertext such as cognitive overload and disorientation. Cognitive overload is the additional effort and concentration necessary to maintain several tasks or trails at one time. Disorientation is the tendency of end-users to lose their way in the maze of non-linear information. This is commonly referred to as the LIH problem.

This thesis is concerned primarily with addressing the LIH problem in hypertext. §2.2 defines what LIH is and examines the reasons for this LIH problem. §2.3 presents the argument justifying the need to address the LIH problem chosen from among the many usability issues associated with hypertext, which some think should take higher priority.

2.2 Revisiting the "lost in hyperspace" problem

Although much research effort has been invested to address the LIH problem, it still remains unresolved. This thesis argues that perhaps wrong or inappropriate solutions are being sought because incorrect or incomplete assumptions are made. Hypertext design is hard, and hypertexts are used less effectively than we would wish. The question to ask is:

Is LIH primarily psychological or engineering?

The answer to this question will have serious implications on the solutions being sought to address the LIH problem. If LIH is a *psychological* problem, then the problem may be entirely due to end-users' inability to exploit computer screens, complex information structures, and that nothing in the design is going to ameliorate this. Thus, as a psychological problem it can be alleviated but not solved by better design.

Though disorientation can arise in conceptual space (within the end-user's mind), which most research findings support — this thesis argues that research should not rest on end-users alone! The LIH problem may not just be a *psychological* problem — it may be an *engineering* problem. This implies that LIH is perhaps attributable to bad system design, and poor design causes psychological problems too.

As an engineering problem, it is meant that in contrast one is concerned with issues such as:

- how do designers design human-machine systems so that LIH is not the issue for the overall performance of the tasks (even though the end-user may experience LIH);
- how do designers design a system to avoid/reduce/manage LIH given that it is a requirement for the task to do so; and
- how do designers identify tasks where LIH is the issue?

2.2.1 Defining the "lost in hyperspace" problem

Ironically, the advantage of hypertext in providing end-users with the freedom of reading, and putting them in control to decide which nodes and links to follow, is the very disadvantage of hypertext that end-users are at the risk of taking the wrong turn or getting "lost". It is possible for end-users to experience this feeling of disorientation even though they may faithfully follow a connected trail of hypertext links, and each link in the trail may make perfect sense. However, by the time end-users are several links deep, the relevance of their current position in the hyperdocument to where they start from may be far from clear.

Researchers in their independent research work identified that the LIH problem experienced by end-users can refer to any of the following states (Conklin, 1987; McKnight, Dillon and Richardson, 1991; Nielsen, 1995b, *etc.*) as shown in table 2.2:

Table 2.2. Different states of the "lost in hyperspace" phenomenon

- the problem of not knowing where they are in the network
- how to get to some other place they know (or think) exists in the network
- how to return to a topic left previously
- the problem of forgetting why they want to go to a certain topic in the first place
- the problem of forgetting what topics they have browsed
- the problem of forgetting which key points covered

To understand the reasons for the LIH problem, the LIH problem is examined from two perspectives: (i) within the end-user's mind (§2.2.2); and (ii) within hypertext (§2.2.3). We argue that "lostness" may occur within the end-user's mind, leading to *conceptual disorientation*, or it may occur within hypertext, leading to *spatial disorientation* (see figure 2.5).

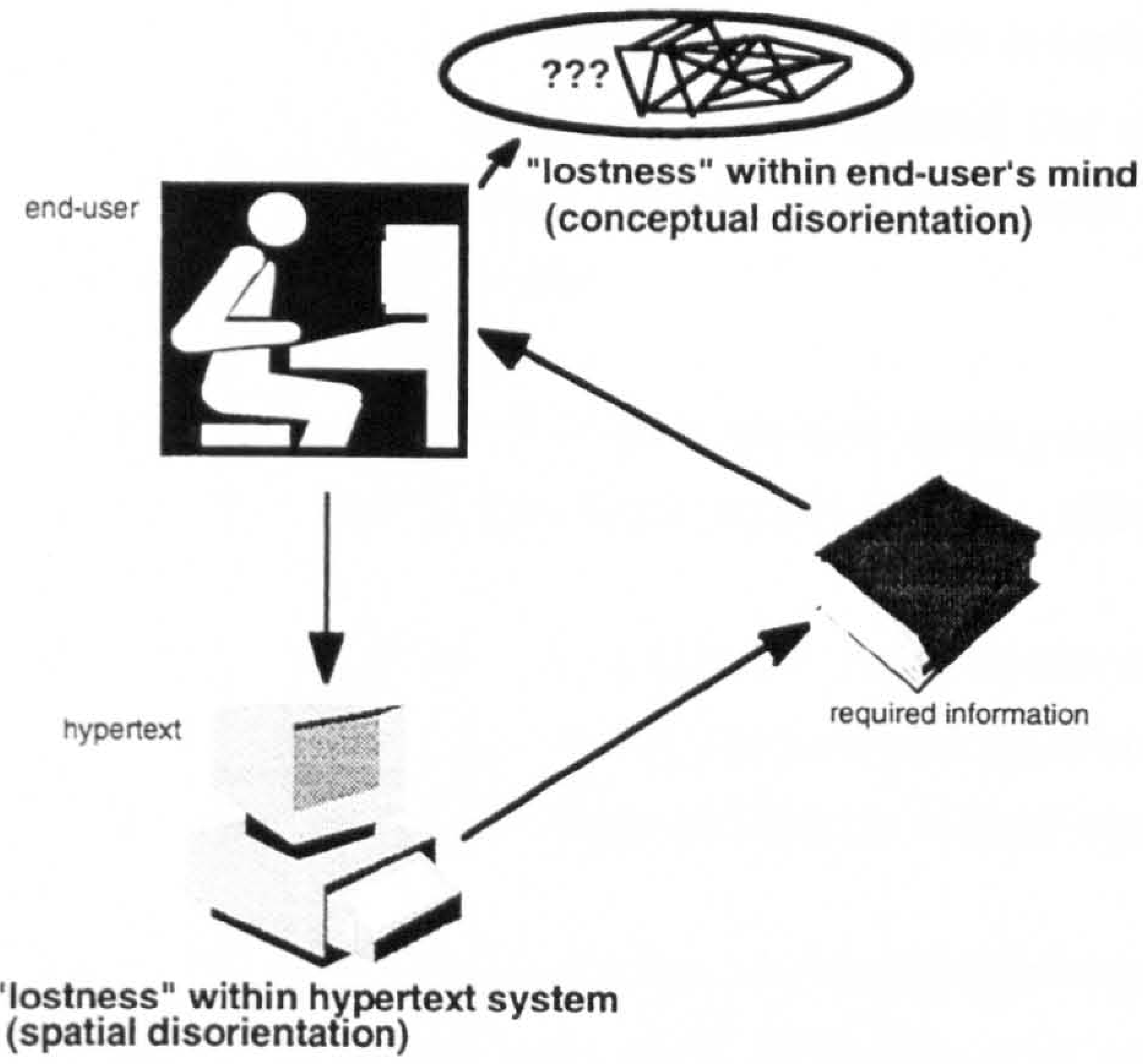


Figure 2.5. "Lost in hyperspace" seen from two perspectives

2.2.2 Conceptual disorientation: Inside the end-user's mind

"...The human mind... operates by association..." (Bush, 1945). Hypertext attempts to mimic the brain's ability to access information quickly and intuitively by reference. However, this grand vision is hindered by the LIH problem experienced by end-users. Elm and Woods (1985) define what it means for end-users to experience this "getting lost" phenomenon while navigating hypertext:

> "Getting lost in a display of networks means that the end-user does not have a clear conception of the relationships within the system, does not know his present location in the system relative to the display structure, and finds it difficult to decide where to look next within the system."

Most people think the LIH is an *end-user's problem*, resulting in improvements being sought in the presentation of information. Not surprisingly, therefore, many research solutions involve the use of graphical browsers and query/search mechanisms. They seem to make the following assumptions:

- End-users have a wrong or incomplete conceptual model of how information is structured and linked within the hypertext (Elm and Woods, 1985).
- End-users experience "lack of closure" since they are not able to tell the extent of a network or what proportion of relevant items remains to be seen (Shneiderman and Kearsley, 1989).
- End-users face the "embedded digression problem" where they lose track of digression since they are distracted from the main tasks by lots of interesting information (Foss, 1989).

- End-users generally lack the experience in using hypertext for learning, and this makes it difficult for them to remember, consolidate and understand the semantic content of nodes, resulting in a lack of detailed memory of any particular item and an inability to summarise what has been covered (Foss, 1989).

- End-users do not have adequate systems, that they should have graphical browsers and query/search mechanisms (for example, Edwards and Hardman, 1993, *etc.*).

- End-users do not use their senses fully. They should have multi-modal experiences, fully using sensory and motor skills, audio, stereolocation and so on (Coutaz, Nigay and Salber, 1993). This view leads to solutions being sought in virtual reality.

Note that all assumptions are referenced to the *end-users'* supposed failings: *end-users* have a wrong or incomplete conceptual model; *end-users* lack experience in using hypertext for performing tasks such as browsing; *end-users* are distracted because of the "embedded digression problem"; and *end-users* do not understand the chosen display conventions. And because hypertexts are difficult to build, many commentators are happy to seek solutions in better understanding *end-users* and helping *them* cope. There has, of course, been some success in this approach — maps, virtual reality visualisation — that, no doubt, seem to confirm it!

2.2.3 Spatial disorientation: Within hypertext and on the World Wide Web

Though disorientation can arise in conceptual space (within the end-user's mind), which most research findings support - the thesis argues the "blame" should not rest on end-users alone! Many hypertexts are, however, poorly designed and built in terms of how information is structured and displayed. When Brown (1990) conducted an assessment on the quality of hypertexts designed by seventy student authors, poor design of the visual appearance of material, overuse of technology with lots of clever effects but no attempt made to re-design for the new medium, lack of a coherent overall structure and presentation style were some prevalent faults identified. It is, therefore, not surprising that some of the problems in hypertexts that contribute to the LIH problem are due to poor designs caused by bad authorship.

Brown (1990) argues that "although getting lost is often claimed to be a great problem, the evidence is largely circumstantial and conflicting. In smallish applications, it is not a major problem at all". Brown's remark raises an important issue: when we speak of documents being so small that end-users cannot get "lost" in them or so large that navigation aids are required to use them effectively, the implication is that information occupies "space" through which end-users "travel" or "move" (McKnight, Dillon and Richardson, 1991).

There are two aspects to LIH (Mayes, Kibby and Anderson, 1990): firstly, navigational problems, where the geography of the network may be too complex to grasp even with the aid of a map or virtual spatial environment; and secondly, the astronomical number of nodes and links creates problems of scale unless some form of filtering mechanism is employed. They argue that navigational aids do not help end-users navigate in conceptual space (within end-users' minds) but only tell end-users where information is located.

This thesis agrees with Mayes *et al.* (1990) that addressing the LIH problem within hypertext goes beyond providing more and more navigational aids. In fact, Tripp and Roby (1990) found that end-users' performance deteriorated if both navigational aids (for example, an advance organiser and a visual metaphor) were used at the same time. They concluded that the aids activated conflicting mental models of the hypertext structure.

Hypertext provides more locations in which to store information and more dimensions in which to travel compared with traditional linear text. There is thus greater potential for becoming lost or disoriented using hypertext. When examining a book, readers can look for orientation cues in the form of overviews, summaries, contents-pages, *etc.* However, if these cues are not present in hypertext, as in most cases, end-users may lose track of the context either through an external interruption or in the course of pursuing momentarily interesting links which prove to be dead ends (Bernstein, 1988).

The book's organisational principles took centuries to develop (Thimbleby, 1992) and we take for granted very many organisational structures in books, like page numbers and alphabetical orders in reference books. Yet, writing a good book is difficult. Creating good hypertext is even more difficult. Hypertext authors are faced with a vast range of potential structures and an astronomically large number of choices when creating a hypertext document (Thimbleby, 1995a). The central organising principles in books cannot be used in hypertext, otherwise we end up with an electronic, conventional book.

Could it be possible that because hypertext authors themselves are "lost" in the process of designing and authoring hypertexts, they inadvertently contribute to poorly designed hypertexts, which in turn leads end-users often to being LIH? How can we more quickly find better organisational principles for hypertext? This thesis argues for a move away from treatment to prevention, from treating the end-user's *symptoms* — themselves a reaction to bad design — to avoiding the bad design. Therefore, the way hypertexts are designed and built needs to be re-examined. There is a need to examine fundamentally how information should be structured and displayed, and not assume that if certain design features work well for some information contents and purposes, it will be appropriate for others.

2.3 Motivating research in addressing the "lost in hyperspace" problem

Is the LIH problem a crucial problem that needs to be addressed? Bernstein (1988) does not think so. In fact, he claims that even though disorientation may be an unsettling experience, a degree of disorientation deliberately introduced and thoughtfully controlled and guided, can be a powerful tool for designers. He cites an example of a public information system where reducing end-users' anxiety may be more important than eliminating disorientation. Mayes *et al.* (1990) suggest that for hypertexts designed for computer-assisted learning, disorientation is a prerequisite for depth learning. They explain that taking wrong turns is a necessary part of the process of a learner trying to map the information being discovered onto a developing framework of understanding, since the learner has to "make sense" of the links between nodes.

In a workshop on "The Missing Link: Hypermedia Usability Research and The Web" (Buckingham-Shum, 1996) held at the Open University, some HCI researchers and practitioners felt that there are more important and pressing issues besides the LIH problem on the web, and hypertexts in general that need to be addressed. This came about because of the results reported in the 4th Web User Survey by the Graphic, Visualisation and Usability Center at Georgia Tech Research Corporation conducted from 10 October through November 1995 (Pitkow and Kehoe, 1995). It was reported, from a sample size of more than 23 000 responses, that end-users were not "lost" and that the classical LIH problem was not a problem (6.5%), as opposed to the most widely cited problem that it takes too long to view/download pages (69.1%).

If the LIH problem only refers to "end-users not being able to determine where they are" as reported in the survey, then perhaps it may not be a pressing issue. But 6.5% of the user population of approximately[2] 30 million in 1994 on the web who reported being "lost", is certainly not a small number. The smallest of usability problems, when multiplied across thousands or millions of end-users, becomes a source of massive inefficiency and untold frustration (Nielsen, 1993). Besides the LIH phenomenon does not only refer to "end-users not able to determine where they are", it can refer to any of the following conditions (as described previously in §2.2.1): end-users cannot identify where they are; end-users cannot return to previously visited information; end-users cannot go to information believed to exist; end-users cannot remember what they have covered; and end-users cannot remember the key points covered. In the same survey, problems such as these: not being able to find a page that they know is out there (34.5%); not being able to find a page once visited (23.7%); and not being able to visualise where they have been and where they can go (14.3%) were identified as "real" problems. This constitutes an enormous number of end-users on the web, who might not report that they were "lost" but experienced different forms and degrees of "lostness". These findings provide a snapshot of the current web user population, and the problems experienced by end-users on the web. The tremendous success of the web is creating problems. The web was never designed to handle so many and such large applications as more and more people are using it (Maurer, 1996). End-users really are lost. It is not just a superficial disorientation.

Yet there seems to be some controversy among the hypertext research community with regard to whether the LIH problem is still a worthwhile problem that needs to be solved. A paper submitted to the Hypertext'97 conference received mixed reactions from four reviewers. One of them thought that there are other more pressing problems regarding hypertext that need to be tackled, while another adamantly felt the LIH problem is one of the most difficult issues in the hypertext/hypermedia community and encouraged continued research in this challenging area. The other two were quite supportive of more work being done to address this

[2]The number of end-users with access to the Internet was around 30 million at the end of 1994 (Nielsen, 1995b). Nobody knows for sure how many end-users are on the Internet because of its distributed nature.

issue. Nielsen (1996) in his CHI'96 Workshop on "HCI and the Web" also supported the fact that "end-users very easily get lost in hyperspace, and much existing hypertext research could benefit web designers as well as designers of web software". To collect more views regarding this issue, the author also corresponded with some prominent and leading researchers in hypertexts and HCI (for example, Jeff Conklin, Peter Brown, Ben Shneiderman, Keith Instone, Patricia Wright, Linda Hardman and Ian Benest), and reached the conclusion that the LIH is still a problem.

Elm and Woods (1985) claim that disorientation is more than just the end-users' subjective feeling of being "lost", it will also lead to a degradation in the end-users' ability to extract the information that is needed to perform their tasks successfully. Tripp and Roby (1990) assert that disorientation in a hypertext environment will lead to increased cognitive load which will reduce the mental resources available for learning. They explain it is logical that achievement would suffer to the extent that navigation is demanding, since the same mental resources needed for learning are also engaged in navigational tasks.

If end-users are frequently lost, they will become frustrated and this may influence the way they interact with the hypertext. Worse still, they may cease to use the system because they may feel that they are wasting their time and overlooking crucial information. This is certainly not desirable as the primary objective of hypertext is to provide end-users with information!

In safety-critical systems, the LIH problem resulting in end-users not being able to locate specific and accurate information on time, can lead to more fatal consequences. A real-life incident that happened on board an Airbus320 is a case in point (Mellor, 1995). In this incident, upon discovering that a technical fault had occurred, pilots and engineers were unable to obtain the relevant information via an on-line hypertext manual to rectify the fault. Fortunately, in this incident someone happened to go over the technical manual (on paper) the night before and scribbled something relating to this fault. Just imagine what would have happened to the many innocent lives on board the airbus if the fault were not rectified on time.

The cost of not solving the LIH problem can be great in terms of time lost, money wasted and in some cases, loss of many precious lives. The LIH problem brings more harm than good and more research efforts should be invested to find ways to eliminate, or at least to ameliorate this problem.

2.4 Methodology taken in this thesis

Building systems with effective design involves a variety of skills, drawing upon and effectively harnessing many backgrounds and disciplines (Baecker, Grudin, Buxton and Greenberg, 1995a). Hypertexts are without exception. If previous attempts drawn from singular fields like human-computer interaction or hypertext were inadequate to solve this problem, then perhaps there is a need to get help from other disciplines. Kim (1995) advocates that where one discipline is weak, we should look for another that is strong. We should also learn to build and use bridges between the many relevant disciplines and perspectives to create the opportunities for creative synergy.

Since LIH is a complex problem, a *multi-disciplinary* approach is necessary to draw upon knowledge and findings from diverse disciplines such as human-computer interaction (HCI), software engineering (SE) and cognitive psychology (CP) and integrate this knowledge into hypertext. By "discipline", this thesis refers to a body of shared concepts and criteria collectively agreed upon by researchers in the same field. The values defined by a discipline are criteria for deciding what good and bad research is. This thesis is concerned with some ways of enabling how the concepts, values, methods and procedures from these disciplines can be integrated into hypertext to address the LIH problem (see figure 2.6). §2.4.1 describes the multi-disciplinary approaches taken in this thesis to address the LIH problem.

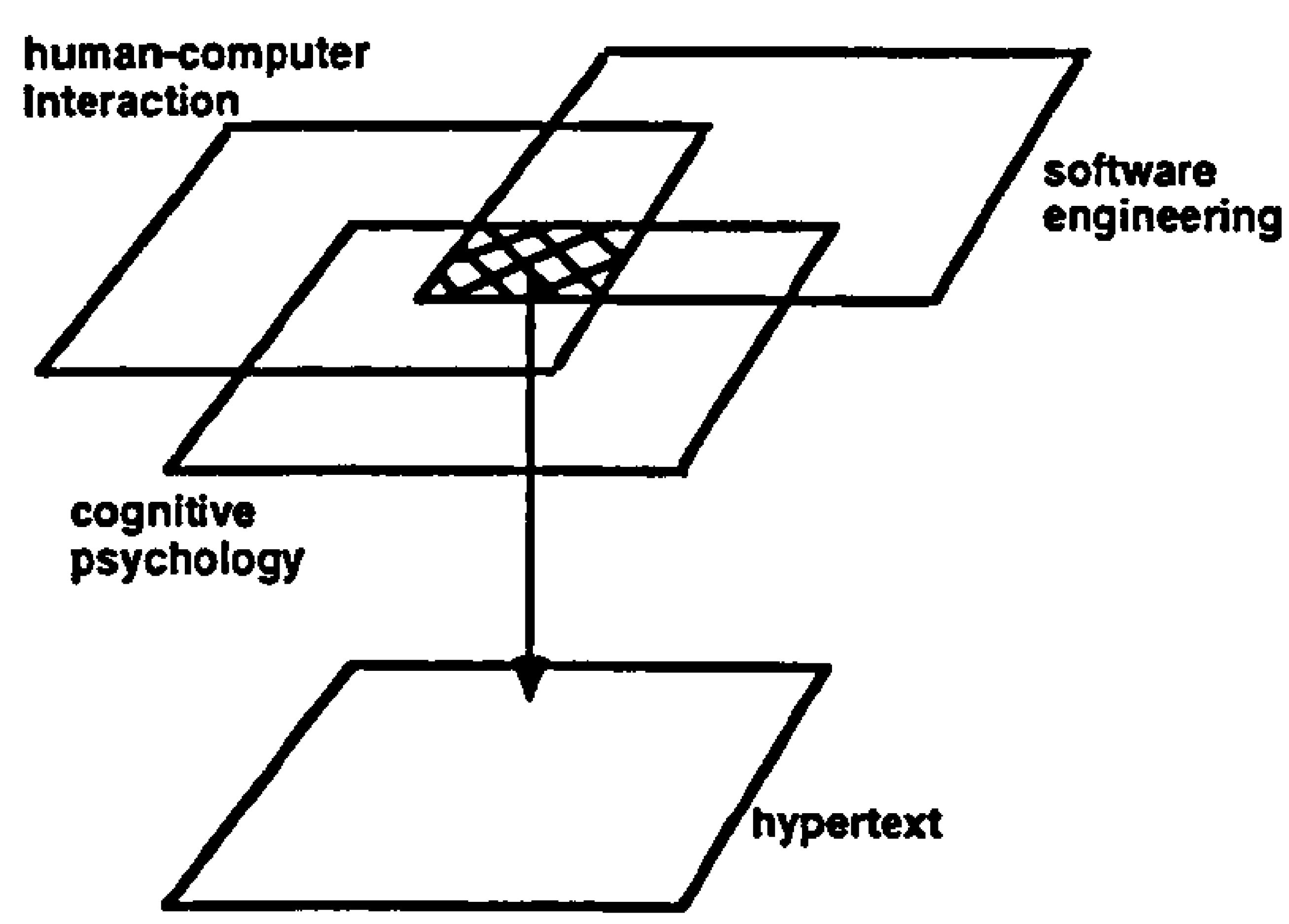


Figure 2.6. Integrating commonalties among the disciplines of HCI, CP and SE onto hypertext

This thesis, however, will not focus on how people from different disciplines can achieve better understanding when working together, which is an important human issue in practice. Understanding someone else's scientific disciplines is difficult because the values, methods and procedures for each discipline differs. For example, computer scientists have a great deal to contribute, both through the demonstration systems they create and through the abstractions they have developed. Psychologists, though concerned with abstraction and models, also place a strong emphasis on measurement and hypothesis testing. Monk (1995) suggests that scientists who consider their work to be of significance should encapsulate the understanding achieved in their investigations as procedures that can be referenced and used. Another suggestion is that individuals with varying backgrounds in design teams should come to understand the concepts and values of other team members, and come to some degree of compromise, to be an "interdisciplinary ambassador" (Kim, 1995). Monk (1995) uses the analogy of travel to describe the advantages and drawbacks of multi-disciplinary approaches to research. Because travel not only broadens the minds of the individual travellers, it also enriches the culture they return to. The effort involved in multi-disciplinary approach should not be underestimated.

2.4.1 Multi-disciplinary approaches to address the "lost in hyperspace" problem

Because this thesis is interested in taking a "breadth-first investigation" of the LIH problem to gain a wider insight and restructuring of the problem, a productive problem-solving approach as opposed to a reproductive problem-solving approach is adopted. A reproductive problem-solving approach could be a hindrance to finding a problem since it focuses on known aspects of the problem, and therefore it is not able to see novel interpretations that might lead to a solution (Dix, Finlay, Abowd and Beale, 1993). This thesis acknowledges that though the productive problem-solving approach provides a wider insight into the LIH problem, however, with constraints of time and human resources (the author herself), each approach could not be dealt with in greater depth.

Figure 2.7 outlines an overview of the methodology taken in this thesis to address the LIH problem so that usable hypertexts are produced. This thesis argues that since the LIH is a complex problem involving both engineering as well as psychological issues, there is a need to take a step back and examine more fundamental issues. Four *proactive, multi-disciplinary* approaches to eliminate, or at least ameliorate the LIH problem are proposed:

- Approach One: Need for good design principles and guidelines
- Approach Two: Need for an engineering, task-based approach to understand end-users' needs
- Approach Three: Need for good hypertext structure
- Approach Four: Need for better support tools for authoring

This thesis investigates how each of these approaches can contribute towards an integrated solution in the form of an authoring tool to address the LIH problem on the web.

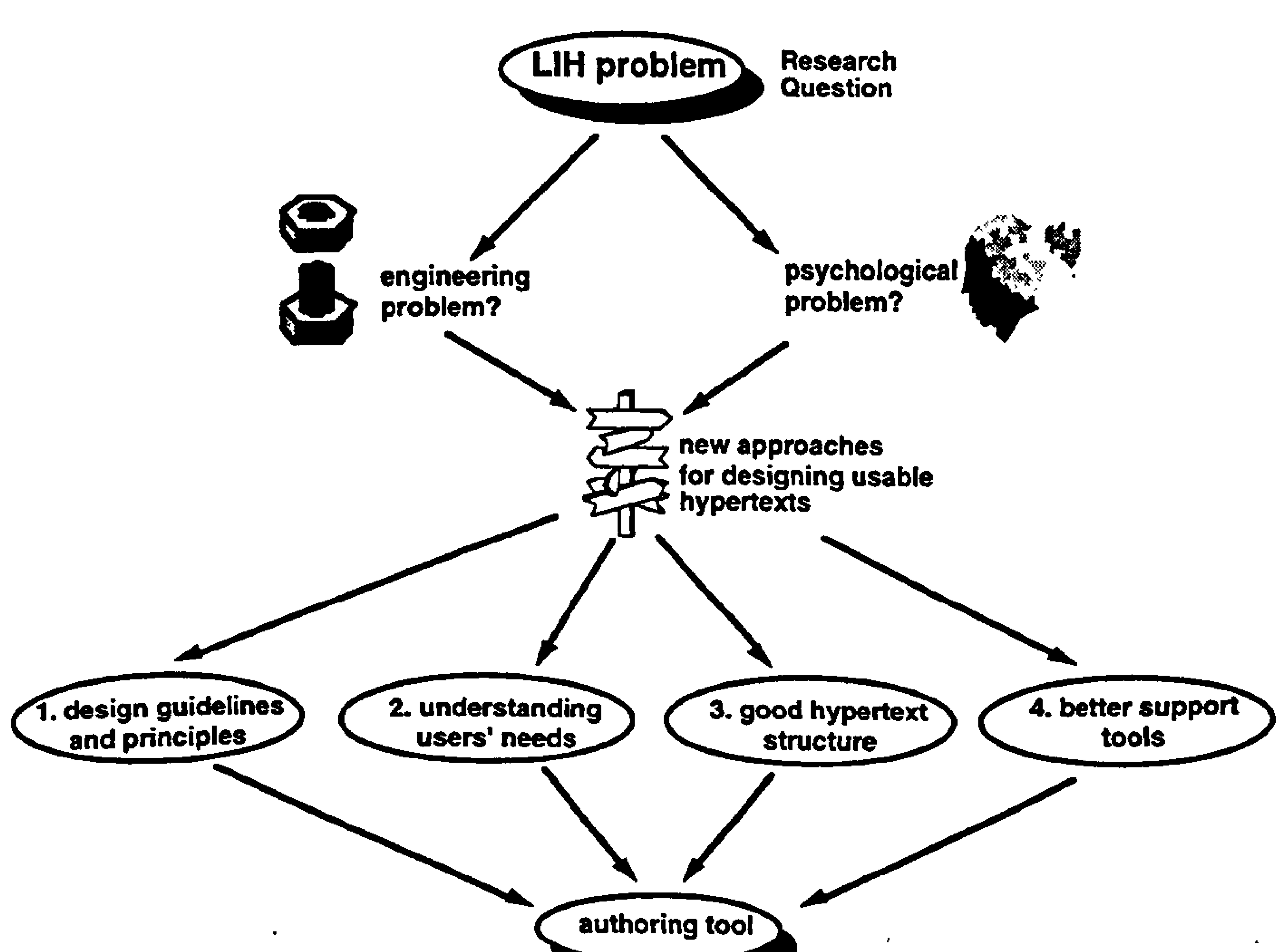
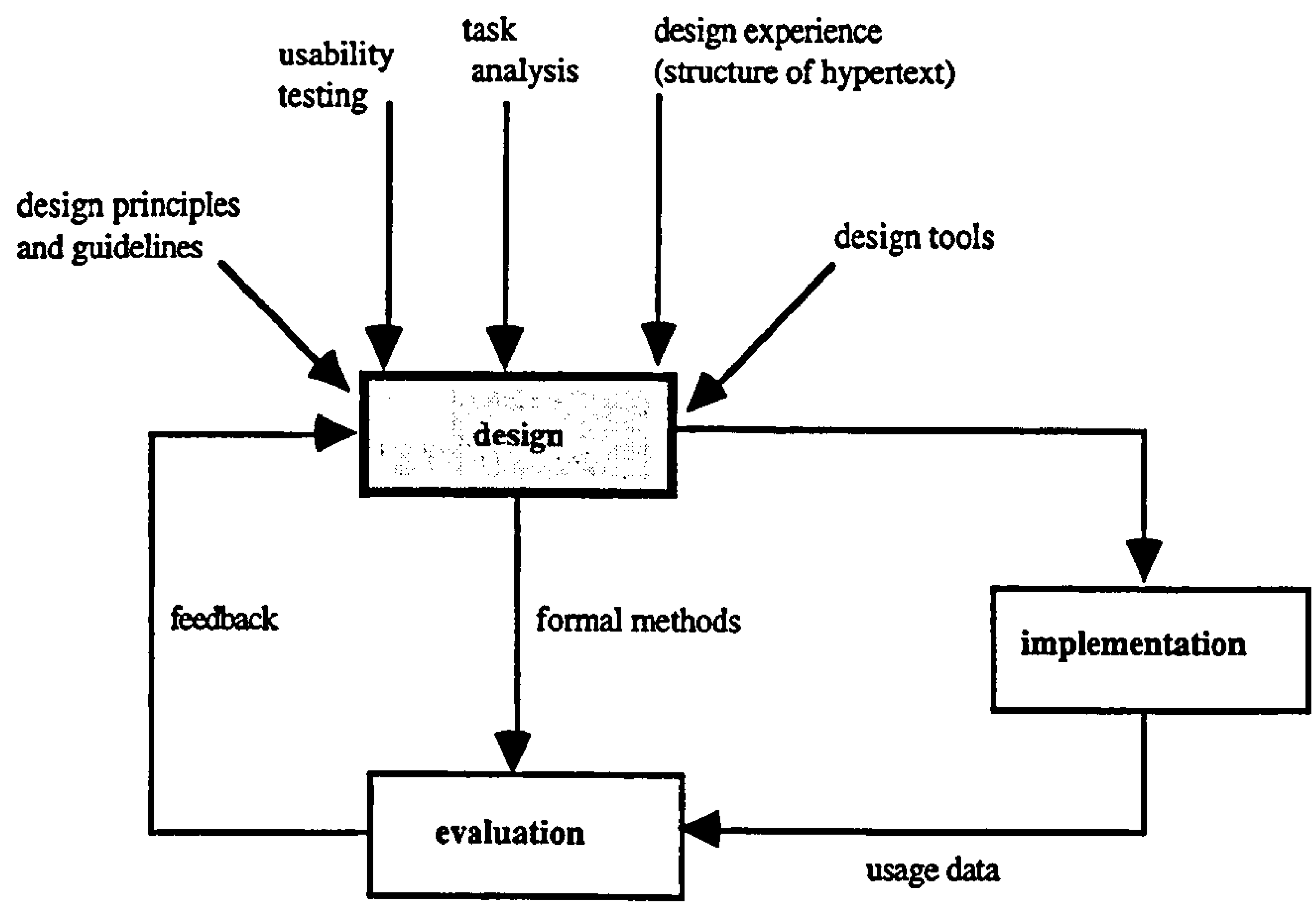


Figure 2.7. Overview of the methodology taken to address the LIH problem

This thesis proposes incorporating into the research methodology all user-centred methods recruited to realise these approaches by drawing upon knowledge and findings from different disciplines and integrating them into hypertext. These user-centred methods used in the methodology focus on the activities in the "Design Stage" of a typical user-centred design cycle as shown in figure 2.8.



Our investigation involves the activities relating to this stage, and not the implementation and evaluation stages

Figure 2.8. User-centred design cycle (adapted from Perlman, 1988)

- *Getting to know end-users and their needs* (Lewis and Rieman, 1993). It is well-known that designers often design for themselves unless they are trained to realise that people are diverse, and that end-users are unlikely to be like them (Landauer, 1995). Solutions to the LIH problem should then address the issues of helping end-users navigate through conceptual space. It is important to have an accurate representation of end-users' behaviour and actions when they perform or try to perform common tasks such as browsing, information search, seeking references and recall. By trying to make sense of what end-users should do or what they actually do, hypertext authors will at least stand a better chance of producing user-centred hypertexts that will meet end-users' needs more effectively. To do that, this thesis will investigate whether task analysis and cognitive user modelling techniques can be used to help designers understand end-users' behaviour and actions, taking into consideration end-users' reasoning and learning processes (Approach Two, see chapter 4).

- *Iterative design* (Gould, Boies and Lewis, 1991). Interactive systems require iterative design. The most promising approach is to iterate design and evaluation until a satisfactory result is achieved (Baecker *et al.*, 1995a). To make iterative design possible when facing budget or schedule constraints, a rapid prototyping approach is used to support mocking up and trying out interfaces and dialogues. Since disorientation can occur in a spatial network of nodes and links, there is a need to re-look at design issues. It is important to ensure that good hypertext design principles and guidelines are incorporated into the building of hypertext in the first place (Approach One, see chapter 3). In considering design principles and guidelines, there is also a need to investigate the "best" structure for hypertext according to their functions (Approach Three, see chapter 5).

- *Early and continual user testing* (Gould, Boies and Lewis, 1991). Because it is also imperative that hypertext authors build into the hypertexts an accurate representation of user models when performing these tasks, early and continual user testing is essential. This will ensure that the system is designed to meet end-users' needs. This thesis will explore the possibility of using [3]computerised users in the form of executable cognitive user models for the design and validation of hypertexts as early as possible in the design process (Approach Two, see chapter 4).

- *Integrated design* (Gould, Boies and Lewis, 1991). All aspects of usability should evolve in parallel, rather than be defined sequentially, and should be under one management. This thesis integrates these approaches onto a single hypertext system, the web, to help end-users navigate around it without getting "lost" (see chapter 7).

- *Prototyping* (Gould, 1988). It is clear that *ad hoc* methods of designing, constructing and validating hypertexts are not enough. If end-users get "lost" in hypertext, designers do too. This suggests that tools for designing hypertext should provide much improved computational support for designers (see chapter 6). A hypertext authoring tool is being built to help designers build well-structured hypertext prototype (see chapter 7).

- *Usability testing* (Nielsen, 1995b; Lindgaard, 1994). To find out if the prototype hypertext authoring tool is "good enough" to help designers produce a "good enough" prototype hypertext, there is a need to know how to express in quantifiable, measurable terms when a "good" system is "good enough", because failure to do so will render designers incapable of substantiating claims of its usability and improvements in a system (see chapter 4).

2.4.2 Experimental strategies

Generally, most people feel that in order to get statistically significant results, a number of people (around 20-25) should be asked to carry out the experimental task, in order to pick up a wider range of problems and to get some sense of the frequency and the consequences of each. However, Nielsen and Landauer (1993) conclude from analysing usability problems described in eleven published projects that the maximum cost/benefit ratio for a medium-large software project could be obtained through the use of *three* test users. They stress that though this is not a "magic number", it is an indication that usability testing can be successfully carried out with a small number and a modest budget. It is open to debate but this thesis argues that producing a commercial, medium-large software project requires more than three test users, since end-users can be so varied in their needs.

---

[3]This work was carried out jointly with Cécile Rigny, a PhD student at Middlesex University, interested in cognitive user modelling.

The main objective for acceptable industrial practice is to produce commercially viable systems. Although doing research also involves producing well-designed systems, the main purpose is not to produce systems for commercial purposes, but rather to gain insight into a problem the systems are meant to address, as well as to use these systems to test new ideas. Therefore, doing research takes a different form in terms of the methodology employed in designing and testing systems, compared to industrial practice.

This thesis will be designing and reporting a series of laboratory experiments to test the hypotheses. These laboratory experiments are pilot experiments in which this thesis typically studies how to improve the interface as part of an iterative process. Since the aim is to learn which detailed aspects of the interface are good and bad, and how the design can be improved, the experiments will be aimed only at *formative evaluation* instead of *summative evaluation*[4] (Nielsen, 1993). Therefore, small sample sizes (about 3-6) to test the hypotheses would be used. Small numbers of users are more cost-effective since common/frequent problems but not infrequent or minor ones are encountered first. (The problem is — and remains for any methodology — how to find the infrequent disasters!) With 3-6 people, this thesis hopes to get qualitative results and impressions. Methods used in evaluation as reported in this thesis include survey, questionnaire and interview.

2.4.3 A note about terminology used in this thesis

To avoid confusion, the distinction between the users of a hypertext authoring tool and users of a hypertext is made. In subsequent chapters, the term "author" or "designer" or "user" refers to the user of a hypertext authoring tool and "end-user" refers to the person who uses the hypertext.

"Authoring tools" and "hypertext systems" are used interchangeably in this thesis, and they refer to a set of software tools used to create a hypertext.

2.5 Conclusion

To summarise, this thesis began by looking at the different definitions of hypertext that have evolved over the years. These definitions are largely dependent on the different perspectives taken by the authors, what they think hypertext could be used for and the then available technology that supported its implementation. Because there is still no general consensus on what hypertext constitutes, this chapter listed several factors that can be used to classify hypertext. The benefits of hypertext are discussed from two perspectives: the positive impact it has on our lives in general; and the operational advantages it has over traditional linear text. This thesis examined problems associated with hypertext. One of

[4]In contrast, summative evaluation aims at assessing the overall quality of an interface, for example, for use in deciding between two alternatives or as part of competitive analysis to learn how good the competition really is. A typical method to use for summative evaluation is a measurement test, providing statistical implications.

them is the LIH problem. This thesis then discussed the motivation behind the investigation, justifying the need to address the LIH problem.

This thesis reviewed and analysed the research efforts done so far to address the LIH problem. It argued that incorrect assumptions were made and therefore, the LIH problem still remains unsolved. The LIH problem was re-examined from fresh angles observed from within the end-user's mind, within hypertext, and on the web. It proposed taking multi-disciplinary approaches to address this problem, and set down the reasons for taking this approach. This thesis then mapped out the research methodology in user-centred designs, and experimental strategies in laboratory experiments, to address the LIH problem.

Section II takes each of the approaches proposed in turn and discusses how the LIH problem can be addressed. Chapter 3 investigates good design principles and guidelines for hypertext, chapter 4 describes a systematic, engineering task-based approach to understand end-users, chapter 5 explores important hypertext structural issues and chapter 6 identifies good features in authoring tools.

Section II

Multi-disciplinary approaches to address the “lost in hyperspace” problem

Chapter 3

Need for good design principles and guidelines

Chapter objectives:

This chapter emphasises that designers should do things well from the start, so that remedial work need not be done to correct the deficiencies within hypertexts because they are (or were) poorly designed and built. Its objectives are:

- to define what design principles and guidelines are
- to describe a framework to define design principles and guidelines for hypertext
- to describe an experimental framework to evaluate design principles and guidelines for hypertext

3.1 Introduction

Chapter 2 examined the LIH problem and set out reasons for the proactive, multi-disciplinary approaches and methodology taken in this thesis to address it. To begin, this chapter is concerned with the need for good design principles and guidelines for hypertext authoring. This chapter proposes a framework (see figure 3.1), a systematic approach to define principles and guidelines for hypertexts, an area which has not been investigated to any great extent (Glushko, 1992). Earlier attempts to define design principles and guidelines for hypertexts (for example, Hardman and Sharratt, 1990; Shneiderman and Kearsley, 1989; *etc.*) concentrated on a list of do's and don'ts of good hypertext authoring, but not on how they were derived. This chapter details how design principles and guidelines for hypertexts are formulated. Briefly, to draw up a list of relevant design principles and guidelines for hypertext, this thesis began with an investigation of general design principles and guidelines for interactive systems. It then examined design principles for hypertext authoring, by assessing their relevance from a cognitive perspective, since the interaction between end-users and hypertext is essentially *cognitive* (Preece, Benyon, Davies, Keller and Rogers, 1993). Based on the design principles proposed, a prototype hypertext was built. Assimilating findings from usability research, the design guidelines for hypertext were categorised into different areas of testing. The design guidelines were drawn up in the form of a user questionnaire to test the design principles that had been implemented in the prototype hypertext. Preliminary investigations were carried out to test the design principles proposed. The findings were fedback to refine the design principles and guidelines.

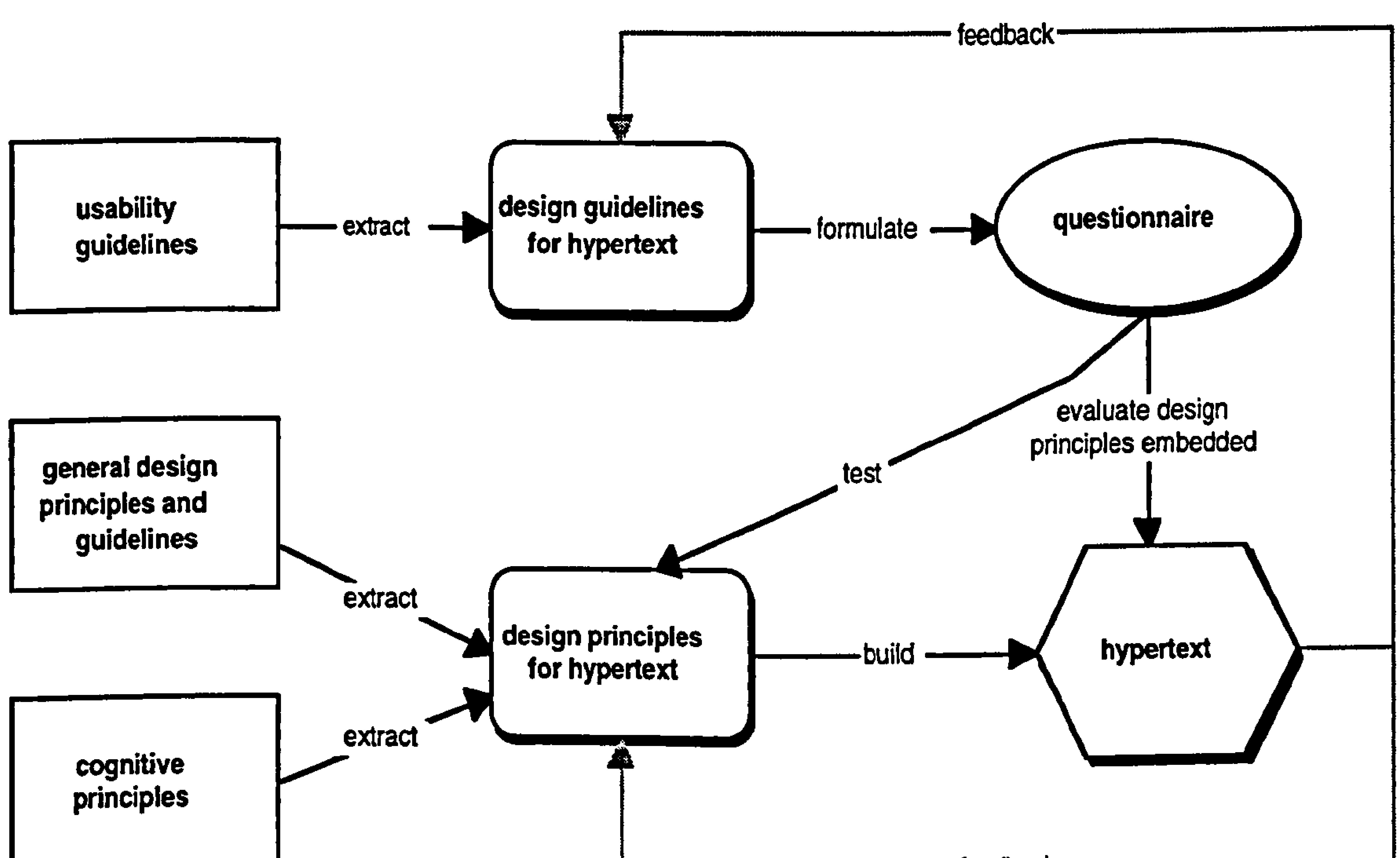


Figure 3.1. Framework for formulating hypertext design principles and guidelines

3.2 General design principles for interactive systems

The need for good design principles and guidelines for hypertexts has spurred the author to do a survey of other research efforts done in this area for general interactive systems, and in particular hypertexts. This survey achieves a two-fold objective: (i) to clarify confusing concepts such as principles, guidelines and methodologies; and (ii) to re-examine general design principles and guidelines, and extract relevant ones for hypertext authoring.

It is a well-accepted HCI fact that good design should take into account characteristics of the intended end-users and the work they do. If designers hope to design good interactive systems, they need answers to these questions: (i) How can knowledge of technological possibilities and user needs be duly considered and synthesised into an appropriate design? (ii) What process should be used? (iii) How can it be done most effectively? The answers to these questions can be summarised as either principles or rules, guidelines or checklists, or methodologies.

According to Baecker, Grudin, Buxton and Greenberg (1995a), *design principles* refer to collections of statements that advise the designer on how to proceed. Much research had been carried out to draw up comprehensive lists of general principles for the design of interactive systems (for example, Hansen, 1971; Foley and Wallace, 1974; Norman, 1983; Gaines and Shaw, 1984; Rubinstein and Hersh, 1984; Heckel, 1991; Shneiderman, 1992; *etc.*). Extracts of design principles taken from some of these well-referenced sources are:

- Rubinstein and Hersh (1984) presents 93 design principles including:
    - Separate design from implementation.
    - Develop an explicit user mental model.
    - Minimise conceptual load.
    - Be consistent in the use of the language.
    - Don't blame the end-user.
    - Make it easy to correct mistakes.
    - Select end-users who are representative of the target population.
    - Use representative tasks.

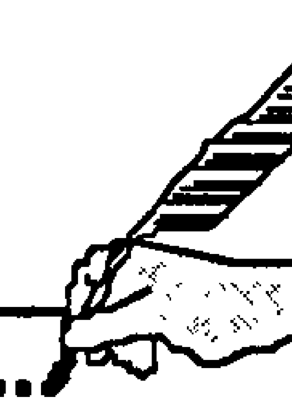
- Heckel (1991) structures around 30 design elements, for example:
    - Know your subject.
    - Maintain the end-user's interest.
    - Communicate visually.
    - Structure the end-user's interface.
    - Serve the novice and the experienced end-user.
    - Build a model in the end-user's mind.
    - Make the design simple, but not too simple.
    - Involve the end-user.

- Shneiderman (1992) formulates 8 rules for dialogue design:
    - Strive for consistency.
    - Enable frequent end-users to use shortcuts.
    - Offer informative feedback.
    - Design dialogues to yield closure.
    - Offer simple error handling.
    - Permit easy reversal of actions.
    - Support internal locus of control.
    - Reduce short-term memory load.

These sources attempt to create universally applicable platform-independent principles that provide sound advice, but lack the contextual information to assess how they can be applied to real design problems in practice (Henninger, Hayes and Reith, 1995; Smith and Mosier, 1986). One criticism against design principles is that they risk over-simplifying the design problem. Another criticism is that they are usually vague, sometimes contradictory and often formulated at an inappropriate level of specificity. For example, because opinions differ as to what "be consistent" means, advice such as "be consistent" with the use of language is not always helpful. For design principles to be useful, they must be translated into something more concrete before they can be applied to a particular system because they are almost completely open to interpretation (Lindgaard, 1994). Kellogg (1988) claim that the field was too immature to offer accurate recommendations for design. Thimbleby (1990) argues that the validity of principles are questionable, having been derived from small studies in artificial settings. Henninger, Hayes and Reith (1995) reason that the combination of diverse end-user backgrounds, diverse application domains, and proliferating technology solutions make it difficult to have a single, coherent, concise and universally applicable theory as a basis for design principles.

Although principles can assist designers as they design, they do not offer help on how to structure the design process. Therefore, design principles must be procedurised. For instance, to apply the principle "know the end-user", a number of methods recommended are: talk with end-users; observe end-users working; learn about the work organisation; perform task analysis; conduct surveys; *etc.* In addition to design principles, formalised procedures known also as *methodologies* (Baecker *et al.*, 1995a), are constructed to guide designers and help them structure the process of effective design.

It is open to debate, but this thesis refers to *design guidelines* as collections of tests that can be applied to an interface to determine if it is satisfactory, as defined by Baecker *et al.* (1995a). Figure 3.2 summarises the definitions of design principles and guidelines as they are applied in this thesis.



Definitions ......

Design principles refer to collections of statements that advise the designer how to proceed.

Design guidelines refer to collections of tests that can be applied to an interface to determine if it is satisfactory.

Figure 3.2. Definitions of design principles and guidelines

In contrast with design principles, design guidelines are usually more specific and in some cases quantitative, and formulated in a way to make them *testable*. For example, the guideline "provide an average response time of less than one second" is more specific and testable compared to a general principle "reduce response time". Designers or evaluators can evaluate whether or not an interface is desirable based on whether it meets this guideline. Design guidelines are used primarily in system and interface evaluation (Baecker *et al.*, 1995a). Extensive sets of guidelines for interactive systems are found from these sources: Smith and Mosier (1986) with 944 guidelines; Brown (1988) with 302 guidelines; Mayhew (1992) with 288 guidelines, *etc.*

For hypertext design guidelines to be useful, they need to be specific and therefore testable. This implies that for different categories of hypertext systems, there are specific sets of design guidelines that should be drawn up for these systems. The next section presents general design principles for hypertexts and draws upon findings from cognitive psychology to ensure that the human element central to the design of hypertexts is carefully considered. §3.4 illustrates how these design principles are formulated into design guidelines, which in turn contribute to the design of a questionnaire for evaluating hypertexts.

3.3 Finding and understanding relevant design principles for hypertext

Because hypertext is a relatively new design field, there are few detailed published case studies or design principles and guidelines that designers can readily use. According to Glushko (1992), this has been one of the several contributing factors that have led to the failures of many hypertexts. One of the major HCI issues in hypertext authoring is how to present information to the reader in an easily comprehensible way. Writers of hypertext need to develop new ways of giving the reader cues to cope with the freedom of choosing any sequence of reading. When Brown (1990) conducted an assessment on the quality of hypertexts. It was found that the student authors did not pay attention to the structure and presentation style of the hypertexts, resulting in poor design of the visual appearance of the material, overuse of technology with lots of clever effects but no real attempt to re-design for the new medium. It is generally agreed that designing hypertext is not an easy task. Hypertext authors need to make initial decisions on (Hardman and Sharratt, 1990): (i) subject matter; (ii) readership; and (iii) actual design which includes overall structure of hypertext, structure of the subject matter, and layout of materials.

Just as early programmers lacked experience, so do hypertext authors. It is, therefore, not surprising that some of the problems in hypertexts may be due to poor designs caused by bad authorship. To overcome this problem, hypertext designers need help (Thimbleby, 1997). Landauer (1995) explains that unless designers provide useful aids to assist end-user navigation, hypertext end-users will get "lost" easily owing to the inherent, non-linear nature of hypertext. Another failure of designers quoted by Addison and Dudman (1995) is the designers' inability to structure and represent information, resulting in end-users having difficulty in comprehending the information access routes provided in the system to get to relevant information. They cited Dorling-Kindersley's multimedia CD-ROM "The Ultimate Human Body" (1994), as a representative multimedia example of such design flaws.

Perhaps designers need to re-examine the way that hypertexts are designed and built by ensuring that good design principles and guidelines are incorporated into the building of hypertexts in the first place (for example, Nielsen 1995b; Signore 1995; *etc.*). If so, then what are these principles and guidelines? Gould, Boies and Lewis (1991) contend that the issues involved in creating an effective, useful and usable system should go far beyond the superficial treatment of interfaces as a facade to functionality.

3.3.1 Cognitive principles

Before listing the design principles applicable to hypertext authoring, this thesis will turn the attention briefly to cognitive principles to understand the way people act and react in their environment, so that user-centred design principles are considered. According to theories in cognitive psychology as described by Preece *et al.* (1993), end-users' interactions with hypertexts can be categorised into these key areas as shown in figure 3.3.

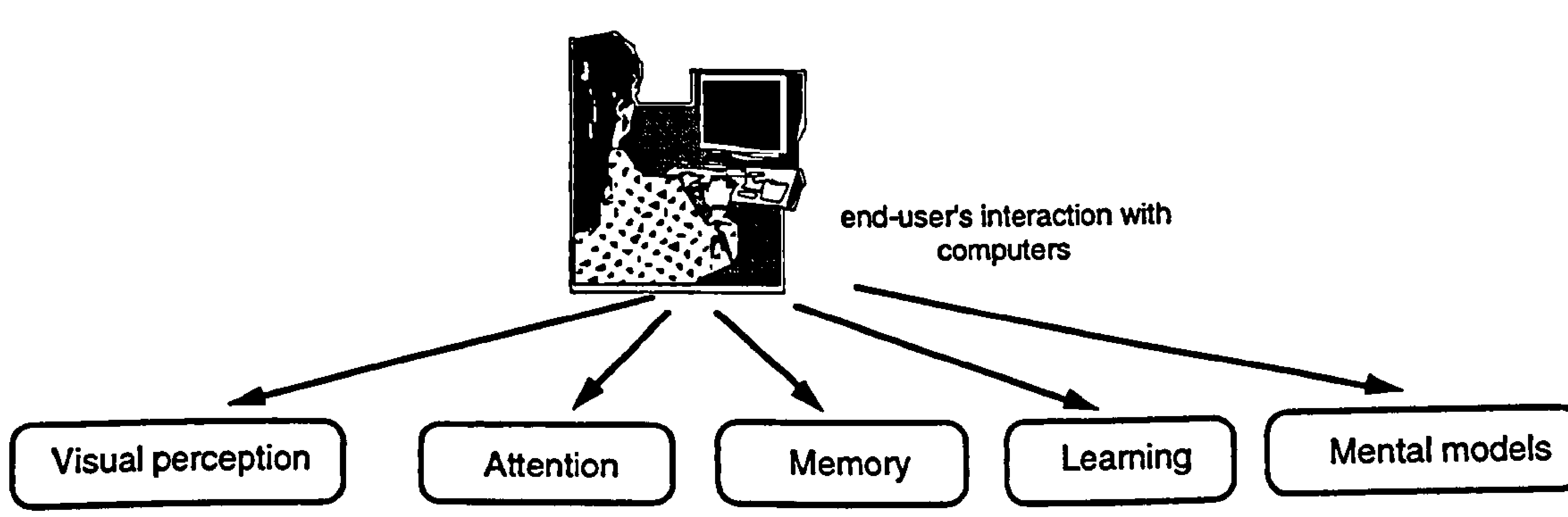


Figure 3.3. Cognitive psychology and end-users' interactions with hypertexts

C1. *Visual perception.* The process of seeing is an active process. During this process, sensory input and information are organised. The main organisation is the separation of the object of interest and focus from the background, so that it can be distinguished from the surrounding background. It is important to ensure that the information (text, graphics, animation, video or combinations of these) displayed on the screen is legible, distinguishable, comprehensible, uncluttered and meaningfully structured.

C2. *Attention.* Because cognitive processes limit the amount of information end-users can attend at any one time, a number of techniques need to be employed to alert and direct end-users' attention. Examples include: presenting information in a logical and meaningful structure to help end-users find relevant information; using various visual markers (flashing lights, underlining, bold) and auditory cues (such as beeps) to get end-users' attention; and partitioning of a screen into discrete or overlapping sections or windows associated with specific information.

C3. *Memory.* Interfaces should be designed to make minimal demands on our memory. Therefore, when using names and icons, they should be meaningful and easily distinguishable from each other in a set. The menu-based system, for example, has been designed to exploit an established finding in memory research, that end-users can recognise material from a display far more easily than they can recall without looking at it.

C4. *Learning.* Learning to use a computer is a complex process. To make learning easier, active methods of teaching which aid the learning processes should be explored.

C5. *Mental models.* When designing a system, designers should ensure that they communicate the structure of the hypertext clearly to end-users. If the end-users' mental model does not match design model, it can easily lead to LIH (Mantei, 1982). Therefore, the design of hypertext should ensure that the information processing activity with hypertext is within the end-users' mental processes, should provide knowledge about what end-users can and cannot be expected to do, leading to more user-centred compatible interfaces.

3.3.2 Applying cognitive principles to design principles for hypertext

Let's return to design principles for hypertext authoring. Much of the work done so far in hypertext design concentrates only on the interface design issues. Design principles should go beyond just providing an attractive interface. According to Hardman and Sharratt (1990), the interface design principles applicable to hypertext authoring are to ensure consistency of presentation, and to maintain minimal mental overload for end-users to remember objects, actions and codes to navigate hypertexts. Other researchers in their independent research work (for example, Fillion and Boyle, 1991; McKnight, Dillon and Richardson, 1991; Allison and Hammond, 1989; Conklin, 1987; *etc.*) identified interface design principles to address the following areas involving navigation aids, screen and information display and on-line assistance. Inspired by their work and cognitive principles (as described in §3.3.1), this thesis suggests six general design principles essential for designing good hypertexts (see figure 3.4).

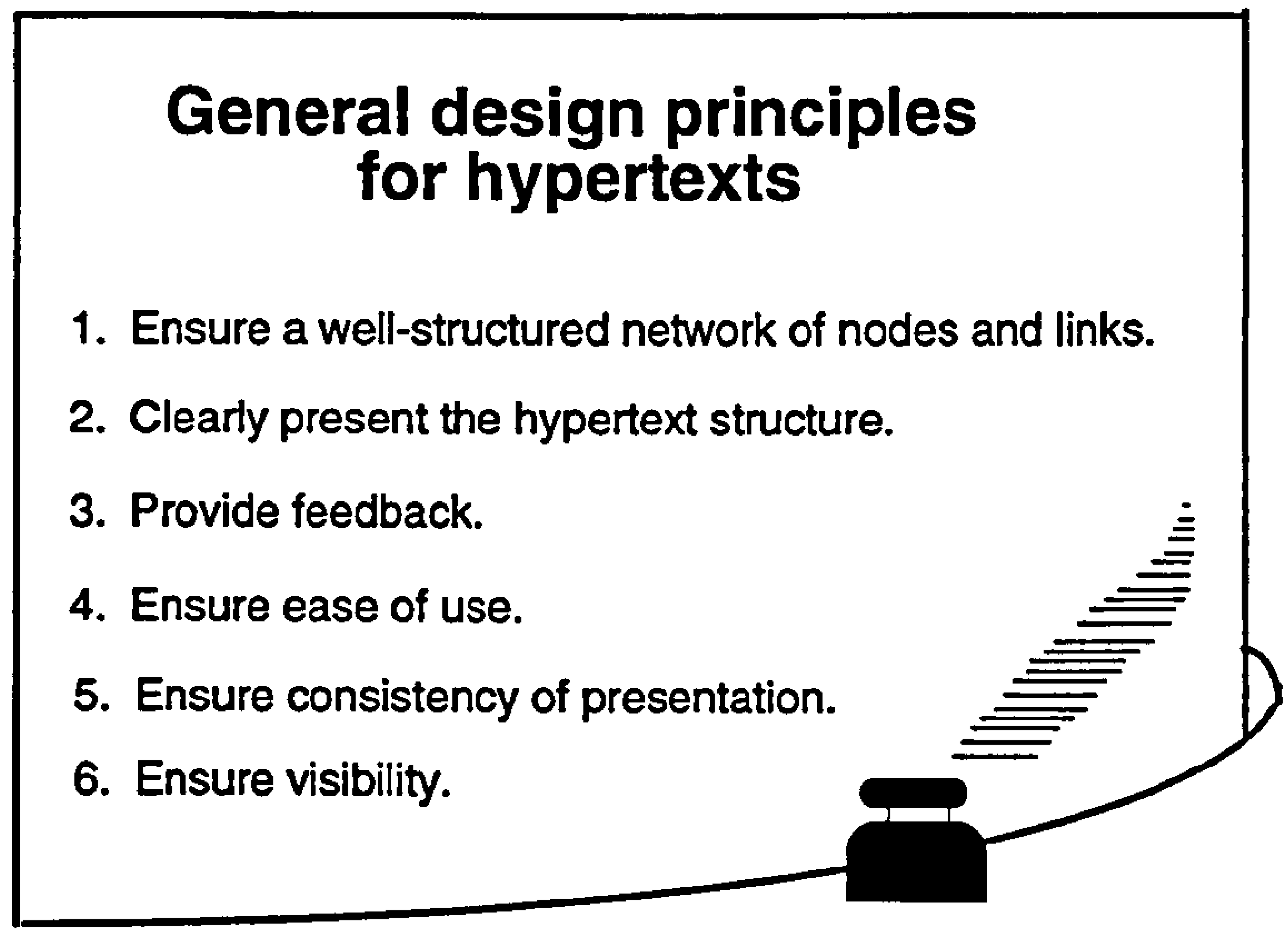


Figure 3.4. General design principles for hypertexts

P1. *Ensure a well-structured network of nodes and links.* Addison and Dudman (1995) argue that hypertext has been developed without real consideration of formal approaches, with significant emphasis being placed on seductive aspects of the interface rather than the development of the underlying structure. It is important to construct a good hypertext structure which supports designers and end-users alike, to prevent the LIH problem. Graph theory has been applied with success to improve interface design (for example, Thimbleby and Addison, 1995; Botafogo, Rivlin and Shneiderman, 1992; *etc.*). By selecting appropriate graph theory algorithms, it is possible to analyse the structure of hypertext, producing metrics that can support designers in making decisions to ensure a well-structured hypertext.

P2. *Clearly present the hypertext structure.* Not only should designers ensure that the hypertext is well-structured to greatly facilitate the information processing abilities of end-users, they should also present the internal structure and representation of nodes and links to the end-users in the simplest and most effective way. An experiment by Allinson and Hammond (1989) confirms that end-users' lack of overview information or wrong understanding of the hypertext structure make them more prone to losing their orientation. Thus, designers should try to communicate the design model of hypertext so that the mental models end-users form when using hypertext would match the design model (refer to C5). Maps of various kinds and functions have been used to achieve the purpose in many hypertexts (for example, Conklin, 1987; McKnight, Dillon and Richardson, 1991; Maurer, 1996; *etc.*). "Global maps" or "overview diagrams" have been introduced to give end-users an overview of the hypertext and the documents and links between them. "Local maps" or "document finders" are used to provide end-users with information on what documents are linked to current user-defined document. These maps aim to provide end-users with as much context information as possible so that end-users are able to answer questions like: Where am I?; Where do I go from here?; How much information is there?; How did I arrive here?, *etc.* Furnas (1986) proposes using fisheye maps to show details about matters at the focus of attention, but as the information moves away, less and less of the details should be shown. Conklin (1987) suggests using graphical document browsers to display the structure of the document and allow end-users to assess what is there. This is because much of the problem in navigation and in particular, disorientation is the lack of understanding of the document structure and the inability of end-users to assess the amount and size of information available.

P3. *Provide feedback.* End-users should be given full and continuous feedback about the results of actions (refer to C2). At any point in the hypertext, there should be sufficient information to orientate end-users such as the use of title, subtitle, page numbers, *etc.* Help should always be available to end-users at the general level, for example, displaying a help icon on the screen. Help should also be available specific to the end-users' current position, providing adaptive help if possible. Search facilities can also be provided to allow end-users to obtain feedback on specific information they might be interested in. Without feedback, end-user would inevitably get "lost".

P4. *Ensure ease of use.* Designers need to ensure that end-users find the hypertext easy to use (refer to C4). Fillion and Boyle (1991) suggest providing end-users with reference points, orientation cues and some notion of "direction" to aid end-users to learn about the hypertext. Table of contents, index and references have also been included in many hypertexts like the *Semistructured Intelligent Navigation System* (Boyle and Snell, 1990), *Book Emulator* (Benest, 1990) and ACM's *Hypertext-on-Hypertext* (ACM, 1989). Metaphors of all kinds have been introduced in hypertexts such as the travel metaphor in the *Learning Support System* (Allinson and Hammond, 1989) and the book metaphor *Book Emulator* (Benest, 1990). To address the problem of returning to known locations, Monk (1990) suggests a "personal browser" to capture information of previously visited node. HyperCard provides a built-in Recent Command, a backtracking facility, to provide end-users with a list of miniature screen dumps of the 42 most recently visited nodes. In the design of *Hypergate,* Bernestein (1988) describes the introduction of personalised features like bookmarks and margin notes for end-users. Buttons such as "go to next page", "return to previous page", "go to home page", *etc.* are useful in providing good navigational support.

P5. *Ensure consistency of presentation.* Hypertext should have a consistent "look and feel" because information presented in a consistent manner will assist end-users to find relevant information more easily (refer to C1). This may in turn draw end-users' attention and interest in using the system (refer to C2). A consistency in the presentation of operations will result in a coherent, consistent system image. This also makes learning to use the system much easier (refer to C4), since there is minimal mental overload for end-users (refer to C3). There should be consistency in the use of terminology, wording and format.

P6. *Ensure visibility.* By looking, end-users should be able to tell the state of the device and the alternatives for action (refer to C1). The display of information should be kept short and simple with only the necessary information displayed. Information should also be grouped to reflect relationships and ordered to capture the structure of the hypertext. Critical information should be highlighted though this should be used with discretion and only for a small proportion of information on the screen. As the links within hypertexts can be of different types, these should be made apparent to end-users before they are actioned. Visual coding for different types of information in the form of graphics, colour, brightness, flashing or combinations of them can be used. Care should be taken, however, not to use too many codes to prevent overloading the end-user (refer to C3). Graphics can be used to illustrate ideas in a more interesting and appealing manner compared to just plain text.

3.3.3 Design of a prototype hypertext

Knowing what some good hypertext design principles are, a prototype hypertext on *Basic Computer Anatomy* using HyperCard version 2.01 was built. The contents of the prototype were adapted from Beekman (1995). For the purpose of illustration, the hypertext only implemented some of the suggestions made in P1 – P6. The design of the hypertext was intended to be as simple as possible, since the objective was to examine the design issues involved in creating an effective, useful and usable hypertext, instead of providing superficial treatment of interfaces as a facade to functionality. The hypertext prototype was made up of 47 cards. Figure 3.5 shows the first four cards of the hypertext, which are accessed in sequence by pressing the "next" button, indicated by the "->" button.

To ensure a well-structured network of nodes and links (refer to P1), the hypertext was structured as a sequence of cards hierarchically linked to each other. The choice of this kind of structure will be explained in greater detail in chapter 5. To present clearly the hypertext structure (refer to P2), a "table of contents" showing the overall structure to ensure ease of use (refer to P4) was implemented.

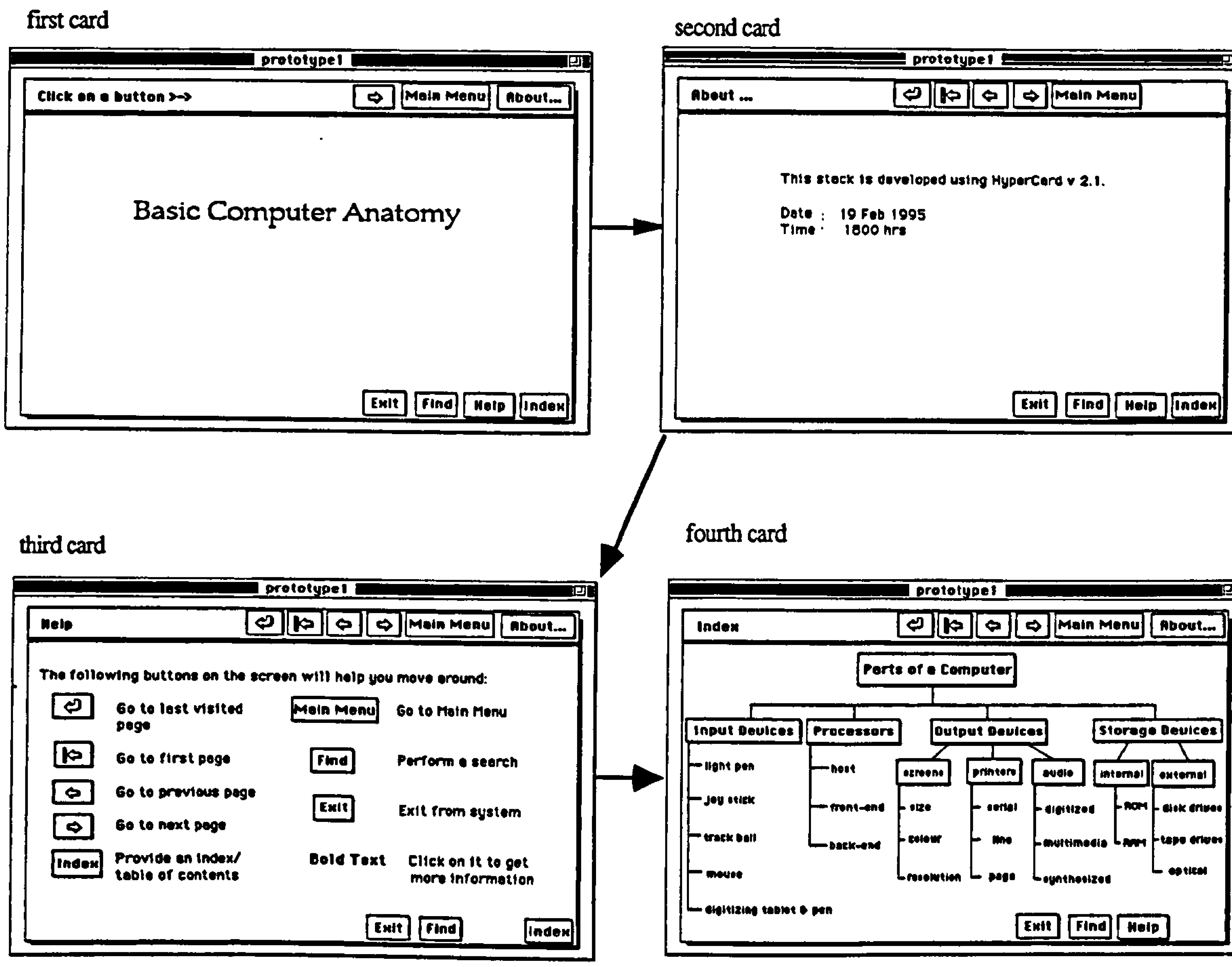


Figure 3.5. First four cards of the prototype

Figure 3.6 shows a sample card of the prototype. For consistency of presentation (refer to P5), every card has the same "look and feel". The top of card shows the title of the card as well as gives end-users feedback to help them orientate themselves (refer to P3). In addition, navigational buttons are also implemented to ensure ease of use and to assist in end-user navigation (refer to P4). At the bottom of the card is a "Help" button (refer to P3) to instruct end-users on how to navigate round this prototype. A "Find" button provides end-users with the facility to locate information within a card more easily (refer to P3). Users move from one card to another using "hotspots", which are easily identified in bold text (refer to P6). Appendix F1 shows the complete stack of cards of the prototype.

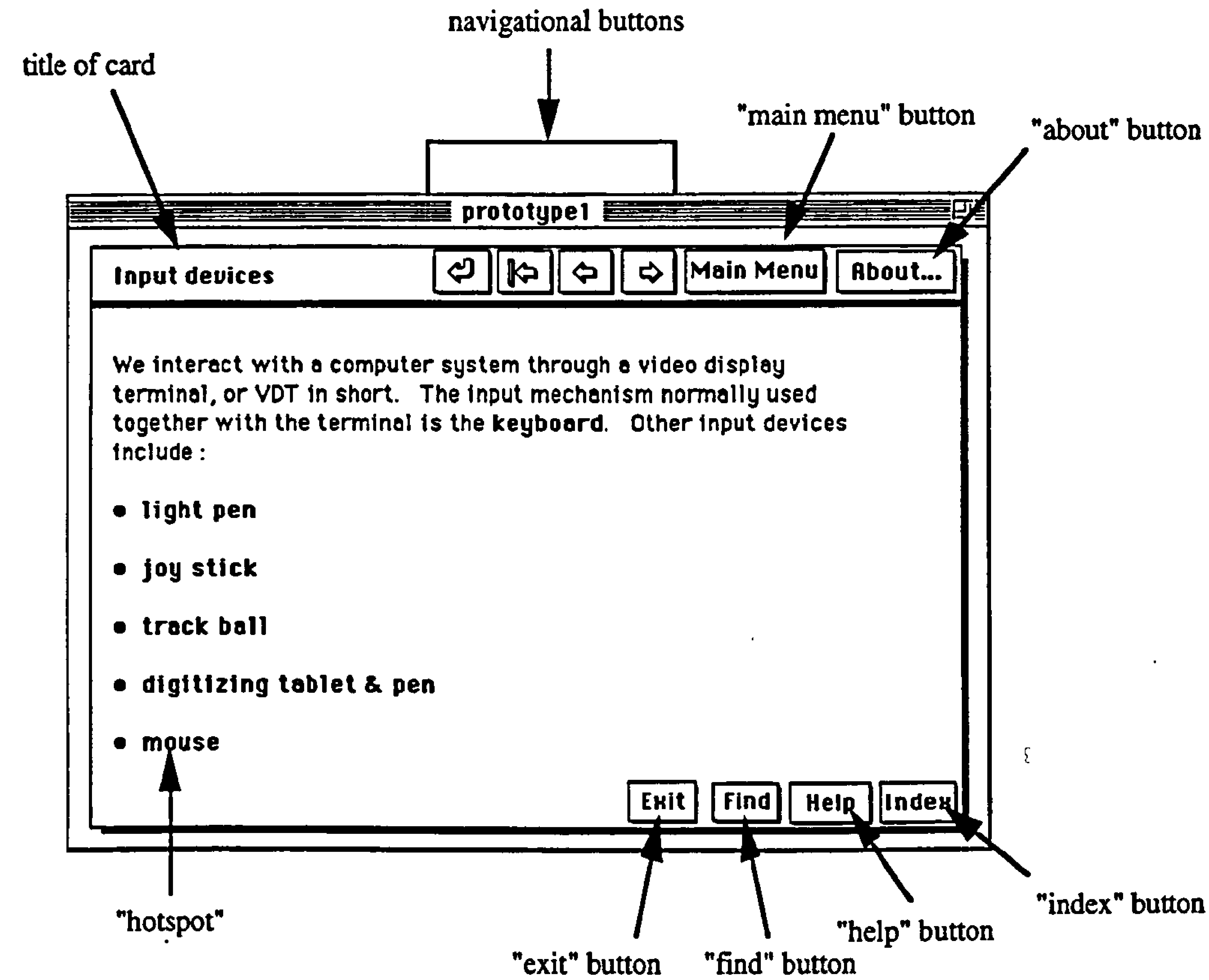


Figure 3.6. Screen shot of a sample card of the prototype

3.4 Formulating design guidelines for hypertext usability

The six design principles for hypertexts P1 – P6 outlined above are by no means exhaustive. As observed, these principles are general, and implementing them is open to designers' interpretation. Having built a hypertext prototype incorporating some ideas suggested in principles P1 – P6, the next thing to do is to test if the prototype is "good". In other words, a set of design guidelines is needed to evaluate when a "good" hypertext is "good enough". In general, a computer system is said to be good, if it is useful and usable (Landauer, 1995; Newman and Lamming, 1995; *etc.*). A system is "useful" if it covers adequately the range of tasks it intends to support. A system is "usable" in relation to end-user's performance in the specific tasks supported by the computer system, and the end-user's attitude towards the system.

Knowing where in the hypertext end-users run into problems or the kinds of problems end-users are likely to encounter, gives a good indication of what needs to be changed, and how to make it better in the next version. Failure to do so will render designers incapable of substantiating claims of improvements in a hypertext (Lindgaard, 1994). Since design guidelines are primarily employed in system and interface evaluation (Henniger, Haynes and Reith, 1995), this thesis turns briefly to established usability dimensions traditionally proposed to measure usability in interactive systems (for example, Shackel, 1991; Nielsen, 1993; *etc.*), before formulating design guidelines to test the design principles P1 – P6.

3.4.1 Usability dimensions

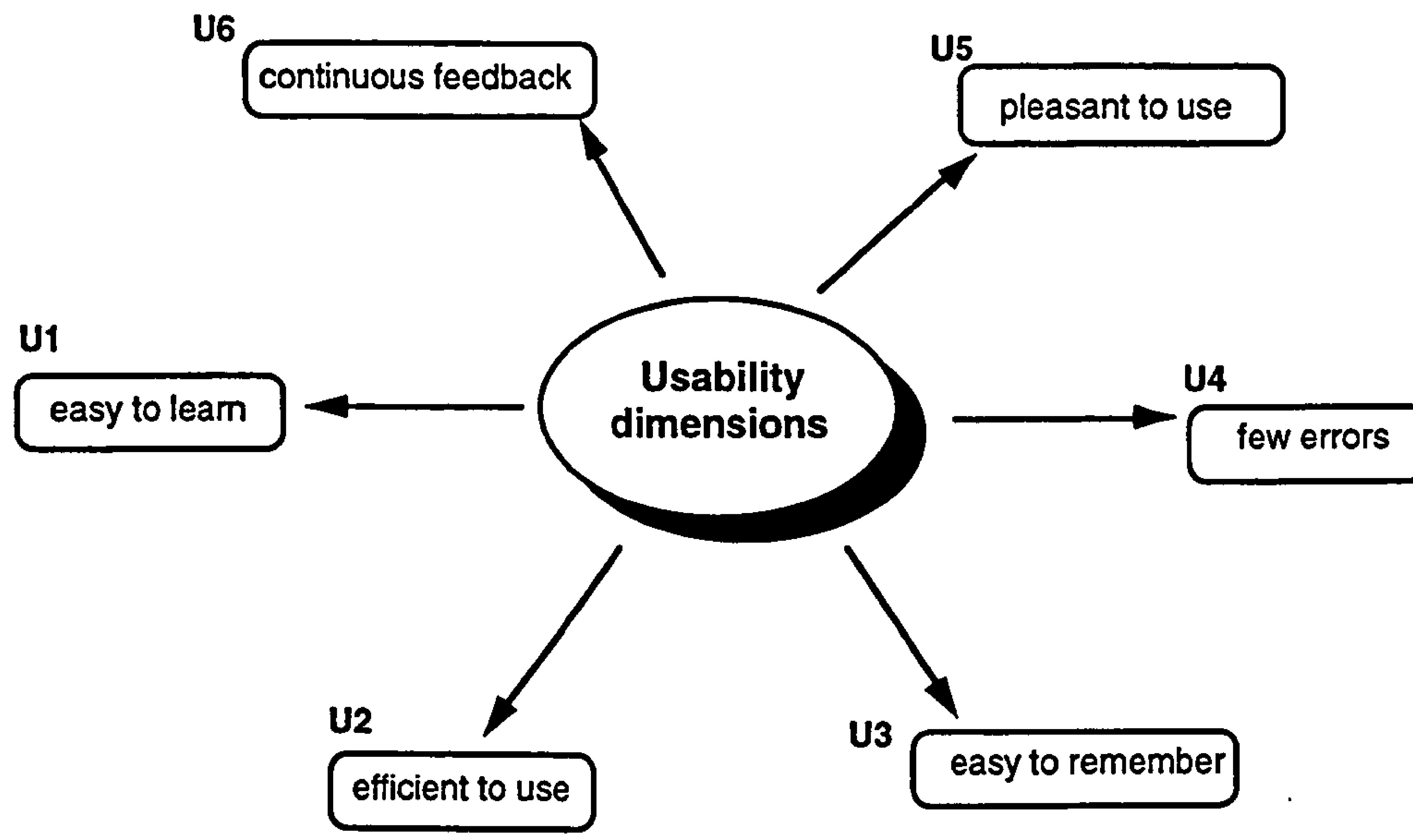


Figure 3.7. Usability dimensions

Figure 3.7 shows the six well-accepted usability dimensions from Shackel (1991) and Nielsen (1993) selected to guide the formulation of design guidelines for evaluating hypertexts:

U1. *Easy to learn.* This measures how quickly end-users can get some work done with the system. It may be measured in terms of the number of trials needed to complete particular tasks with/without training. It may also be measured in terms of the number of times end-users look up manual, consult help, ask questions of the observer, or go back to the opening screen while attempting to complete a certain task or tasks.

U2. *Efficient to use.* Once the end-user has learned the system, a high level of productivity is possible. It can refer to levels of end-user performance, measured in terms of speed and/or accuracy, in terms of proportion of task(s), proportion of end-users, or probability of completion of a given task.

U3. *Easy to remember.* End-users who have used the system before are able to return to using the system after not using it for some period, without having to learn everything all over.

U4. *Few errors.* End-users should not make many errors during the use of the system. Or, if they do, they should be able to recover easily.

U5. *Pleasant to use.* End-users are satisfied by using the system. They are measured in interviews or surveys, which enable response categories to be quantified.

U6. *Continuous feedback.* End-users should be provided with feedback on where they are, and how they can go to a different location from where they are presently.

For design dimensions to be usable, they should be translated into meaningful quantitative statements. Unfortunately, in practice this is not easy to accomplish because the notion of usability is intuitive and difficult to translate into quantifiable measurements. Many factors such as types of end-users and the tasks they engage in, could have significant impact on the types of usability dimensions used, and how they are measured (Dickerson and Hedman, 1993). Only a few have attempted to list down measurement criteria, for example, Tydesley (1988). Eason and Harker (1991) contend that even though usability metrics have been proposed and could be calculated, relatively few studies discussing or seeking to apply these metrics have been reported (Lindgaard, 1994; Babiker, Fujihara and Boyle, 1991). Perhaps this is because designers find it difficult to attach a number to the usability dimensions. The next section describes an attempt to formulate design guidelines by quantifying the usability dimensions U1 – U6 and applying them to hypertext, in the hope that the idea of using design principles and guidelines to guide and shape hypertext design would not remain a farce.

3.4.2 Design guidelines for hypertext

There are many possible ways to present guidelines to designers, for example, providing designers with a list of do's and don'ts in hypertext authoring. A relatively new area of research in recent years is to automate usability metrics, thus providing designers with quick results and analyses. Chapter 7 will discuss this in greater detail on how the prototype hypertext authoring tool implemented some usability metrics that can be automated upon request from the users of the tool. This thesis has also chosen to formulate design guidelines using a measurement instrument such as the questionnaire to measure end-users' subjective opinions. Questionnaires provide a rich source of data fairly quickly compared other popular survey methods such as interviews.

The formulation of the questionnaire (based on design principles) was greatly inspired by the development of a measurement tool called the Questionnaire for User Interface Satisfaction (QUIS) by Chin, Diehl and Norman (1988). QUIS measures end-users' subjective ratings of the interface of an interactive system. According to Chin, Diehl and Norman (1988), even though several questionnaires have been developed to assess end-users' perceptions of interactive systems, their weaknesses ranged from a lack of validation (Callager, 1974) to low reliability (Larcker and Lessig, 1980). Chin, Diehl and Norman (1988) claimed QUIS is reliable. Since the primary focus of the investigation is to come up with a list of hypertext principles and guidelines quickly for the development of a prototype hypertext authoring tool and the evaluation of hypertext (see chapter 7), this thesis is not concerned with validating the questionnaire. The design of the questionnaire on design guidelines was modelled closely after QUIS, adapted for hypertext authoring, because of its reliability as claimed by the authors.

To select the relevant areas to measure usability, this thesis turns to Lingaard's classification of typical usability defects for interactive systems which include (Lingaard, 1994): navigation; screen design and layout; terminology; feedback; consistency; modality; redundancies; end-user control and match with end-user tasks. Inspired by Lingaard's classification of usability defects in interactive systems, this thesis integrates usability dimensions U1 – U6 into QUIS by formulating the general design categories for evaluating hypertext into seven areas G1 – G7 (see figure 3.8).

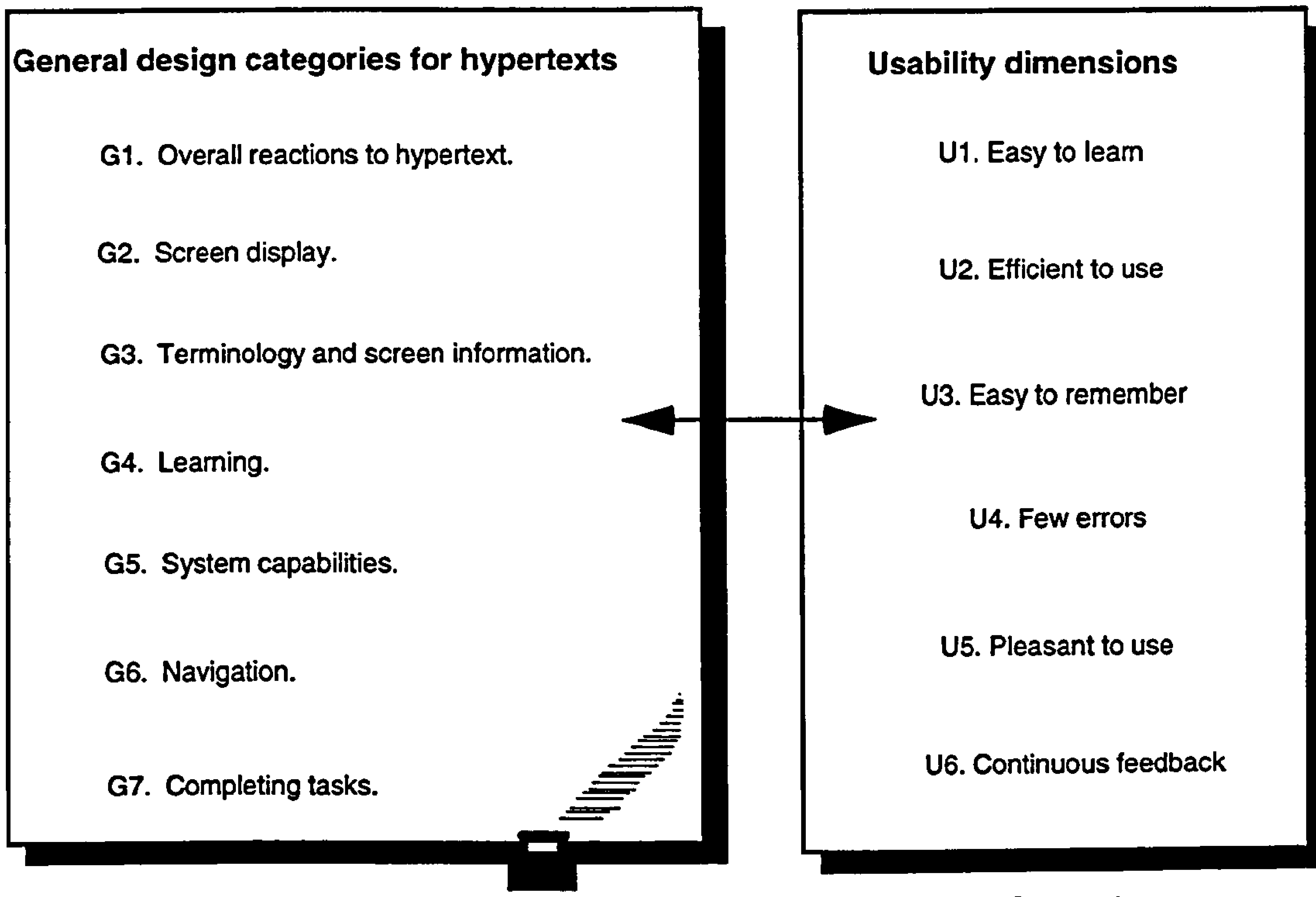


Figure 3.8. Design categories for hypertexts and usability dimensions

G1. *Overall reactions to hypertext.* This area evaluates end-users' overall perception of the performance of hypertext in terms of satisfaction (refer to U4), completion of tasks (refer to U2) and appeal (refer to U5).

G2. *Screen display.* This area measures how clearly information is organised and displayed on the screen (refer to U3, U4).

G3. *Terminology and system information.* This area examines whether hypertext is consistent in the use of terminology, word and format (refer to U3). It also asks if the system provides feedback, and whether error messages are useful (refer to U6).

G4. *Learning.* This area investigates the ease of use of the hypertext (refer to U1).

G5. *System capabilities and user control.* This area examines hypertext's response time, reliability and recovery process (refer to U4). It also asks whether hypertext design is simple or complex, for experienced or inexperienced end-users (refer to U1).

G6. *Navigation.* This area asks questions on how clearly are the navigational elements such as maps, table of contents, etc. displayed (refer to U1 and U3). It also investigates whether the end-user is "lost", and the reasons why (refer to U6).

G7. *Completing tasks.* This area examines the extent of usefulness of facilities in hypertext in helping end-users complete their tasks in browsing, seeking references, information search and revision (refer to U2).

To test the design principles P1 – P6, questions based on the seven areas of design categories G1 – G7 were formulated. According to Nunnally (1978), the larger the number of items and scaling steps, the higher the reliability of the questionnaire. Therefore, the first version of the questionnaire began with a large number of questions, forty in total covering all the seven areas G1 – G7. For the actual questionnaire used in the experiments, please refer to appendix B1. Table 3.1 shows the questions and the corresponding design principle(s) tested by the questions. A total of 8 questions were asked to test design principle P1, followed by 9 for P2, 8 for P3, 13 for P4, 5 for P5, and 8 for P6. (The total does not add up to 40 since some questions are suitable to test more than one principle.) The questions to test the different design principles were not clustered together, but spread out in order to get more spontaneous feedback from the respondents.

For the purpose of the investigation, 34 out of 40 questions were closed questions since they are generally easier to analyse than open questions (Preece *et al.*, 1993). Responses obtained from closed questions can be easily converted into numerical values and a statistical analysis can be performed. Six open questions (Questions 26, 33, 35, 37, 38 and 40) were asked since they encourage "freer" answers from respondents, and hence provide a rich source of data, which may otherwise go undetected.

Generally, end-users prefer concrete adjectives for evaluations (Coleman, Williges and Wixon, 1985), therefore the questionnaire used a semantic differential scale, a popular form of attitude scale used in HCI research, to measure end-users' responses (Preece *et al.*, 1993). This scale has bi-polar adjectives at the end-points, and respondents rate the interface on a scale between these paired adjectives by putting a tick in the appropriate column (see figure 3.9).

	extremely	quite	slightly	neutral	slightly	quite	extremely	
poor								good

Figure 3.9. A semantic differential scale to measure end-users' responses

Though it is important to establish external, construct or predictive validity of the questionnaire, it is outside the scope of this thesis. This thesis proposes validation of the questionnaire for future investigation (see chapter 9).

Table 3.1. Questionnaire and corresponding principles tested

Question		Principle(s) tested
	Overall reactions to the system	
1	How would you rate the performance of the system?	P1
2	How would you rate the usability of the system?	P4
3	How would you rate the level of satisfaction when using the system?	P4
4	How would you rate the power of the system in helping you to complete the tasks?	P1
5	How would you rate the appeal of the system in encouraging you to use it?	P6
6	How flexible is the system in helping you to complete the task?	P4
	Screen display	
7	How easy are the characters on the computer screen to read?	P6
8	How does highlighting on the screen simplify the task?	P6
9	How would you rate the organisation of information on the screen?	P6
	Terminology and system information	
10	Is the use of terminology, word and format consistent throughout the system?	P5
11	Does the computer terminology used relate to the task you are asked to do?	P4
12	Is the positioning of the messages on the screen consistent?	P5
13	Do you find prompts for inputs clear?	P6
14	Does the system provide feedback about what it is doing?	P3
15	Are the error messages useful?	P3
	Learning	
16	Do you find it easy to learn how to use the system?	P4
17	How easy is it for you to explore new features by trial and error?	P4
18	Do you find it easy remembering names and use of commands?	P4
	System capabilities and user control	
19	Do you find the system responds fast enough?	P3
20	Do you find the system reliable?	P3
21	How easy does the system enable you to correct your mistakes?	P3
22	Has the system taken both experienced and inexperienced users' needs?	P4
	Navigation	
23	Are the links to other information within the system clearly displayed and highlighted?	P6
24	Are there sufficient information on screen to help you know where you are in relation to the structure (e.g. use of title, page numbers, etc.)?	P2
25	Does the system provide sufficient information to help assess the structure of the system (e.g. browsers, maps, table of contents, etc.)?	P2
26	Which 'map' do you think best represents the structure of the hypertext?	P1
27	Does the system provide good features/facilities to help you recall what you have browsed or the key points covered?	P4
28	Is it difficult to identify where you are in the system?	P2
29	Is it difficult to find a particular topic thought to exist?	P2
30	Is it difficult to return to a topic left previously?	P2
31	To what extent do you feel 'lost' when using the system?	P1
32	To what extent do you feel that being 'lost' is waste of time?	P4
	Completing tasks	
33	How long did you take to complete your tasks?	P1
34	To what extent are you familiar with the system after completing the tasks with respect to navigation tools, the way information is organised, and information covered?	P5
35	Which features/ tools of the system are useful in helping you complete your tasks in browsing, seeking references, information search and revision?	P2, P4
36	How confident did you feel when using these features/tools to complete the tasks?	P1, P2
37	Do you have the feeling you are overlooking crucial information when completing the tasks?	P3
38	List 3 features/tools you wish the systsm would provide.	P1-P6
39	How would you rate the effectiveness of the system in helping you complete the tasks?	P4
40	Other comments.	P1-P6

3.5 Experimental framework to evaluate hypertext design principles

Having postulated the design principles and guidelines for hypertext, the questions to ask are: are the principles sound? What effects do implementing or not implementing these principles have on end-users' satisfaction and end-users' perception of the effectiveness of the hypertext? Are the guidelines useful in testing the principles, and hence provide designers with reference and guidance in the design process? To answer these questions, two experiments were conducted which shall be described in experiments 3.1 and 3.2.

In these experiments, a usability goal is defined as "a means of studying and evaluating hypertext's functionality and interface design". According to ISO 9241-11, usability is defined as "the extent in which a product can be used by specified end-users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO, 1990). Therefore, this thesis has chosen to quantify "level of satisfaction expressed by end-users" and "effectiveness of system perceived by end-users" as usability goals to evaluate the functionality and interface design of hypertext. Because of the small samples used, only qualitative results and impressions would be reported.

3.5.1 Experiment 3.1: A sample hypertext

Objectives

Experiment 3.1 achieves the following objectives:

1. To examine whether end-users' ratings on the implementation of design principles in a hypertext have any influence on their satisfaction and perception of the effectiveness of it.

2. To investigate whether design principles are embedded in the design of a sample, commercially-produced hypertext, such as ACM's *Hypertext-on-hypertext*.

Figure 3.10 illustrates how experiment 3.1 is fitted into the framework for formulating and evaluating design principles and guidelines for hypertext described at the outset of this chapter. Experiment 3.1 provides an experimental framework to illustrate how feedback on the design principles suggested for hypertexts is obtained.

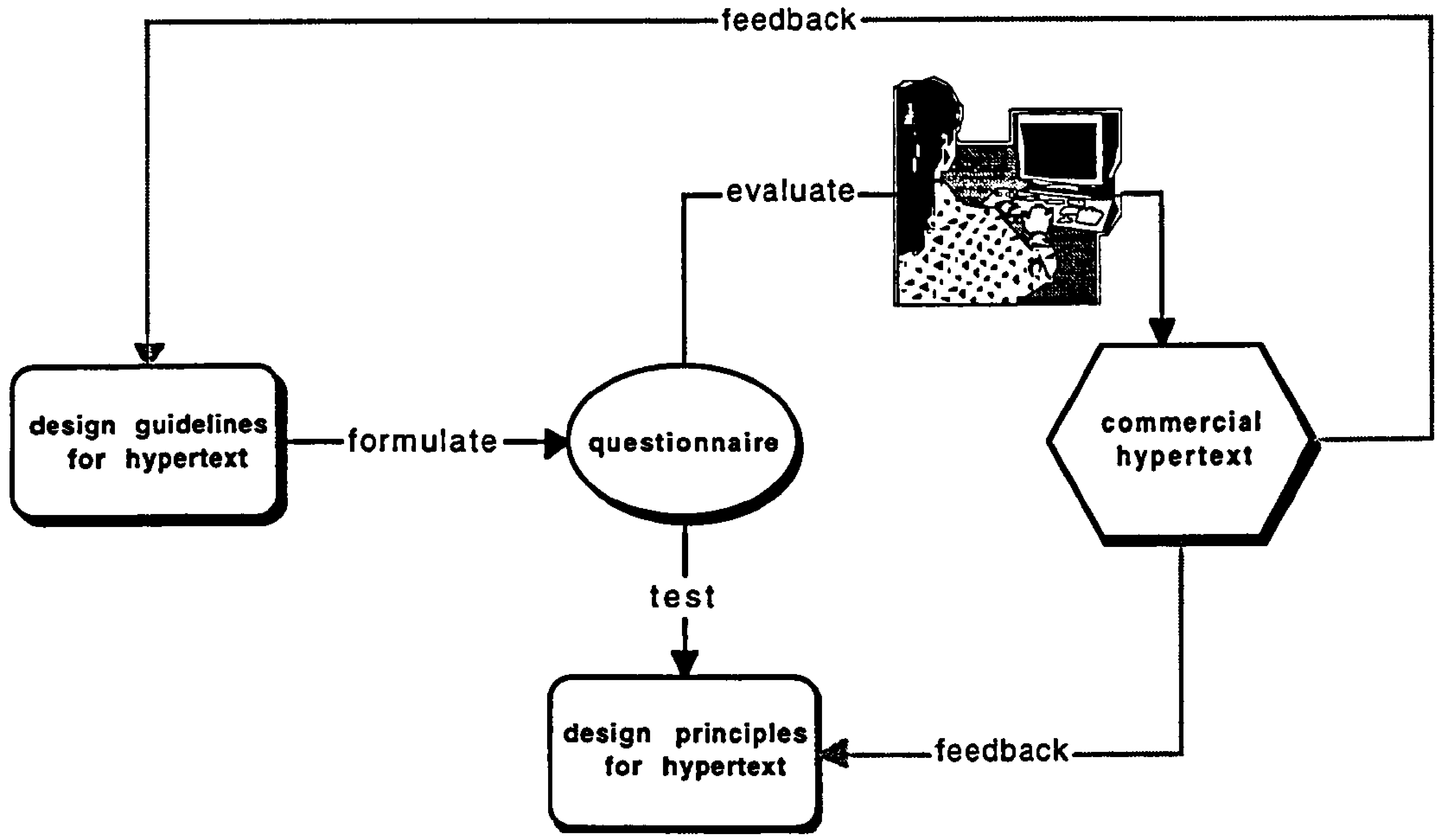


Figure 3.10. Framework for experimental illustration and analysis (experiment 3.1) of design principles and guidelines for hypertext

Protocol

Six subjects were selected to evaluate a hypertext called ACM's *Hypertext-on-Hypertext* (1987). The *Hypertext-on-hypertext* contains position papers presented at Hypertext'87 Conference held at Chapel Hill, California. It was chosen because none of the subjects were familiar with contents presented, hence this thesis made the assumption that subjects' completion of tasks were based upon successful navigation and retrieval of information, and not upon their prior knowledge. Table 3.2 shows the profiles of these six subjects.

The subjects were asked to use the ACM's *Hypertext-on-hypertext* by answering questions that involved tasks such as [1]browsing, seeking references, information search and revision (see appendix C1). After that, they were asked to complete a questionnaire (see appendix B1) commenting on how satisfied they were with the design and structure of the hypertext in helping them to complete the tasks successfully. "Satisfaction" refers to the "feeling of being pleased with the hypertext in helping to complete the task successfully". "Being pleased" is defined in terms of the subject's perceived ease of use, rate of errors, and time taken to perform the task successfully. The subjects were then interviewed after they had completed the questionnaire to ask for further comments and/or clarification regarding their responses on the structure of the questionnaire and the hypertext.

[1]In Chapter 4, we will discuss in greater detail what 'browsing' entails. In this chapter, browsing refers to navigating hypertext without any specific goal or purpose. Seeking references refers to "looking or trying to find something". Information search refers to "examining or looking carefully in order to find information". Revision refers to "bringing back to mind something or succeeding in remembering".

Table 3.2. Profiles of subjects

	Subjects' profile	Subjects' experience in hypertext
Subject 1	Researcher	yes
Subject 2	Researcher	yes
Subject 3	Researcher	no
Subject 4	Researcher	no
Subject 5	Undergraduate	no
Subject 6	Undergraduate	no

Objective 1: Discussion on results and analyses

For easier analysis and display of results, the semantic differential scale used in the questionnaire was translated into a 7-point scale (see figure 3.11). For example, number 1 represents "extremely poor" and number 7 represents "extremely good". A value "5 and above" is considered "good", implying that end-users are generally pleased with the hypertext and designers need not make any changes. A value "3 and below" is deemed "unsatisfactory" indicating end-users' dissatisfaction with the hypertext, and designers should make necessary changes to correct the deficiency. A mid-value of 4 is taken to be "neutral", and designers should find out more from end-users and make changes if required.

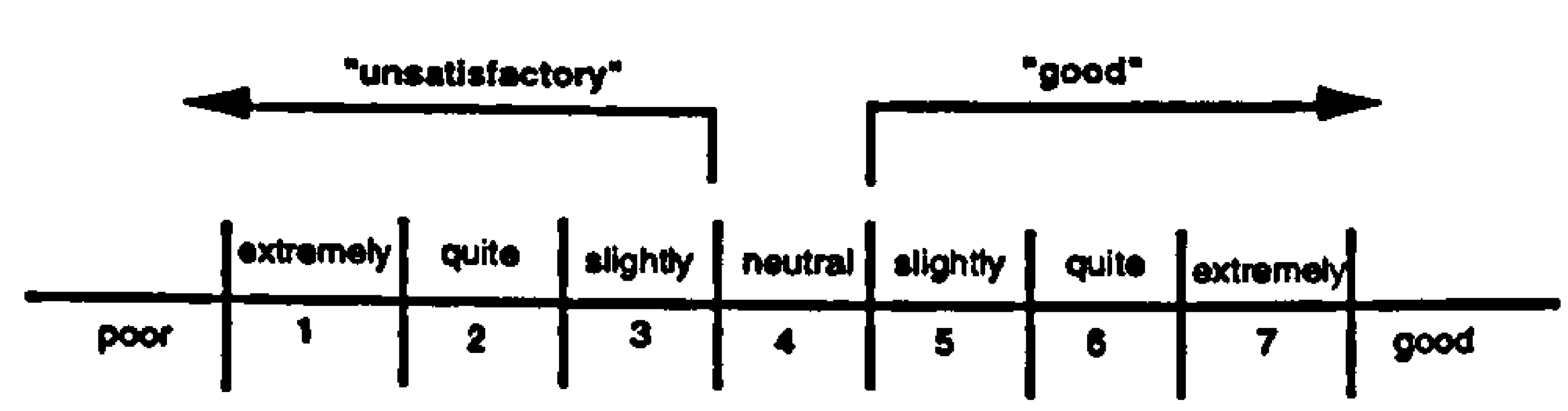


Figure 3.11. A 7-point scale to measure end-users' responses

Table 3.3 presents subjects' evaluation of the ACM's *Hypertext-on-hypertext* translated into the 7-point scale. Only subjects' feedback on the *closed* questions are reported in table 3.3. Because of the rich amount of data captured by the questionnaire, there can be many ways data can be analysed. However, it is not the objective of this thesis to perform a thorough analysis of the data, it will limit the analysis to that which relates directly to the two objectives set out to investigate in this experiment. This thesis will also analyse the data relating to issues such as LIH, navigation and structure in chapter 5.

Table 3.3. Tabulation of subjects' ratings of closed questions

Question	ACM's Hypertext-on-hypertext	S1	S2	S3	S4	S5	S6
1	Performance of the system.	4	6	5	1	4	4
2	Usability of the system.	5	6	6	2	6	3
3	Satisfaction when using the system.	4	5	5	2	5	4
4	Help in completing tasks.	5	6	6	2	3	5
5	Appeal of system.	3	5	4	1	4	2
6	Flexibility of system.	6	3	5	1	5	3
7	Ease of reading characters off screen.	6	3	6	2	6	6
8	Helpfulness with highlighting used.	5	6	6	7	-	6
9	Organisation of information on screen.	6	6	6	2	2	6
10	Consistency in terminology, word and format.	6	6	6	4	4	6
11	Match between terminology used and tasks to do.	6	5	6	4	4	6
12	Consistency in positioning of messages.	6	7	7	3	4	6
13	Prompts for input.	6	6	6	2	5	6
14	Feedback provided.	1	3	6	2	4	2
15	Error messages provided.	6	-	-	-	-	4
16	Ease in learning to use system.	6	7	6	3	3	3
17	Ease in learning to use new features.	7	6	5	4	6	6
18	Ease in remembering names and use of commands.	7	6	-	3	6	6
19	Response of system.	7	1	6	2	5	4
20	Reliability of of system.	-	6	7	4	7	4
21	Recovery of system	5	6	4	2	5	6
22	Adaptability of system to users' needs.	5	6	4	1	6	2
23	Link to other information.	3	3	6	2	4	6
24	Information to orientate users.	1	6	7	4	6	6
25	Information to assess structure of system.	1	6	6	4	6	3
27	Features/facilities to recall key points browsed.	4	5	5	4	6	3
28	Ease in identifying location.	5	6	5	4	2	4
29	Ease in finding a particular topic thought to exist.	7	6	4	2	6	2
30	Ease in returning to a topic left previously.	3	5	7	6	4	6
31	Feeling of "lostness".	3	3	none	1	3	3
32	"Lostness" versus time wasting.	5	2	-	1	4	3
34a	Familarity of navigation tools after use.	6	6	7	1	6	5
34b	Familarity of structure of system after use.	6	6	6	1	6	5
34c	Familarity of information/content covered after use.	6	3	4	1	7	3
36	Confidence in using tools to complete tasks in:						
	- browsing	7	6	5	1	3	6
	- seeking references	5	5	6	1	5	6
	- information search	6	6	6	1	6	3
	- recall	6	2	7	1	6	2
39	Effectiveness of system to complete tasks.	5	6	6	1	6	4

To examine whether a successful implementation of the design principles has any influence on subject's satisfaction and perception of the effectiveness of the ACM's *Hypertext-on-hypertext* (objective 1), and to investigate whether design principles are embedded in the design of ACM's *Hypertext-on-hypertext* (objective 2), the raw data in table 3.3 were collated and analysed using the method described as follows:

1. Group questions according to the design principles to be tested. For example, group questions 1, 4, 31 and 36 together since they test principle P1. Repeat for all the questions. Table 3.4 shows the questions and the corresponding principles tested. Since this thesis is interested to investigate the relationships between end-users' ratings on the implementation of design principles and their perceived satisfaction (asked in question 3) and effectiveness of the hypertext (asked in question 39), questions 3 and 39 are left out of the table 3.4 under P4 as they would not be used in the calculation.

Table 3.4. Questions and corresponding principles tested in the original questionnaire

Principle tested	Closed questions
P1: Ensure a well-structured network of nodes and links.	1, 4, 31, 36
P2: Present clearly the hypertext structure.	24, 25, 28, 29, 30, 36
P3: Provide feedback.	14, 15, 19, 20, 21
P4: Ensure ease of use.	2, 6, 11, 16, 17, 18, 22, 27, 32
P5: Ensure consistency of presentation.	10, 12, 34a, 34b, 34c
P6: Ensure visibility.	5, 7, 8, 9, 13, 23

2. Convert subjects' ratings of these questions into three categories: "3 and below"; "4"; and "5 and above" as indicated in figure 3.11 to be "satisfactory", "neutral" or "good" rating. As an illustration (see figure 3.12), the first row of subjects' ratings in table 3.3 is selected to show how the raw data in table 3.3 were converted to respective tables 3.5a – 3.5c. For example, rating "4" for subject 1 is considered a count of "1" in table 3.5b under S1, rating "6" for subject 2 is considered a count of "1" in table 3.5c under S2, rating "1" for subject 4 is considered a count of "1" in table 3.5a under S4.

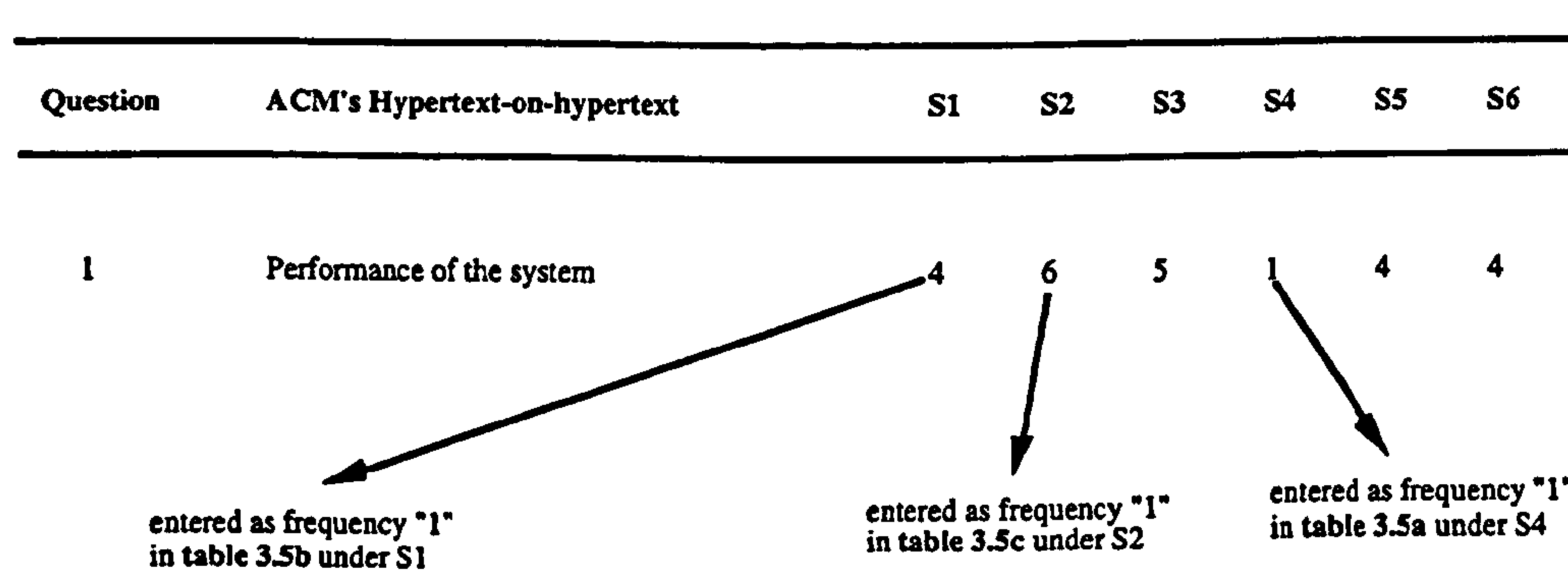


Figure 3.12. An illustration on how the raw data in table 3.3 are converted to tables 3.5a – 3.5c

Tables 3.5a to 3.5c show subjects' ratings on the implementation of design principles in ACM's *Hypertext-on-Hypertext* by adding the respective frequency of the subjects for each principle.

Table 3.5a. Subjects' ratings of value "3 and below" on the implementation of design principles of the ACM's Hypertext-on-hypertext.

	S1	S2	S3	S4	S5	S6
Principle 1	1	2	0	6	3	3
Principle 2	3	1	0	5	2	4
Principle 3	1	2	0	3	0	1
Principle 4	0	2	0	6	1	6
Principle 5	0	1	0	4	0	1
Principle 6	2	2	0	5	1	1
Total	7	10	0	29	7	16

Table 3.5b. Subjects' ratings of value "4" on the implementation of design principles of the ACM's Hypertext-on-hypertext.

	S1	S2	S3	S4	S5	S6
Principle 1	1	0	0	0	1	1
Principle 2	0	0	1	3	1	1
Principle 3	0	0	1	1	1	3
Principle 4	1	0	1	3	2	0
Principle 5	0	0	1	2	2	0
Principle 6	0	0	1	0	1	0
Total	2	0	5	9	8	5

Table 3.5c. Subjects' ratings of value "5 and above" on the implementation of design principles of the ACM's Hypertext-on-hypertext.

	S1	S2	S3	S4	S5	S6
Principle 1	5	5	6	0	3	3
Principle 2	6	8	8	1	6	4
Principle 3	3	2	3	0	3	1
Principle 4	8	7	6	0	6	3
Principle 5	5	4	4	0	3	4
Principle 6	4	4	5	1	3	5
Total	31	30	32	2	24	20

From tables 3.5a to 3.5c, subjects' ratings on how well they perceived the design principles were implemented were computed as shown in table 3.6 . The method used to convert the data from tables 3.5a to 3.5c to table 3.6 is described as follows:

1. Add subjects' ratings of all the three categories. Take S1 as an example. From tables 3.5a – 3.5c, it is noted that S1 had given the rating "3 and below" 7 times; rating "4" 2 times and rating "5 and above" 31 times. The total for each column is shown in tables 3.5a – 3.5c under the row "total". Repeat for subjects 2 to 6.

2. This thesis defines a "successful" implementation of design principle to be of a rating "5 and above". The formula to calculate the successful implementation of all six design principles by a subject is given as:

   $$\frac{\textit{number of times subject has given ratings "5 and above"}}{\textit{total number of times of subject's ratings for that principle}}$$

   Using S1 as an example, the successful implementation of all six principles in ACM's *Hypertext-on-hypertext* as perceived by S1 is: *31/40 x100% = 76%,* where 31 is the number of times S1 had given ratings "5 and above", and 40 is the total number of ratings given by S1. Repeat for subjects 2 to 6.

3. Rating on subject's satisfaction of the hypertext is converted into percentage using this formula:

   $$\frac{\textit{subject's rating on satisfaction x 100\%}}{7}$$

   where 7 is maximum rating that can be given in the 7-point scale.

   For example, S1's % rating of satisfaction when using ACM's *Hypertext-on-hypertext* is (see table 3.3, question 3): *4/7x100% = 57%*. Repeat for subjects 2 to 6.

4. Rating on subject's perception of the effectiveness of the hypertext is converted into percentage using this formula:

   $$\frac{\textit{subject's rating on effectiveness x 100\%}}{7}$$

   where 7 is maximum rating that can be given in the 7-point scale.

   For example, S1's % rating of perception of effectiveness of ACM's *Hypertext-on-hypertext* is (see table 3.3, question 39): *5/7x100% = 71%*. Repeat for subjects 2 to 6.

Table 3.6 compares subjects' ratings of the implementation of design principles calculated in percentages with subjects' ratings of their satisfaction and perception of the effectiveness of ACM's *Hypertext-on-hypertext* in helping them to complete the tasks.

Table 3.6. Percentage of subjects' ratings on the implementation of design principles versus subjects' satisfaction and perception of the effectiveness of ACM's Hypertext-on-hypertext

	S1	S2	S3	S4	S5	S6
Implementation of all 6 principles (%)	78	75	86	5	62	49
Satisfaction (%)	57	71	71	28	71	57
Effectiveness of hypertext (%)	71	86	86	14	86	57

From table 3.6, some interesting observations are made:

Except for subjects 4 and 6 (both inexperienced end-users as indicated in table 3.2), the perceived percentage of success of the implementation of the design principles was quite high for the other subjects: 78% for subject 1 (experienced end-user); 75% for subject 2 (experienced end-user); 86% for subject 3 (inexperienced end-user); and 62% for subject 5 (inexperienced end-user). Subject 4 registered a very low 5% while subject 6 a moderate 49%.

All subjects except subject 4 were generally felt quite pleased with ACM's *Hypertext-on-hypertext* registering 57% to 71%. Their perception of the effectiveness of the hypertext ranged from 57% to 86%, seems to be in positive agreement with their ratings on their satisfaction with the hypertext. The converse is also true. Subject 4's low rating of his satisfaction (28%) with the hypertext positively correlates with a low rating of 14% regarding his perception of the effectiveness of ACM's *Hypertext-on-hypertext*. Following from subject 6's moderate rating on the success of implementation of design principles, the ratings on satisfaction and effectiveness were also moderate.

Given the data from the six subjects, intuitively one would expect, more generally, a positive correlation between perceived success of an implementation of design principles and satisfaction and perception of the effectiveness of the hypertext in helping end-users to complete the tasks.

Objective 2: Results and analyses

To investigate whether good design principles were implemented in the design of ACM's *Hypertext-on-hypertext*, this thesis discusses subjects' responses under the seven areas G1 to G7 as captured in the questionnaire (see table 3.1). Tables 3.7 to 3.13 indicate subjects' ratings under the seven areas extracted from the raw data presented in table 3.3. To convert from table 3.3 to tables 3.7 to 3.13, the following method was used:

1. For each question, group subjects' ratings under three categories: "3 and below"; "4"; and "5 and above".

   Take question 1, for instance. Subjects 2 and 3 rated it "6" and "5" respectively giving a frequency of 2 under "5 and above" category. Subjects 1, 5 and 6 rated it "4" giving a frequency of 3 under the "4" category. Subject 4 rated it "1" giving a frequency of 1 under the "3 and below" category. Repeat for all the questions.

2. Tabulate the frequencies under the respective areas of questions in table 3.7 to table 3.13.

3. It is debatable but this thesis makes the assumption that if an area scores a percentage of 75 and above for ratings given in the "5 and above" category, it implies that the area is well-implemented in the hypertext in question.

This thesis turns now to the discussion of subjects' evaluation of ACM's *Hypertext-on-hypertext* according to the seven areas:

- **G1: Overall reactions to ACM's Hypertext-on-hypertext**

  Table 3.7 tabulates subjects' overall reactions to ACM's *Hypertext-on-hypertext.* These six questions evaluate subjects' overall perception of the performance of ACM's *Hypertext-on-hypertext* in terms of satisfaction, completion of tasks and appeal.

  Question 1 on the performance of ACM's *Hypertext-on-hypertext* did not score well since only 2 out of 6 rated it "good". Similarly, the appeal of the system (question 5) was rated poorly as only 1 out of 6 thought it was "good". As for subjects' satisfaction with the system (question 3) in helping with the completion of tasks, ACM's *Hypertext-on-hypertext* did not fare too badly. Overall, subjects' reactions to ACM's *Hypertext-on-hypertext* were moderate with only (17/36x100%)=47% of their ratings indicated that it was "good".

Table 3.7. Subjects' overall reactions to ACM's Hypertext-on-hypertext

	'3 and below'	'4'	'5 and above'
Question 1	1	3	2
Question 2	2	0	4
Question 3	1	2	3
Question 4	2	0	4
Question 5	3	2	1
Question 6	3	0	3
Total	12	7	17

36

- **G2: Screen display**

This area evaluates how information is organised and displayed on the screen. Subjects rated this area well with an overall percentage of (13/17x100%)=76% (see table 3.8) under the category "5 and above".

Table 3.8. Subjects' evaluation of screen display of ACM's Hypertext-on-hypertext

	'3 and below'	'4'	'5 and above'
Question 7	2	0	4
*Question 8	0	0	5
Question 9	2	0	4
Total	4	0	13

* one person did not answer this quesion.

- **G3: Terminology and system information**

Table 3.9 shows subjects' evaluation of the terminology used and system information provided in ACM's *Hypertext-on-hypertext*.

Except for poor or no feedback provided by ACM's *Hypertext-on-hypertext* as indicated by subjects (question 14), the system displayed consistency in the use of terminology, word and format (questions 10 to 12). On the whole, the subjects did not rate this area well with only a moderate (18/32x100%)=56%.

Table 3.9. Subjects' evaluation of the terminology used and system information provided in ACM's Hypertext-on-hypertext

	'3 and below'	'4'	'5 and above'
Question 10	0	2	4
Question 11	1	2	3
Question 12	1	1	4
Question 13	1	0	5
Question 14	4	1	1
*Question 15	0	1	1
Total	7	7	18

* 4 persons did not answer this quesion.

- **G4: Learning**

  This area investigates the ease of use of the ACM's *Hypertext-on-hypertext* with regard to how easy it is to learn to use the system, to explore new features by trial and error, and to remember names and use of commands. A total of (12/17x100%)=71% of subjects' responses rated it "good" (see table 3.10), implying that it was not too difficult to learn to use the system.

Table 3.10. Subjects' evaluation of learning features provided in ACM's Hypertext-on-hypertext

	'3 and below'	'4'	'5 and above'
Question 16	3	0	3
Question 17	0	1	5
*Question 18	1	0	4
Total	4	1	12

* 1 person did not answer this quesion.

- **G5: System capabilities and user control**

  Table 3.11 shows subjects' evaluation of the system capabilities and user control in ACM's *Hypertext-on-hypertext*. This area examines hypertext's response time, reliability and recovery process. It also investigates whether the hypertext design is simple or complex, taking into consideration subjects' needs and experience. A moderate (13/23x100%)=57% of subjects' responses indicates that this area is well-provided for.

Table 3.11. Subjects' evaluation of system capabilities and user control in ACM's Hypertext-on-hypertext

	'3 and below'	'4'	'5 and above'
Question 19	2	1	3
*Question 20	0	2	3
Question 21	1	1	4
Question 22	2	1	3
Total	5	5	13

* 1 person did not answer this quesion.

- **G6: Navigation**

Table 3.12 presents subjects' evaluation of navigation aids provided in ACM's *Hypertext-on-hypertext*. Only (23/52x100%)=46% of subjects' responses indicates that ACM's *Hypertext-on-hypertext* provides "good" navigation aids to help subjects navigate round it. Except for subject 3, all five subjects indicated feeling "lost", and all of them thought that LIH is a waste of time. Poor response to question 23 indicates that links to other information within hypertext is not clearly displayed and highlighted.

Table 3.12. Subjects' evaluation of navigation aids provided in ACM's Hypertext-on-hypertext

	'3 and below'	'4'	'5 and above'
Question 23	3	1	2
Question 24	1	1	4
Question 25	2	1	3
Question 27	1	2	3
Question 28	1	2	3
Question 29	2	1	3
Question 30	1	1	4
*Question 31	5	0	0
*Question 32	3	1	1
Total	19	10	23

* 1 person did not answer this quesion.

Question 26 asked subjects to identify the structure of hypertext from a choice of ten different structures. Responses to Question 26 will be discussed in chapter 5 which deals with how information is structured in hypertexts.

- **G7: Completing tasks**

This area examines the extent of usefulness of facilities in ACM's *Hypertext-on-hypertext* in helping subjects complete tasks in browsing, seeking references and revision. Table 3.13 shows subjects' evaluation of ACM's *Hypertext-on-hypertext* in helping them complete the tasks. Except for question 34(c), subjects rated the provision of navigation aids and structure "good" in helping them complete tasks successfully, with an overall moderately high score of (32/48x100%)=67%.

Table 3.13. Subjects' evaluation of ACM's Hypertext-on-hypertext in helping them complete tasks

	'3 and below'	'4'	'5 and above'
Question 34a	1	0	5
Question 34b	1	0	5
Question 34c	3	1	2
Question 36			
- browsing	2	0	4
- seeking references	1	0	5
- information search	2	0	4
- revision	3	0	3
Question 39	1	1	4
Total	14	2	32

Summary

Subjects' responses to open questions (35, 37, 38 and 40) were generally positive. However, they commented that they preferred hypertext with a user interface that is simple and easy to use. They also said it would be added bonus if the hypertext is interesting and fun, with clever use of colours, sound and graphics. Carroll's 1982 paper "The adventure of getting to know a computer" advocates the game metaphor to present end-users with an exploratory environment and turn obstacles such as disorientation into challenges. Perhaps designers can exploit the game metaphor in making hypertexts easier to learn and more fun to use as well. This could be a seventh design principle of "making it fun to use", on top of the six proposed previously.

Table 3.14 compiles subjects' evaluation of ACM's *Hypertext-on-hypertext* according to the seven areas listed in the questionnaire to test the design principles. As discussed previously, it is open to debate, but this thesis has chosen 75% as a benchmark score to decide whether an area as perceived by the subjects is well-implemented. By analysing the results, it would seem that except for G2 on screen design, all other areas were not well-implemented in ACM's *Hypertext-on-hypertext* according to subjects' responses.

Table 3.14. Compilation of subjects' ratings of the success of implementation of design principles based on the seven areas tested in the questionnaire

	Ratings of implementation	Remark
G1: Overall reactions	47%	not well-implemented
G2: Screen display	76%	well-implemented
G3: Terminology and system information	56%	not well-implemented
G4: Learning	71%	not well-implemented
G5: Systems capabilities and user control	57%	not well-implemented
G6: Navigation	44%	not well-implemented
G7: Completing tasks	67%	not well-implemented

Experiment 3.1 was designed to achieve two objectives. The first objective was to examine whether end-users' perception of the success of the implementation of design principles has any influence on their satisfaction and perception of the effectiveness of ACM's *Hypertext-on-hypertext*. From analysing subjects' responses, there is a positive correlation between subjects' perception of the success of implementation of design principles and subjects' satisfaction and perception of the effectiveness of ACM's *Hypertext-on-hypertext* (see table 3.6).

The second objective was to investigate whether the six design principles, as described previously in §3.3.2, were effectively embedded in the design of a commercially-produced hypertext, ACM's *Hypertext-on-hypertext*. The results have shown that ACM's *Hypertext-on-hypertext* was not too well rated, based subjects' responses to the questionnaire (see table 3.14). Of the seven areas tested in the questionnaire, only the screen display (G2) scored well. In general, subjects felt that ACM's *Hypertext-on-hypertext* could be improved by incorporating better features such as: (i) history list to trace the paths taken; (ii) guide to use the system; (iii) map to show where subject is, where he has been and where he can go; (iv) better help facility; and (v) intelligent recommendation facility/context-sensitive index.

Chapter 5 will discuss in greater detail the relationships among these factors in areas G6 and G7 on: subjects' representation of the structure of ACM's *Hypertext-on-hypertext*, structure of ACM's *Hypertext-on-hypertext*, subjects' experience, subjects' satisfaction, subjects' perception of the effectiveness of ACM's *Hypertext-on-hypertext* in helping them to complete tasks, LIH, time spent in completing tasks and subjects' actual performance in terms of tasks completed. Chapter 5 will also discuss subjects' responses to question 26 regarding their representation of the structure of ACM's *Hypertext-on-hypertext* and the phenomenon of feeling "lost".

3.5.2 Experiment 3.2: A prototype hypertext

Objectives

Experiment 3.2 investigates whether the six design principles proposed in §3.3.2 and embedded in the prototype hypertext are "good enough".

Protocol

Figure 3.13 illustrates the framework to carry out an experiment to get feedback on design principles for hypertexts.

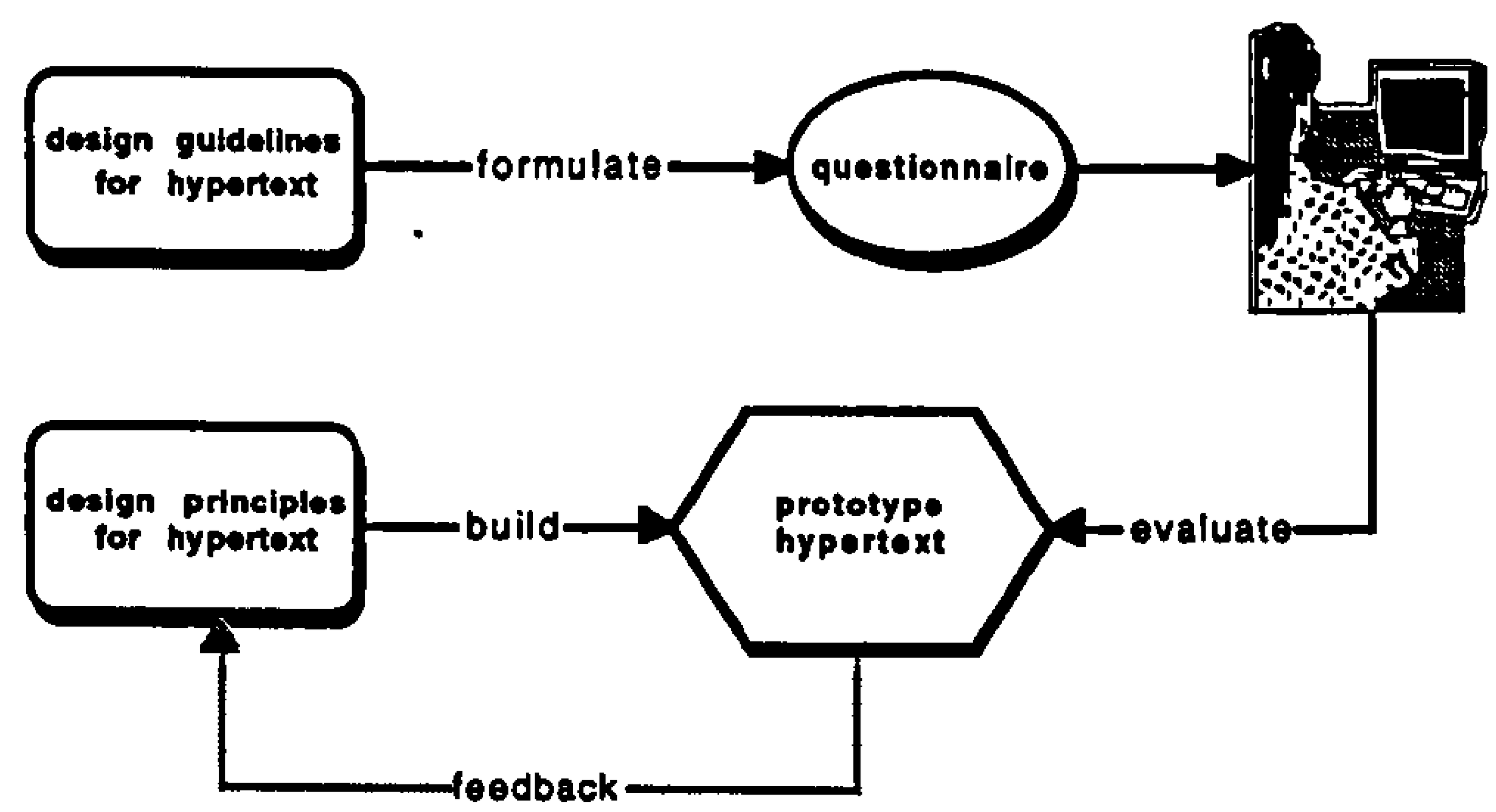


Figure 3.13. Framework for experimental illustration and analysis (experiment 3.2) of design principles for hypertexts

Three "typical" subjects were asked to evaluate the prototype as described in §3.3.3. The three subjects were students from Middlesex University. One of them was an undergraduate in applied computing, the second was an undergraduate in mechanical engineering and the third was a postgraduate in business computing. By "typical", this thesis refers to an end-user who has general knowledge about navigation with hypertexts. The subjects were asked to complete some exercises associated with browsing. They were then interviewed and asked to complete the revised questionnaire on the performance of the prototype. The revised questionnaire contains twenty-four questions (see appendix B2). The number of questions was reduced because the subjects in experiment 3.1 commented that there were too many questions asked. Questions were carefully selected to ensure that the questionnaire was still able to test the different design principles as well as to retain the most relevant design categories to assess navigation and the LIH problem. Instead of testing seven areas as outlined in the original questionnaire, this revised questionnaire tests only five areas concentrating on the navigational, control and learning aspects. They include: (i) overall reactions to the system; (ii) learning; (iii) system capabilities and user control; (iv) navigation; and (v) completing tasks.

Table 3.15 shows the questions and the corresponding principles tested (questions 3 and 23 are left out in table 3.15 under P4 as they should not be used in converting the data. Principle P3 is evaluated by open questions 21 and 22). As before, the relationships between subjects' ratings on the implementation of design principles and their perceived satisfaction (asked in question 3) and effectiveness of the hypertext (asked in question 23) can be investigated using the method described in §3.5.1 (under "results and analyses" section for objective 1). This, however, would not be illustrated in this section.

Table 3.15. Questions and corresponding principles tested in the revised questionnaire

Principle tested	Closed questions
P1: Ensure a well-structured network of nodes and links.	1, 15, 20
P2: Present clearly the hypertext structure.	8, 9, 12, 13, 14, 20
*P3: Provide feedback.	-
P4: Ensure ease of use.	2, 4, 5, 6, 11, 16
P5: Ensure consistency of presentation.	18
P6: Ensure visibility.	7

*P3 is evaluated using open questions 21 and 22

Results and analyses

To determine whether the six design principles embedded are "good enough", the method described in §3.5.1 (under "results and analyses" section for objective 2) to convert the raw data was used. Table 3.16 shows subjects' ratings of the prototype.

Table 3.16. Subjects' ratings on the success of implementation of the design principles in the hypertext prototype

	Questions	Subject's rating		
		S1	S2	S3
	Overall reactions to the system			
1.	How would you rate the performance of the system?	4	6	6
2.	How would you rate the usability of the system?	5	5	7
3.	How would you rate the level of satisfaction when using the system?	3	6	5
	Learning			
4.	Do you find it easy to learn how to use the system?	4	6	7
5.	How easy is it for you to explore new features by trial and error?	3	5	7
	System capabilities and user control			
6.	Has the system taken both experienced and inexperienced users' needs?	5	-	6
	Navigation			
7.	Are the links to other information within the system clearly displayed and highlighted?	1	7	6
8.	Are there sufficient information on screen to help you know where you are in relation to the structure (e.g. use of title, page numbers, etc.)?	4	6	5
9.	Does the system provide sufficient information to help assess the structure of the system (e.g. browsers, maps, table of contents, etc.)?	4	7	5
11.	Does the system provide good features/facilities to help you recall what you have browsed or the key points covered?	6	6	2
12.	Is it difficult to identify where you are in the system?	7	6	6
13.	Is it difficult to find a particular topic thought to exist?	2	7	7
14.	Is it difficult to return to a topic left previously?	2	7	7
15.	To what extent do you feel 'lost' when using the system?	3	no	no
16.	To what extent do you feel that being 'lost' is waste of time?	4	-	-
	Completing tasks			
18.	To what extent are you familiar with the system after completing the tasks with respect to navigation tools, the way information is organised, and information covered?			
	18a. familarity of navigation tools after use	7	6	6
	18b. familarity of structure of system after use	6	5	6
	18c. familarity of information/content covered after use	7	5	7
20.	How confident did you feel when using these features/tools to complete the tasks?			
	- browsing	7	7	7
	- seeking references	7	7	7
	- information search	7	7	7
	- recall	-	7	7
23.	How would you rate the effectiveness of the system in helping you complete the tasks?	4	6	6

Tables 3.17 to 3.21 show subjects' evaluation of the prototype hypertext according to the five areas tested. This section will not elaborate on how the tables were tabulated since the method for converting the raw data presented in table 3.16 was described in detail in §3.5.1.

- **Overall reactions to prototype hypertext**

Table 3.17 tabulates subjects' overall reactions to the prototype. The three questions evaluate subjects' overall perception of the performance of the prototype in terms of satisfaction, completion of tasks and performance.

Subjects rated the overall impression of the prototype well since their ratings indicated that it was "good" with a percentage score of (7/9 x 100%)=78%.

Table 3.17. Subjects' overall reactions to the prototype hypertext

	'3 and below'	'4	'5 and above'
Question 1	0	1	2
Question 2	0	0	3
Question 3	1	0	2
Total	1	1	7

- **Learning**

These two questions evaluate subjects' responses with regard to how easy they found it to use the system and to explore new features by trial and error. A total of (5/6x100%)=83% of subjects' responses rated it "good", implying that it is not difficult to learn how to use the prototype (see table 3.18).

Table 3.18. Subjects' evaluation of the learning features provided in the prototype hypertext

	'3 and below'	'4'	'5 and above'
Question 4	0	0	3
Question 5	1	0	2
Total	1	0	5

- **System capabilities and user control**

One of the subjects did not answer question 6. The other two felt that the prototype was simple to use, taking into consideration subjects' needs and experience (see table 3.19).

Table 3.19. Subjects' evaluation of system capabilities and user control in the prototype hypertext

	'3 and below'	'4'	'5 and above'
Question 6	0	0	2
Total	0	0	2

* 1 person did not answer this quesion.

- ### Navigation

Table 3.20 presents subjects' evaluation of navigation aids provided in the prototype. A total of (15/23x100%)=65% of the subjects' responses indicates that the prototype provides "good" navigation aids. One out of the three subjects indicated feeling "lost".

Table 3.20. Subjects' evaluation of navigation aids provided in the prototype hypertext

	'3 and below'	'4'	'5 and above'
Question 7	1	0	2
Question 8	0	1	2
Question 9	0	1	2
Question 11	1	0	2
Question 12	0	0	3
Question 13	1	0	2
Question 14	1	0	2
*Question 15	1	0	0
*Question 16	1	0	0
Total	6	2	15

* 2 persons did not feel "lost" so questions 15 and 16 were not answered.

- ### Completing tasks

Table 3.21 shows subjects' evaluation of the prototype in helping them to complete the task. A high score of (22/23x100%)=96% indicates that subjects were pleased with the provision of navigation aids and structure was "good" in helping them complete the tasks successfully.

Table 3.21. Subjects' evaluation of the prototype hypertext in helping them to complete tasks

	'3 and below'	'4'	'5 and above'
Question 18a	0	0	3
Question 18b	0	0	3
Question 18c	0	0	3
Question 20			
- browsing	0	0	3
- seeking references	0	0	3
- information search	0	0	3
*- revision	0	0	2
Question 23	0	1	2
Total	0	1	22

* 1 person did not answer

Summary

Subjects' responses to the open questions of 19, 21 and 22 were generally positive. They liked the prototype because it was simple to use and easy to navigate to look for information. Because the prototype had a consistent "look-and-feel", the subjects felt that the prototype gave them useful "feedback" in relation to where they were, and how they could navigate using the navigation buttons. However, they thought that the appearance of the prototype could be improved with colours and graphics (which of course it can, but this has nothing to do with LIH). There were also some buttons which did not prove to be useful. One subject found the "move to first page" button confusing. Another subject felt that all the navigation buttons should be placed at the top left corner, instead of putting them also at the button of the screen. However, the consistency of the presentation compensated for the initial "confusion". On the whole, from question 23, it is to be noted that the subjects rated the prototype (2/3x100%)=67% effective in helping them to complete the task. The subjects were generally satisfied (2/3x100%)=67% with the prototype, as observed in question 3.

Table 3.22 compiles subjects' evaluation of the prototype according to the five areas listed in the questionnaire to test the design principles that have been implemented in the prototype. Except for the navigational support provided by the prototype, it would seem that the subjects thought the other areas were well-implemented scoring above 75%, the benchmark score for "good" implementation, as considered by this thesis.

Table 3.22. Compilation of subjects' ratings of the success of implementation of design principles based on the five areas tested in the questionnaire

	Ratings of implementation	Remark
Overall reactions	78%	well-implemented
Learning	83%	well-implemented
Systems capabilities and user control	100%	well-implemented
Navigation	65%	not well-implemented
Completing tasks	96%	well-implemented

3.6 Assessment

As the author reflects on what had been done, there were things done that were novel, but there were also many interesting things the author wished she had done at the time of writing.

3.6.1 Framework

This chapter proposed a framework on how to formulate design principles and guidelines for hypertexts. The idea of drawing up this framework was not an easy task. Design principles and guidelines are difficult to formulate and assess because they are "intuitive", subjective and often vague. This is perhaps one of the reasons why there has been so little research done in this area. In retrospect, the framework proposed did not look as neatly structured as when implemented! The framework evolved and became clearer as the investigation continued. Great care was taken to ensure that this thesis reported as accurately as possible what was implemented and how it formulated the list of design principles and guidelines for hypertext authoring, since the purpose was to simplify exposition and more clearly to highlight the modelling issues involved.

The ultimate use of principles and guidelines is to provide reference and guidance for designers during the development process so as to shorten and reduce the number of iterations involved in the design-evaluate-redesign cycle and make them more effective. This thesis borrowed findings from general design principles and guidelines, cognitive principles and usability testing methods. The steps taken in this framework may be open to debate! As one goes through the steps suggested in this framework, one can think of many interesting ideas to improve the framework. What distinguishes science from other disciplines is that it is "prepared for, indeed encourages, its own criticism, testing, refutation and eventual replacement" (Thimbleby, 1990), and if what has been done in this chapter is able to stimulate deeper thinking and further research into this whole area of design principles and guidelines, it has fulfilled its purpose.

3.6.2 Design principles: Prototype hypertext

Intuitively speaking, the six design principles when implemented should have a positive impact on the end-users' satisfaction and perception of the hypertexts. Designers should ensure that design principles be incorporated into the design of the hypertexts right from the start, so that many design blunders can be avoided.

When building the prototype, conscious efforts were made to ensure that the six design principles were implemented. However, the subjects still felt that not all the design principles were successfully implemented. For example, they felt that the navigational support provided by the prototype though not too "bad", could be further improved. Suggestions were made to arrange the navigation buttons. The thesis learned from this experiment that designing hypertext is not an easy task. Even though one may be aware of the design principles, the resultant prototype might not be reflective of the implementation. Many hypertexts are not well-designed (Thimbleby, 1995a). The question to ask is: why then are hypertext authors not able to design hypertexts that they want, or end-users need? Can it be the case of "they do not want to do the things they do, but they do it because they have no choice or help"? The thesis will be exploring this issue later in chapter 4, where the intention is to better understand end-users' needs, and provide better navigational help. Furthermore, chapter 7 describes how design principles can be integrated into an authoring tool so that designers need not worry about the implementation of the principles, but are free to concentrate on the content of the hypertext.

3.6.3 Design categories: Questionnaire

Since the purpose of this chapter is to get quick results and impressions about the design categories proposed, a small sample size was used to test the questionnaire (Nielsen, 1993). With more data, more sophisticated statistics would have yielded more reliable results, but really the problem with the data in the experiments is the experimental biases (for example, experiments were not double blind, all subjects were students, *etc.*) and lack of psychological control over individual differences (for example, field dependency). Despite these easily corrected limitations, the experiments were usefully informative.

This thesis has shown from preliminary findings in experiments 3.1 and 3.2 that the questionnaire containing the design categories provides a means to test whether design principles have been satisfactorily implemented. However, this thesis would suggest as recommendation for future work that more testing should be carried out to establish external, construct or predictive validity of the questionnaire, that is, the design categories.

3.7 Conclusion

This chapter began with a discussion of general design principles and guidelines for designing interactive computer systems. The need for good design principles and guidelines for the design of hypertexts is obvious since not many good hypertexts are built, resulting in a set of problems associated with hypertexts, of which the LIH problem is one of them. This thesis began by describing a framework used in deriving a set of design principles and guidelines for hypertext authoring. Six design principles have been proposed: (1) ensure a well-structured network of nodes and links; (2) clearly present the hypertext structure; (3) provide feedback; (4) ensure ease of use; (5) ensure consistency of presentation; and (6) ensure visibility. Based on these six design principles, a prototype hypertext was built using HyperCard. To assess whether the prototype was "good enough", design categories testing these six areas were proposed: (i) overall reactions to hypertext; (ii) screen display; (iii) terminology and screen information; (iv) learning; (v) system capabilities; (vi) navigation; and (vii) completing tasks. These design categories were presented in a form of a questionnaire to evaluate the hypertext prototype.

Initial findings from the experiments support the need also to make the hypertext fun to use, which this thesis would recommend as a seventh design principle. From the experiments, it is to be noted that designers need more than just design principles and guidelines. They need authoring support to help them cope with the complexity of the design process. This thesis will examine other design issues such as looking into the need to understand end-users better (chapter 4), the need for good hypertext structure (chapter 5) and the need for better authoring tools (chapter 6).

Chapter 4

Need for an engineering, task-based approach to understand end-users

Chapter objectives:

To build quality hypertexts, designers need to understand end-users. This is uncontentious wisdom, but so far, doing this is not easy. This chapter explains a systematic, engineering task-based approach to understand end-users' behaviour and the common tasks they perform when navigating hypertext. Its objectives are to explain:

- how end-users' common tasks can be understood by breaking them down into simpler subtasks
- how the results of this cognitive task-based approach can guide design of hypertext
- how the usability of hypertext can be determined using evaluation methods and techniques
- how the conventional development lifecycle can be improved

4.1 Introduction

Authoring hypertexts is complex and difficult. Designing usable hypertext requires knowledge about who will use it, what will it be used for, the work context and the environment in which it will be used, and what is technically and logistically feasible. It is a well-known fact that designers often design for themselves unless they are trained to realise that people are diverse, and that end-users are unlikely to be like them (Landauer, 1995; Thimbleby, 1990). The more errors that can be avoided "up front" by the right method, the less work both test users and designers will have to put in to make prototypes acceptable. Unfortunately the conventional exhortations do not help make much impact on hypertext design problems. The hypertext design space is astronomically large (Thimbleby, 1995a). Iterative design, that usually helps improve systems, has problems in that end-users involved in the process experience too little of a system to help make significant design contributions.

Figure 4.1 shows the six stages described in the well-accepted iterative development process (Newman and Lamming, 1995).

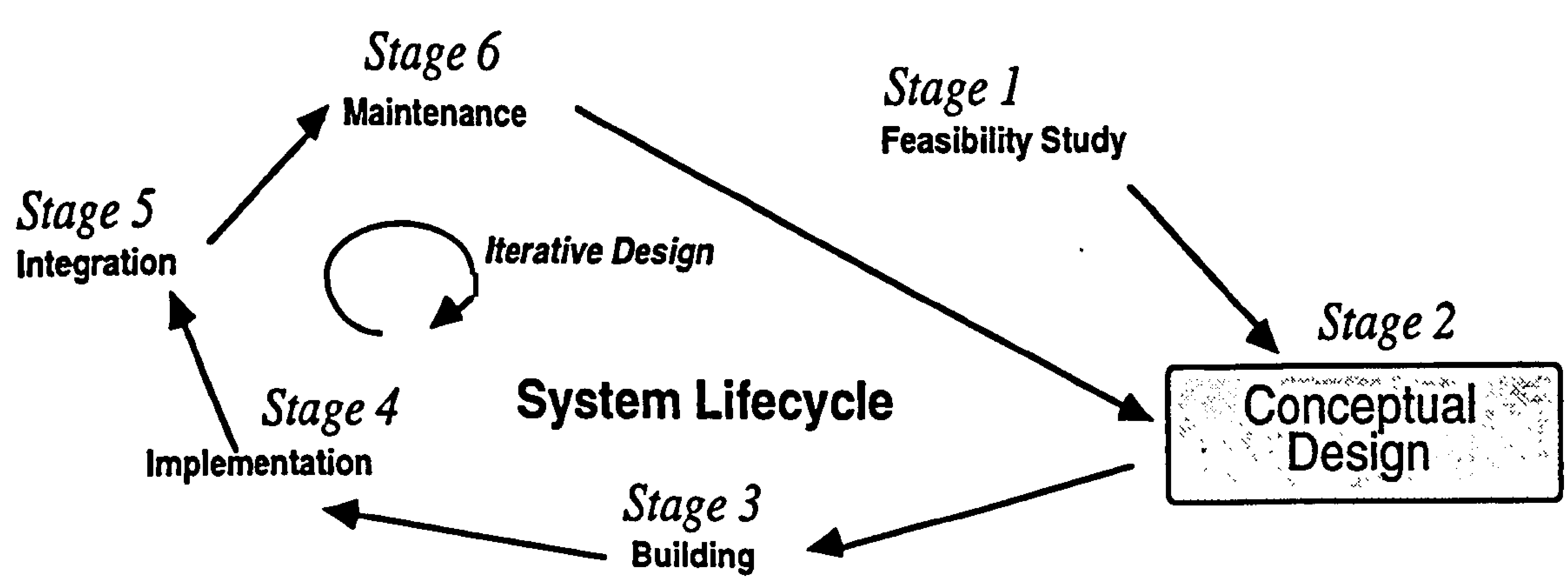


Figure 4.1. Conventional iterative development process

In the conventional iterative process, Conceptual Design (see figure 4.2) consists of these activities: analysing end-users' needs; building prototype hypertext; and testing and validating prototype with real users. The validated prototype is ready for building in Stage 3 after the iterative test-modify-test cycles. The Conceptual Design is vital to the success or failure of the final system as it encompasses all activities relating to gathering and analysing end-users' needs, specifying end-users' requirements, building and testing prototype.

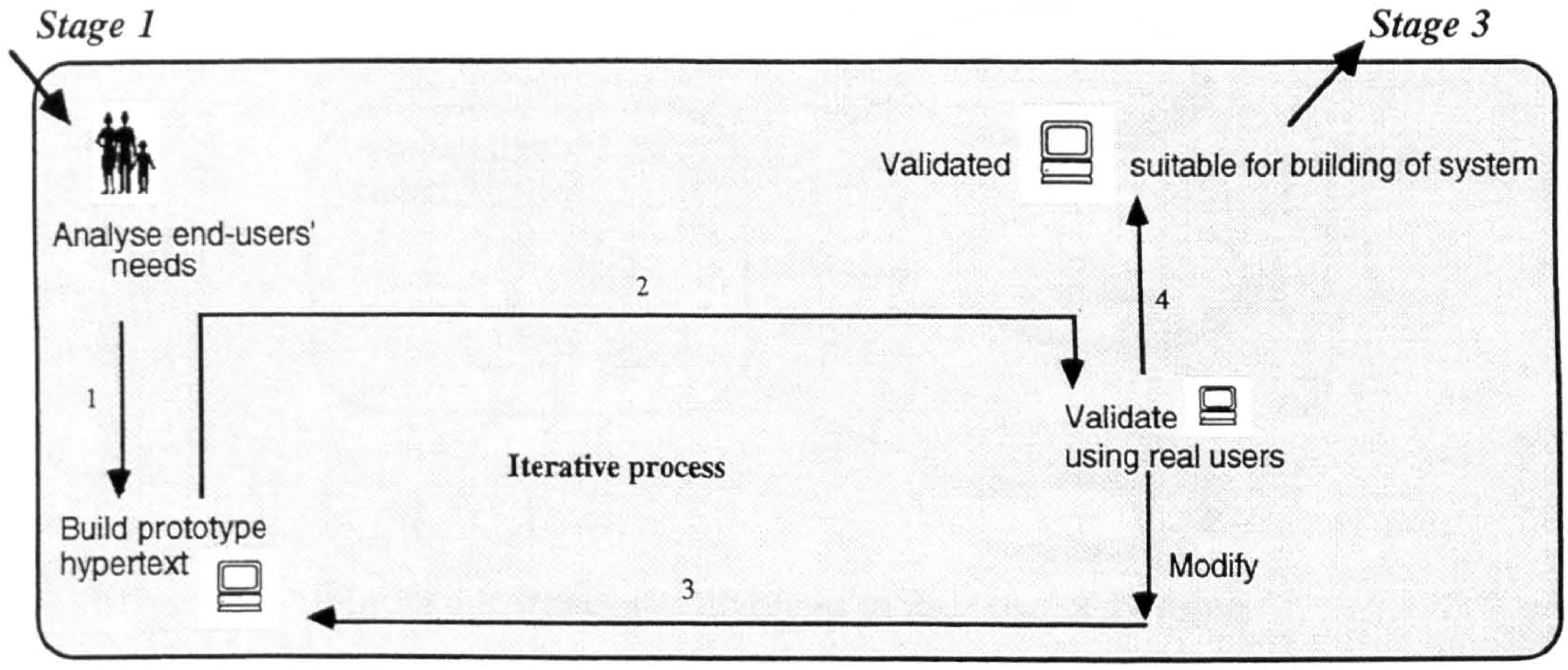


Figure 4.2. Conceptual design stage in conventional iterative development process

However, the Conceptual Design Stage in the conventional iterative process poses some problems:

Problem 1: There is a lack of a disciplined and systematic approach to designing well-structured hypertexts that will meet end-users' behaviour and navigation needs (Glushko, 1992; Carmel, Crawford and Chen, 1992).

Problem 2: The prototype hypertext is not thoroughly tested before developing into the final system. Glushko (1992) points out that often researchers would lavish time on a demonstration portion of a planned hypertext, but will have less time to customise other parts of the hypertext the same way.

If conventional iterative development process is found lacking, better ways to ensure that good hypertexts are produced are needed. A key aim of the design of good interactive systems is to make end-users the focus of design activity (Landauer, 1995; Newman and Lamming, 1995; *etc.*), hence the term *user-centred design*. This can be achieved by involving end-users and taking their needs throughout the design process. This is uncontentious wisdom, but so far, doing this is not easy. Figure 4.3 shows some principles involved in user-centred design, which include carrying out task analysis, early testing and evaluation, designing iteratively with many test–modify-test cycles (Preece, Benyon, Davies, Keller and Rogers, 1993). Through the process of iterative design, designers will be able to tailor hypertexts more and more to end-users' needs.

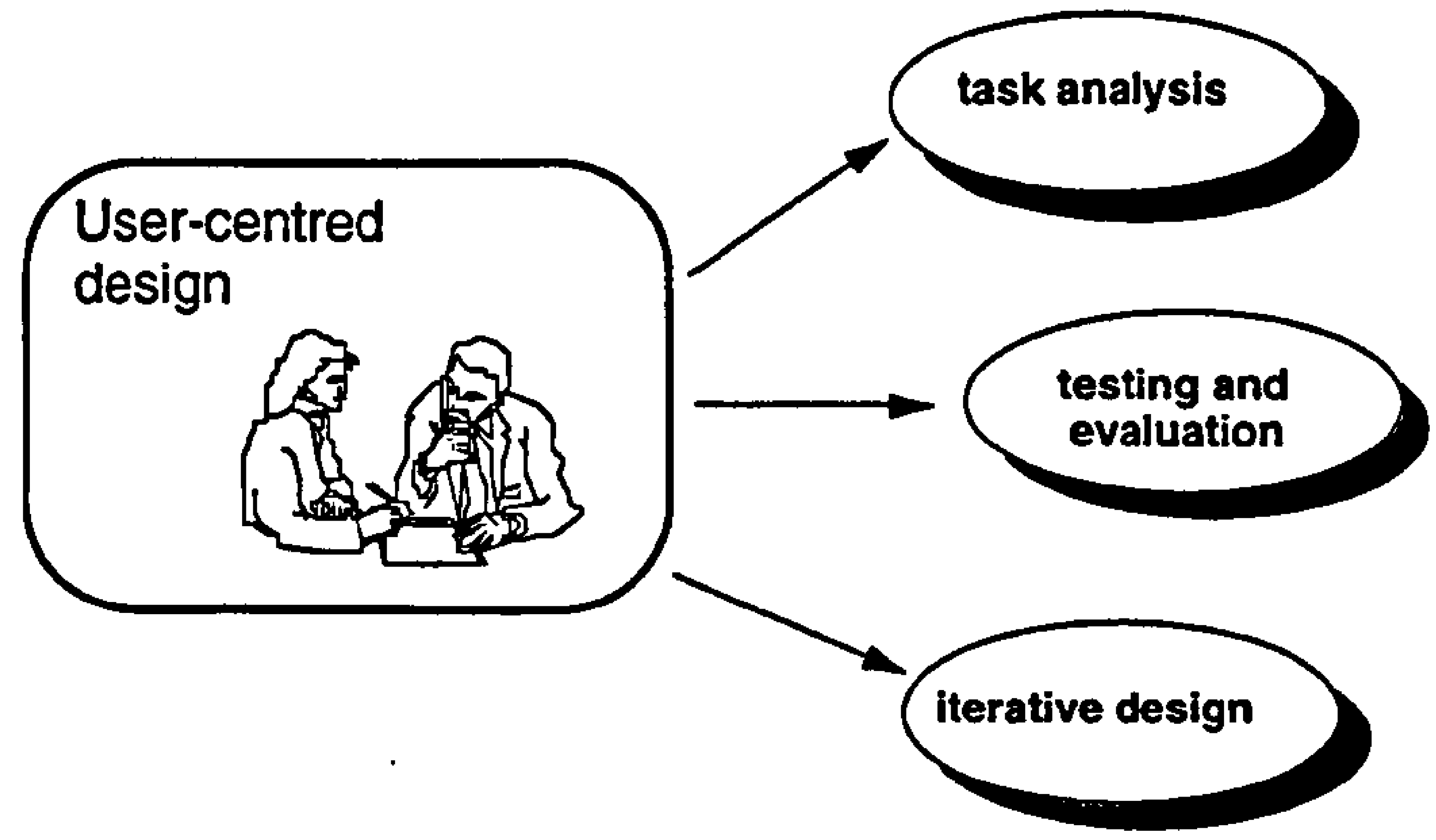


Figure 4.3. Principles involved in user-centred design

This chapter proposes a systematic, engineering task-based approach to understand end-users' behaviour and the common tasks they perform when navigating hypertexts. By incorporating user-centred design principles into the *Conceptual Design* (Stage 2 of the conventional iterative process), this chapter aims to address the preceding criticisms of the Conceptual Design Stage:

- To address Problem 1, §4.2 presents a cognitive, task-based approach to understand end-users' behaviour and navigation issues when using hypertext by giving them more structure.

- To address Problem 2, §4.3 investigates established evaluation methods and techniques, and discusses how some of these can be effectively employed by designers to measure the usability and usefulness of hypertexts, thus ensuring that quality hypertexts are produced.

4.2 A cognitive, task-based approach

Understanding end-users and their tasks has always been an important, research issue ranging from Hansen's "know the user" (Hansen, 1971), to Lewis and Rieman's "task-centred design" (Lewis and Rieman, 1993). If designers attempt to design an interactive system without an adequate understanding of end-users' tasks, Newman and Lamming (1995) caution that designers would not get very far. By trying to make sense of what end-users should do or what they actually do, designers should stand a better chance of producing user-centred hypertexts that will meet end-users' needs more effectively. Nielsen (1993) argues that an understanding of end-users and their tasks guarantees better design. From the study carried out by Rosson, Mass and Kellogg (1990), user-task analysis was among one of the several techniques used by designers for "idea generation". Other techniques involve using external sources such as available literature, meta-strategies such as concentration, design activities such as charts and diagrams, as well as trial and error.

Before describing a cognitive, task-based approach postulated in this chapter, §4.2.1 begins with a literature review to better understand task modelling and analysis. Because browsing is a very common exploration strategy, and that little research is done to understand browsing (Carmel, Crawford and Chen, 1992), §4.2.2 defines browsing, examines strategies end-users employ in browsing, identifies the factors that affect browsing, and surveys experimental research on browsing. Using the task *browsing* as a concrete example, §4.2.3 describes how cognitive task diagrams for browsing are built. The results of this cognitive task-based approach are used to guide the design of a prototype hypertext.

4.2.1 Task modelling and analysis

One set of methods concerned with modelling activities in terms of tasks is known widely as task analysis (TA). Generally, TA analyses tasks in terms of human behaviour (Johnson, 1992). By focusing on specific user goals, designers can often build a fairly precise model of the end-user's activity (Newman and Lamming, 1995). Early work on TA was strongly influenced by training needs, for example, task analysis for knowledge description (Johnson, Diaper and Long, 1984; Diaper, 1989), and hierarchical task analysis (Annett and Duncan, 1967). More recently, TA has been given greater importance in the context of interactive system design (Newman and Lamming, 1995; Johnson, 1992). TA needs to be done in addition to a general requirements analysis, which tends to focus on *what* functionality is required and not *how* to provide the functionality (Preece *et al.*, 1993).

What is a task?

The concept of a task is central to TA. A task is a unit of human activity carried out in order to achieve a specific goal (Preece *et al.*, 1993; Diaper, 1989). Johnson (1992) extends the definition of a task to include agents that bring about some change of state in a given domain; agents include people, animals or machines. The performance of a task usually involves a sequence of steps, contributing towards achieving the task's goals. However, the sequence of steps taken is influenced by surrounding circumstances, a phenomenon described by Suchman (1987) as situated action.

There are many ways to classify tasks. Marchionini (1989) classifies tasks as *closed* and *open*; closed tasks have a specific objective whereas open tasks have a general objective. Johnson (1992) suggests that one way of grouping tasks is in terms of roles or jobs agents undertake. However, it is a well-known fact that the grouping of tasks is not as clear-cut as it appears. For a task to be more precisely identified and defined, Johnson (1992) suggests that designers take into account the context in which it is found.

This thesis makes a distinction, and groups hypertext tasks into *generic* and *context-specific* tasks. Generic tasks are defined as common tasks end-users need to perform when using any hypertext, which include activities such as browsing, information search, recall and seeking references. Generic tasks are, therefore, application-independent. On the other hand, non-generic or context-specific tasks are tasks that are unique for different applications. For instance, an end-user using an educational hypertext will most certainly have goals that are different from using, say an information retrieval hypertext. Because this thesis focuses on addressing the LIH problem in general hypertexts, only generic tasks would be considered.

Methods and techniques

To make the process of collecting and abstracting information about end-users' tasks easier and more systematic, a number of task analysis techniques have been developed over the years. These techniques can broadly be divided into three areas (Preece *et al.*, 1993) namely macro methods, intermediate methods and micro methods. Figure 4.4 shows a taxonomy of simple task analysis techniques.

- *Macro methods* analyse the whole system in terms of organisational, social and environmental aspects. Examples are Open System Task Analysis (Eason and Harker, 1980) and User System Task Match (MacCaulay, Fowler, Kirby and Hutt, 1990). Flowcharts and natural language descriptions constitute the main ways of describing information collected with these methods.

- *Intermediate methods* decompose discrete tasks into hierarchical tree of subtasks. They are used mostly by industrial specialists. An example is the Hierarchical Task Analysis, pioneered by Annett and Duncan (1967), produces many variants adapted to different contexts. For example, hierarchical planning emphasises end-users' goals rather than the operations that end-users carry out at a computer interface (Sebillotte, 1988); and techniques for charting the events and necessary responses in air traffic control (Phillips, Bashinski, Ammerman and Fligg, 1988).

- *Micro methods* are concerned with cognitive modelling. The aim is to model the cognitive knowledge and physical actions that end-users must carry out in order to do a task. These are very fine grain techniques, which are also used in analytic evaluation. Cognitive modelling techniques create a variety of models describing end-users, their interactions and interfaces used to support their interactions. These models have been applied to predict end-user behaviour, communicate designs, provide consistency between levels of design, and evaluate aspects of a design (Howes, 1995). Well-known cognitive modelling techniques include: Common Language Grammar (CLG) (Moran, 1981); Task Action Language (TAL) (Reisner, 1981); Task Action Grammar (TAG) (Payne and Green, 1986); Goals, Operators, Methods and Selection Rules (GOMS) (Card, Moran and Newell, 1979); Cognitive Complexity Theory (CCT) (Kieras and Polson, 1986); Interacting Cognitive Subsystems (ICS) (Barnard, 1987); and Programmable User Models (PUMs) (Blandford and Young, 1995).

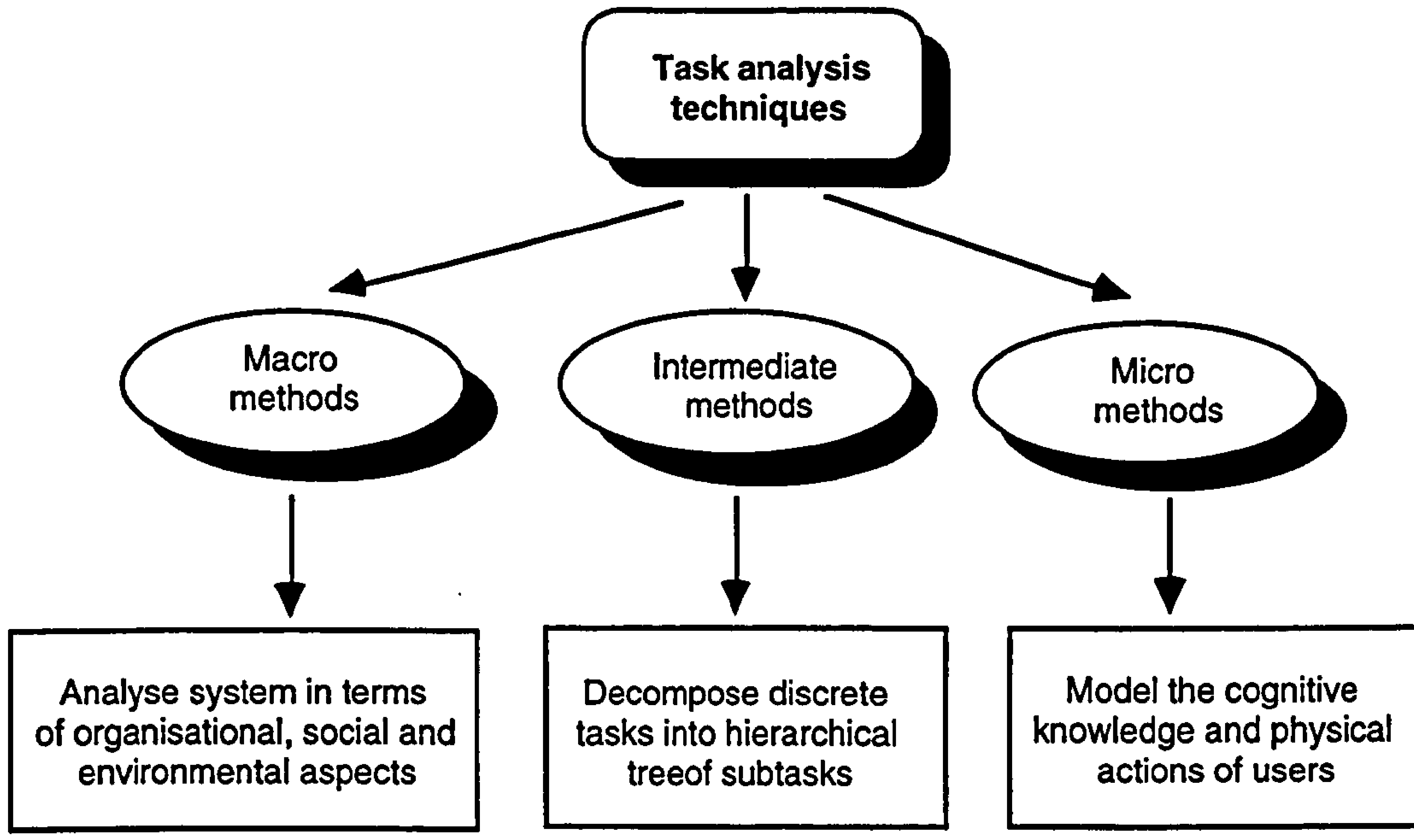


Figure 4.4. Taxonomy of task analysis techniques

Opposing views on task modelling and analysis

Although task models are aimed at supporting the design process, and the design of the human-computer interfaces in particular, by taking into consideration end-users and their tasks, in most cases, there is little information about the formalism and how to use them to develop models of interaction (Phillips *et al.*, 1988). They are mostly theoretical models, and most design experiments carried out with these models have shown that they are difficult and time consuming to use (Rigny, 1997; Anderson, Carroll, McGrew, Grudin and Scapin, 1990; Wilson, Barnard and MacLean, 1986). Another criticism concerning TA is its applicability which include (Lim, 1996): failure to contribute to novel designs; failure to contribute directly to design specification; addressing only a limited part of the design specification; lacking in method; lacking rigour; and procedurally inadequate.

On the other hand, Johnson (1992) believes that TA would have the most impact specifically during the feasibility and requirements stages of design. He advocates that TA at these stages could be used to identify the functionality of the system, as well as the design of the interaction dialogue. Lim (1996) goes on to suggest that for TA to contribute effectively, they should be structured and incorporated at each stage of the development lifecycle.

Browsing: An example of generic hypertext task

This thesis acknowledges the limitations of TA. However, to produce usable and useful hypertexts, it is important to understand end-users' behaviour and navigation issues when using hypertext. To do so would demand that end-users' needs be analysed, and one way to do that is to analyse the tasks they perform. Since navigating hypertext is very much a cognitive activity (Carmel, Crawford and Chen, 1992), this thesis proposes a cognitive, task-based approach to give more structure to end-users' behaviour and navigation issues. Inspired by Buckley and Johnson (1987) that the purpose of TA is to contribute to design, the outputs of TA could then be used to guide the design of a prototype hypertext during the Conceptual Design stage.

To develop an accurate understanding of end-users, it is essential that representative tasks which provide reasonably complete coverage of a system's intended functionality are chosen (Lewis and Rieman, 1993). Therefore, to identify the representative tasks in hypertext, this thesis considers the kinds of support hypertexts generally provide: *applications support* (for example, tutorial and educational needs); *support for collaboration* among a number of individuals and production of on-line manuals; *functional support* (for example, browsing and authoring); and *cognitive support* (for example, reading, annotating, collaborating and learning). Based on the kinds of support provided by hypertext, four representative tasks end-users want to perform (or try to perform) when using hypertext are identified. These tasks are browsing, information search, recall and seeking references (see figure 4.5). It is debatable but according to the Oxford dictionary, browsing generally refers to "reading (parts of a book or books) without any definite plan". Information search refers to "examining or looking carefully in order to find information". Recall refers to "bringing back to mind something or succeeding in remembering". Seeking references refers to "looking or trying to find something". These definitions shall be further refined to take into consideration navigation strategies and cognitive processes associated with these tasks.

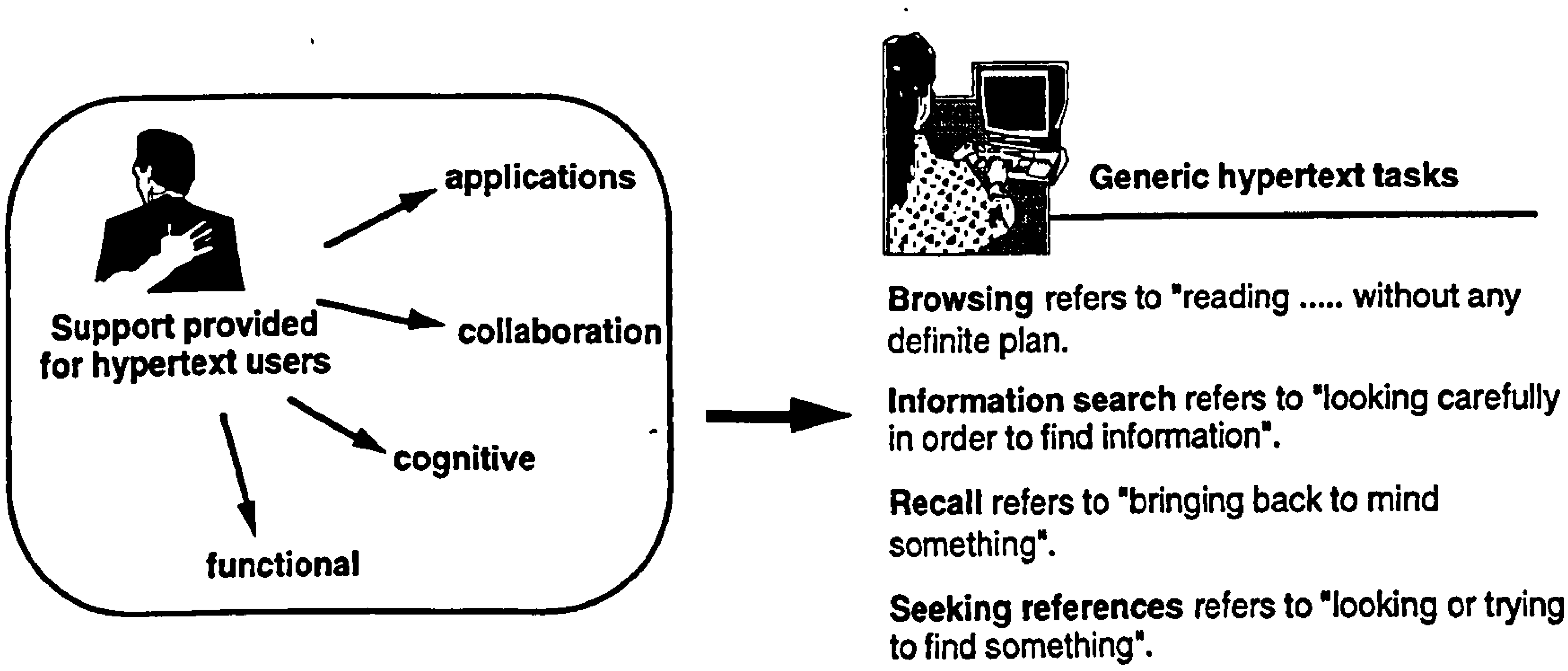


Figure 4.5. Hypertext support and generic hypertext tasks

Since browsing is an activity that hypertext end-users commonly engage in (McAleese, 1993; Rada and Murphy, 1992; Carmel, Crawford and Chen, 1992), this thesis selected the task called *browsing* out of the four to illustrate how cognitive task diagrams can be built. §4.2.2 describes browsing and why it is selected to study end-users' navigation behaviour.

4.2.2 What is browsing?

Browsing is a very common exploration strategy employed by end-users when navigating hypertexts. Many attempts have been made to define browsing, however, there is still some controversy about its definition (Carmel, Crawford and Chen, 1992). To better understand browsing in hypertext, it is important to draw learning points from information scientists' perspectives on text browsing. Although text browsing is often specific and constrained by the context (books, journals, *etc.*), it is a useful benchmark for a more precise definition within hypertext (McAleese, 1993). Some worthwhile sources include: Morse (1973) describing browsing as seeking for new information; Hildreth (1982) suggesting the serendipitous nature of some types of browsing; and Batley (1989) claiming browsing is more purposeful and focused. These definitions from information scientists seem to suggest that there are different types of text browsing involved, which O'Connor (1985) categorises as systematic, purposeful and serendipitous.

In fact, the hypertext research community seems to come to similar conclusions. Marchionini and Shneiderman (1988) define browsing as "an exploratory, information seeking strategy that depends on serendipity .... especially appropriate for ill-defined problems and for exploring new task domains". Cove and Walsh (1988) use a three-stage model to describe browsing: (i) search browsing where the goal is known; (ii) general purpose browsing where the goal is to consult sources that have a likelihood of items of interest; and (iii) serendipity browsing which is purely random. Bates (1989) sees browsing as a semi-structured, goal-oriented activity, distinct from Boolean search.

Consequently, this thesis defines browsing to be unfocused and focused. *Unfocused browsing* refers to reading and navigating in hypertext without a definite or explicit goal for the end-user to accomplish. It is open to debate but this thesis also includes *focused browsing*, where end-users have an idea of what they want to do to accomplish a task. Inspired by the theory of exploratory learning (Polson and Lewis, 1990), focused browsing assumes that end-users perform four steps in sequence to achieve a goal (see figure 4.6). The first step involves setting goal, that is, end-users start with a description of what they want to accomplish – a task. Then, in the next step, end-users may explore the hypertext's interface to discover actions useful in helping them accomplish their current goal. Based on a match between what they are trying to do and the interface's description of actions, end-users then select actions they think will help them accomplish their current goal. In the final step, end-users when deciding what action to do next, will try to understand responses from the hypertext based on the action they have just performed.

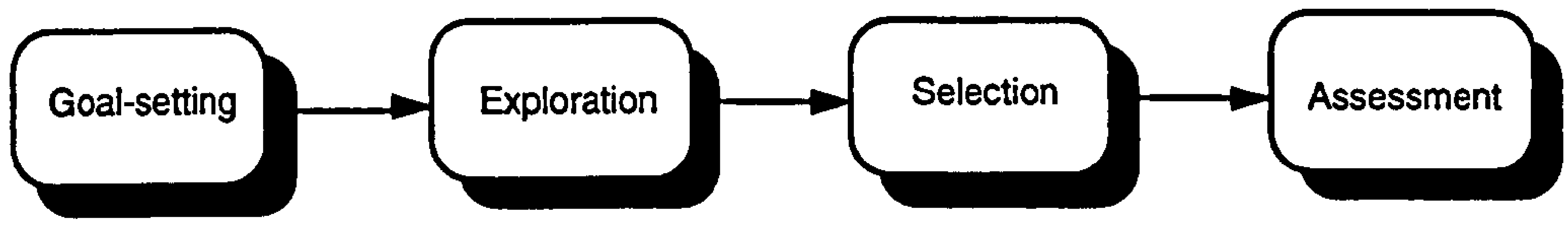


Figure 4.6. Sequence of steps taken to perform focused browsing

Factors that affect browsing

Because of the embedded digression problem and the art museum phenomenon in hypertext (Foss, 1989), supporting browsing in hypertext is not easy. Borrowing insights from information science, Hildreth's definition (1982) in browsing claims that structure imposed on information facilitates browsing. This highlights the importance of summaries and indexes to provide an overview to achieve purposeful browsing (McAleese, 1993). Besides structure, Conklin (1987) added that cueing is important to facilitate effective browsing. Designers should tap on hypertext end-users' experience with conventional linear text, since they bring to their tasks old ideas, for example, typographical cueing such as headings, font, size, *etc.* (McAleese, 1993).

To identify the factors affecting browsing, this thesis refers to Marchionini and Shneiderman's framework (1988) originally defined to identify determinants of success in seeking information. By adapting it to browsing, Carmel, Crawford and Chen (1992) identifies both external and internal factors (see figure 4.7), which are categorised broadly to reflect the concepts of conceptual and mental models. A well-accepted distinction made by Norman (1983) contrasts "a mental model as what the end-user has in mind and that is not observable" from "a conceptual model as given to the end-user by an outside source". External factors affecting browsing include: (i) the setting which refers to the physical environment, time allocated, *etc.*; (ii) the browse system which refers to how the knowledge is structured and accessed; and (iii) the task domain which refers to the subject-area knowledge (conceptual model) encoded by the designer. Internal factors refer to the end-user's mental models and experience with the task, system and browsing skills.

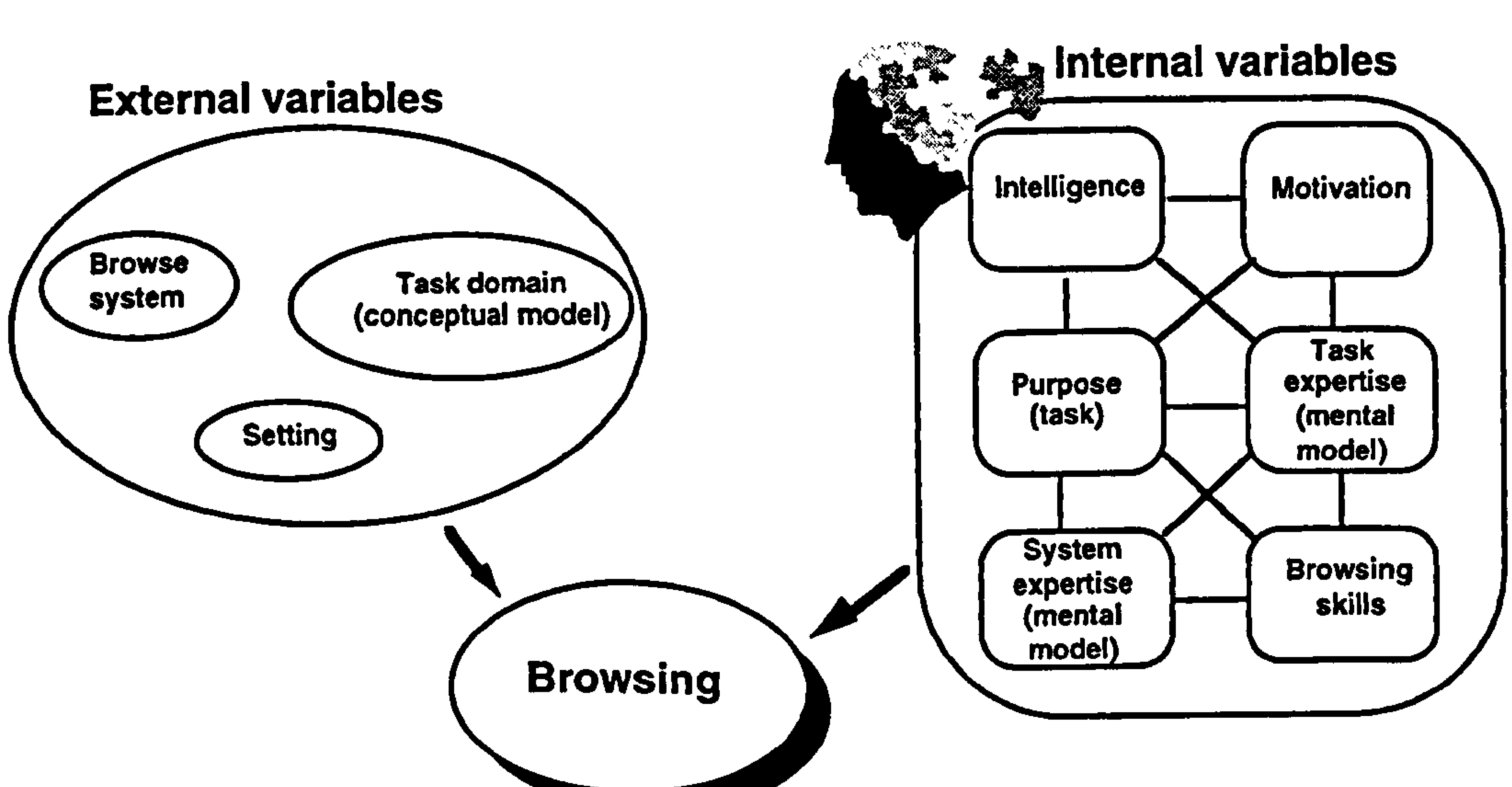


Figure 4.7. Factors affecting browsing

Navigation strategies

Before going further, this thesis notes the distinction between navigation and browsing with regard to hypertext. According to Allinson and Hammond (1988), navigation refers to browsing with the system in control in contrast to browsing with the end-user in control. However, McAleese (1993) sees navigation as involving the use of graphical browsers to move around, but browsing as using the link mechanism of the hypertext elements (for example, cards, windows, nodes). This thesis, however, defines navigation as determining one's position within the overall hyperspace, and/or directing one towards a particular piece of information (Pittas, 1995). Following from this definition, browsing is therefore considered in this thesis as one of the strategies employed by end-users when they navigate round hypertext. As described previously, this thesis considers both unfocused and focused browsing.

Because hypertext presents several different options to readers, individual readers, therefore, must employ different navigation strategies to determine which of them to follow at the time of reading the text. A study conducted by Canter, Rivers and Storrs (1985) to find out the way end-users navigate a database, provided some useful parallels for hypertext with regard to navigation strategies. The study concluded that end-users make use of five navigation strategies when navigating a database (see table 4.1). Except for the task *recall*, these findings provide useful insights in refining the definitions of the other three navigational tasks commonly engaged in by end-users when they navigate round hypertext (as illustrated in figure 4.5). From table 4.1, scanning and exploring are strategies used to get an overview of the system, and can be used to describe the task *information search*. Browsing and wandering can be used to define strategies used in *focused browsing* and *unfocused browsing* respectively. Searching is following an idea until a goal is found or when a dead end is reached, and this strategy can be used to define the task *seeking references*.

Table 4.1. Different navigation strategies and related hypertext tasks

Strategies	Brief description	Hypertext tasks
Scanning	Covers a large area without depth.	*information search* *unfocused browsing*
Browsing	Follows a path until a goal is achieved.	*focused browsing*
Searching	Strives to find an explicit goal.	*seeking references* *focused browsing*
Exploring	Finds out the extent of the information given.	*information search* *unfocused browsing*
Wandering	Does not have a purpose or goal.	*unfocused browsing*

From table 4.1, it is noted that since the tasks "information search" and "seeking references" employ the likely strategies of "unfocused" and "focused" browsing respectively, this thesis decided to group them together. The following section describes how the cognitive task diagrams for browsing are built. (Cognitive task diagrams for recall can be developed similarly. Please refer to appendix D for the task diagrams on recall.).

4.2.3 Building a cognitive task diagram for browsing

Because the concept of a task is central to user-centred design, many TA techniques have been developed that focus on different aspects of tasks, such as the ease of learning a task, the knowledge end-users require in order to accomplish a task, or the task structure. These different aspects generally refer to the cognition, practice or logic of the task (Payne and Green, 1989). Therefore, to break the task browsing into unit tasks, it is an imperative to represent it in terms of how it is usually done (practice), the flow of actions (logic) and the cognitive processes (cognition).

Figure 4.8 shows the steps taken in the task-based approach to build and validate cognitive task diagrams for browsing (step 1). This thesis began by defining the main goals for hypertext browsing. As discussed earlier, this thesis postulates both focused and unfocused browsing. To gather inputs for building the cognitive task diagrams for browsing (step 3), it started with defining the different subtasks associated with browsing (step 2) based on known facts about browsing such as knowledge about hypertext browsing and cognitive strategies for hypertext browsing (as described in §4.2.1 and §4.2.2) and the author's experience with hypertexts.

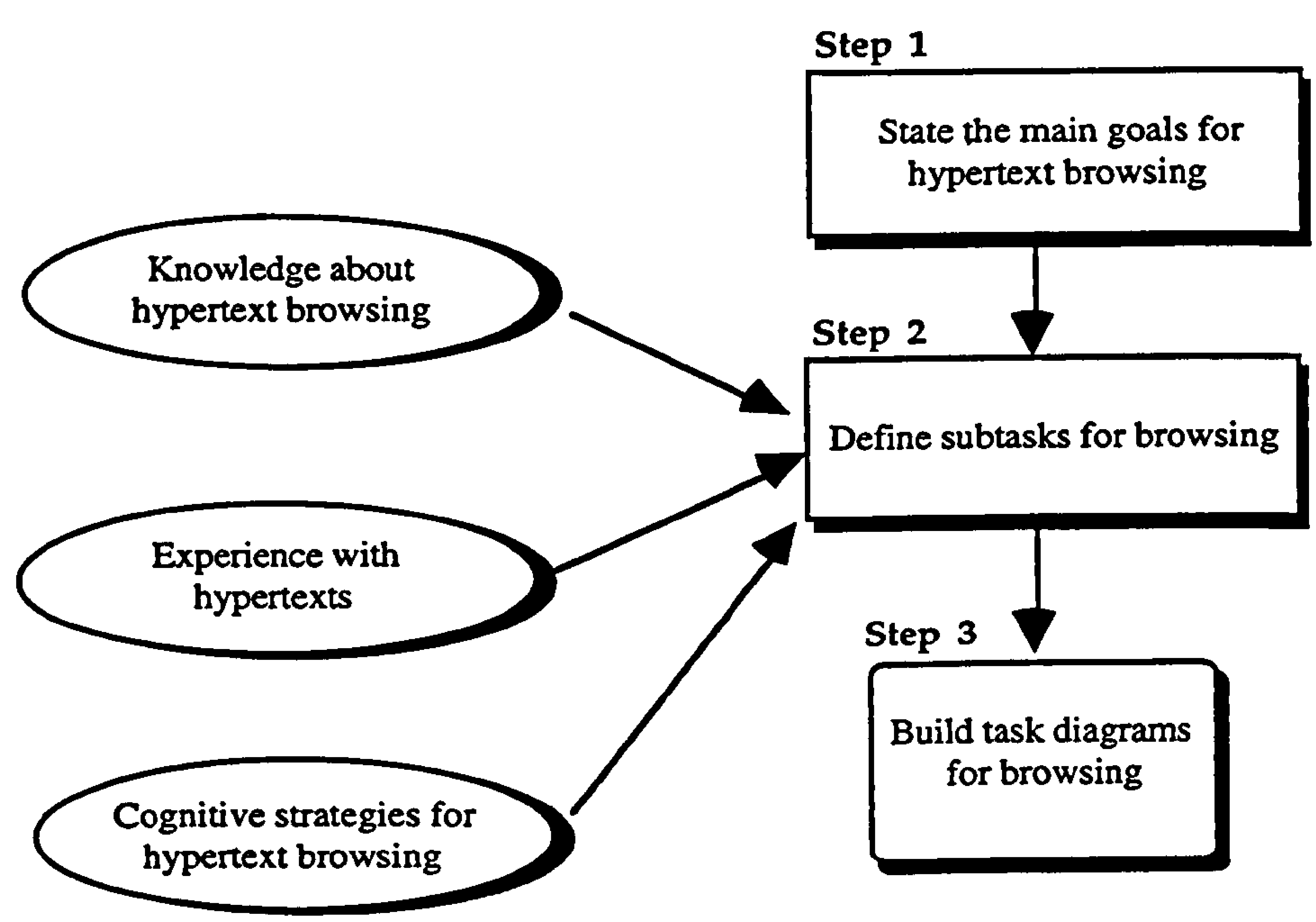


Figure 4.8. A cognitive, task-based approach to build task diagrams for browsing

Because the intention was to model the practice and logic aspects of the task browsing, this thesis adapted the task decomposition (TD) technique (Johnson, 1992) to break down the task browsing into subtasks end-users need to have in order to perform browsing, and then described browsing in terms of a flowchart of actions to achieve goals. Goals are the desired state(s) of the system and actions are the lowest level units of behaviour. A rough cut of the task diagrams for browsing was made. Figure 4.9 shows the task diagram for browsing and the possible subtasks that are associated to it: (1) getting general information; (2) understanding the system; (3) finding a particular topic; (4) exploring; and (5) finding out what topics or areas are covered in the system.

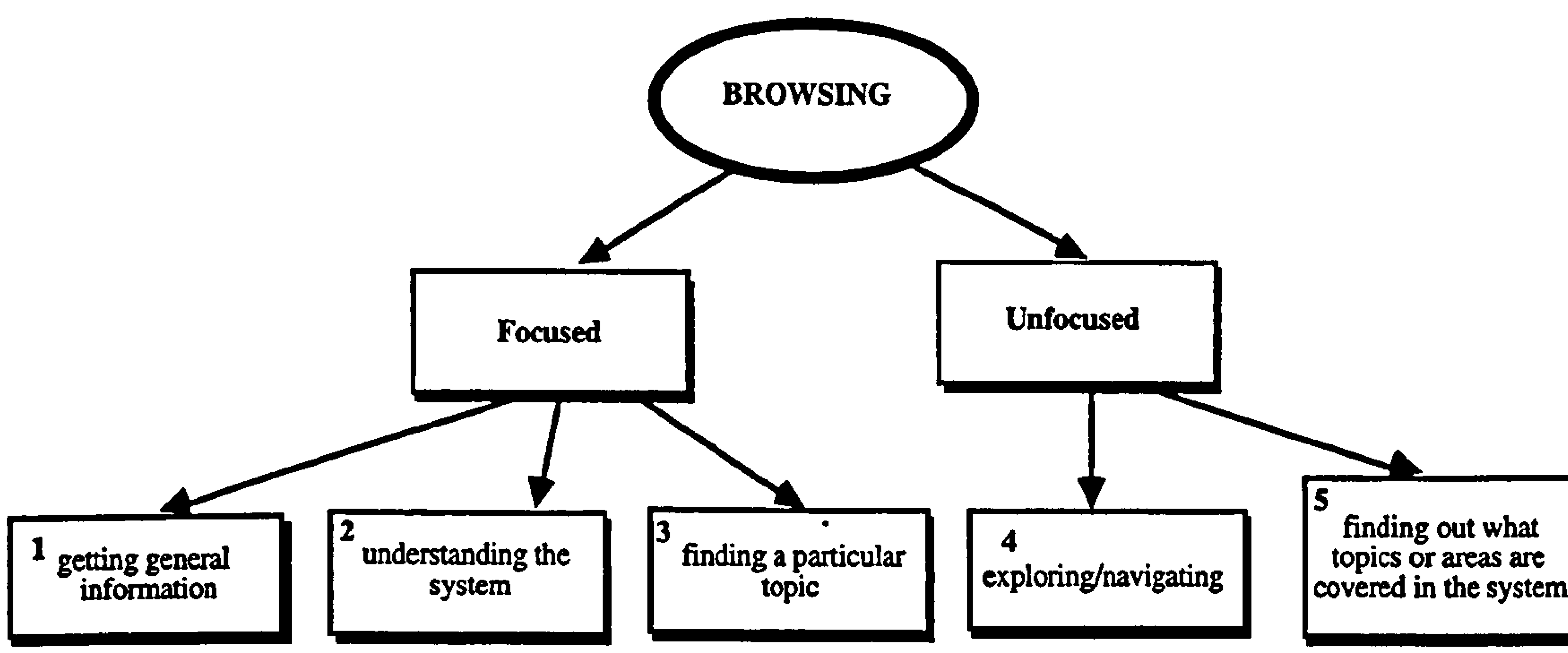


Figure 4.9. Task diagram for browsing

These subtasks can be classified into two categories namely, focused browsing and unfocused browsing. For focused browsing, in subtasks 1, 2 and 3, end-users have a goal in mind when executing them. In contrast, end-users performing unfocused browsing (for example, subtasks 4 and 5) do not have a particular goal to accomplish. Each subtask is then further broken down into simpler, executable subtasks or actions. Although these subtasks are interrelated, the order in which they are performed depends on several parameters: end-users' domain knowledge of hypertexts; end-users' domain knowledge of the contents covered in a particular hypertext; end-users' intention or goal of using the system; and the context of end-users' tasks based on situated actions described in (Suchman, 1987). Unlike TD and related techniques that tend to focus on what actually happens, this thesis also built into the task diagrams subtasks that describe what *should* happen. This is an important inclusion since browsing is dynamic and therefore, hypertexts supporting browsing should be adaptive as well as predictive.

Focused browsing

Figure 4.10 shows the task diagram for getting general information (subtask 1). End-users can obtain general information through various ways either by trying out buttons option, reading paper manuals, or asking someone. To execute the subtask "looking for About ...", the action end-user needs to take is to click onto the "About" button. To simplify the task diagram, the subtask "trying out buttons" option is represented as a shadowed box to indicate that it can be further broken down into many simpler, executable subtasks (see figure 4.11). The subtask "trying out buttons" identifies all possible buttons normally found in hypertext such as "find", "help", "about", "next", "back", "previous", "home", "exit", "table of contents" and "map". Because browsing in hypertext is a cognitive activity, it is also essential that two aspects of cognition that are important to end-users browsing hypertext should be considered: reasoning and learning. Established TA techniques that are concerned with describing some aspects of the cognitive characteristics of end-users' tasks include the MHP, GOMS, CCT, ICS, PUMs — there are many proposals (as described in §4.2.1). These cognitive TA techniques, however, are difficult to use and implement (Preece, Rogers, Sharp, Benyon, Holland and Carey, 1994; Anderson *et al.*, 1990; Wilson *et al.*, 1986). Instead, because reasoning and learning are two important aspects of cognition that need to be captured, the cognitive task diagrams are inspired by Polson and Lewis' Cognitive Walkthrough (Wharton, Rieman, Lewis and Polson, 1994). This thesis made the assumption that learning is achieved when end-users are able to understand, induct, and subsequently, build links and concepts, as illustrated by the dotted links in figure 4.11 for clicking the "find" button, for instance. Hence, when an end-user clicks on a button, he learns what it does.

To obtain general information (see figure 4.10), end-users can also look for the "help" option. To execute the subtask "looking for help option", the action end-users need to take is to click onto the "help" button. Although the subtasks "reading paper manuals" and "asking someone" may not have any direct influence on the design of the interface, they are included in the task diagram to remind designers that these are also possible end-user actions.

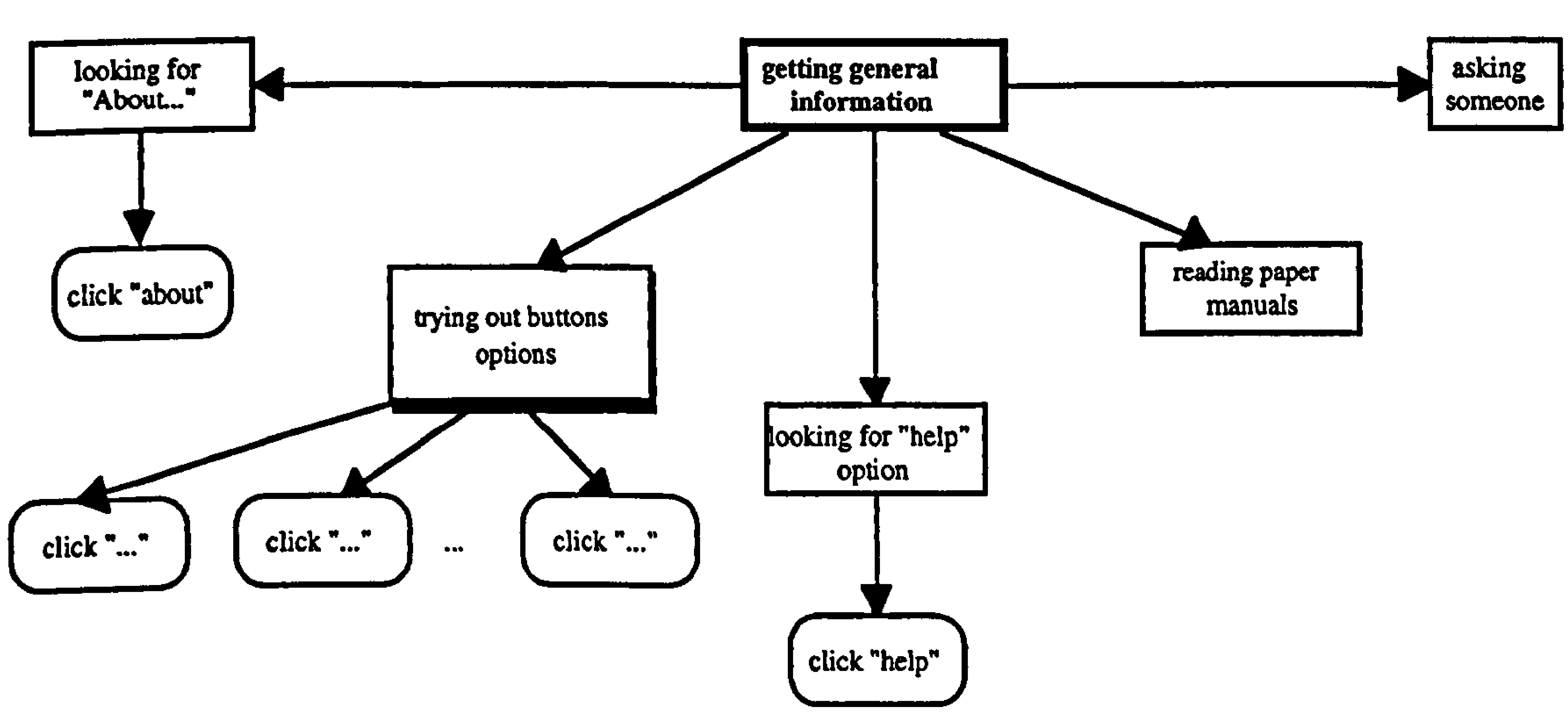


Figure 4.10. Task diagram for subtask 1: Getting general information

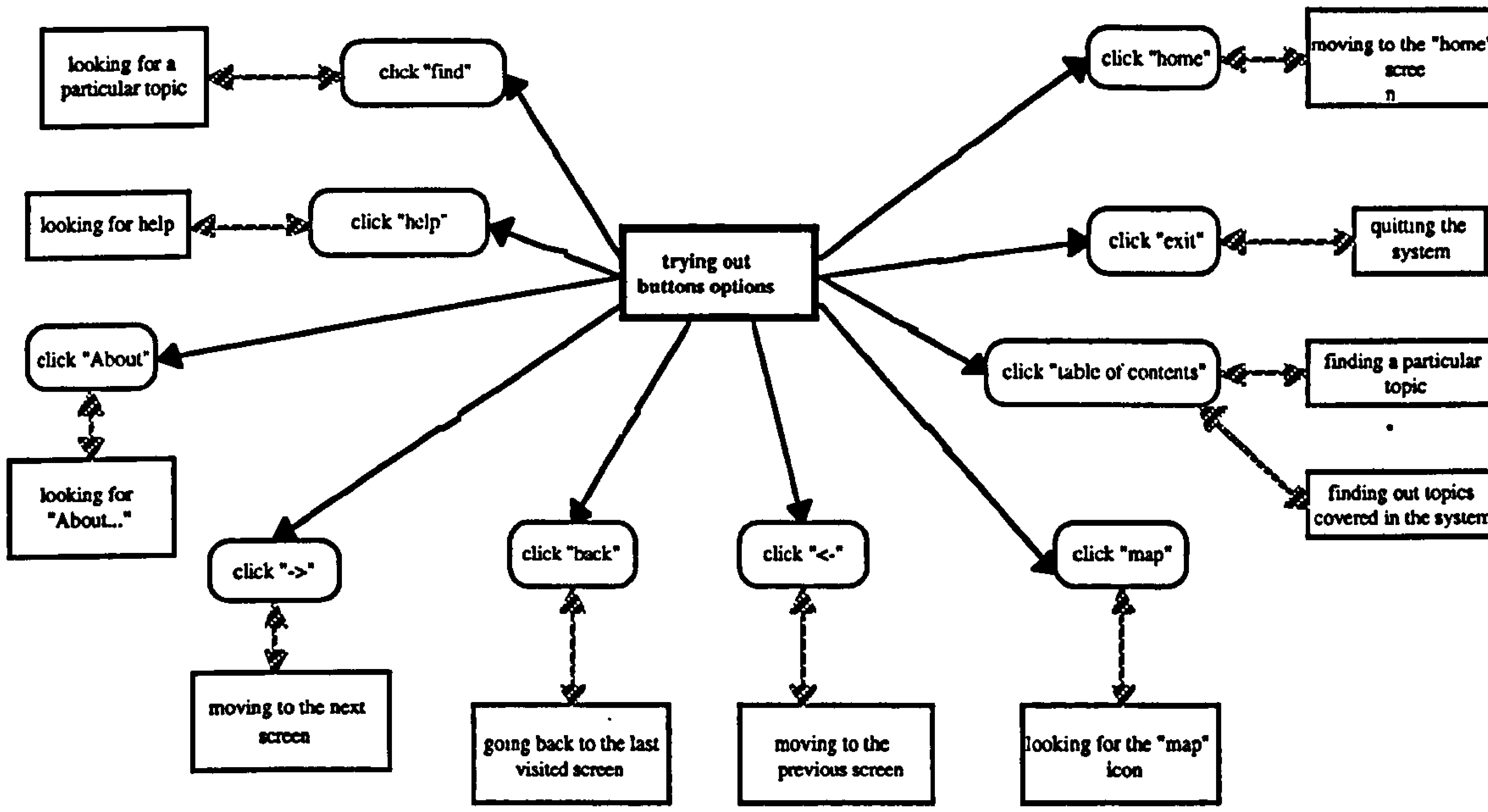


Figure 4.11. Task diagram for subtask: Trying out buttons options

To understand the system (subtask 2), end-users can choose to perform any of these further subtasks: clicking highlighted word; trying out buttons option; getting to know the metaphors used; and reading paper manuals (see figure 4.12). Executing the subtask "click on highlighted word" results in the action "moving to another page". End-users learn that there is a relationship between "clicking highlighted word" and "moving to another page related to the highlighted word", which is indicated by a dotted arrow in the diagram. As described in chapter 3, metaphors (for example, travel and book metaphors) have been introduced in hypertexts to make learning easier. Understanding the system also involves understanding the metaphors used, and the icons selected to represent them. In executing the subtask "trying out buttons options", learning can be accomplished as end-users understand the icons used to represent the buttons, and induct the metaphors used.

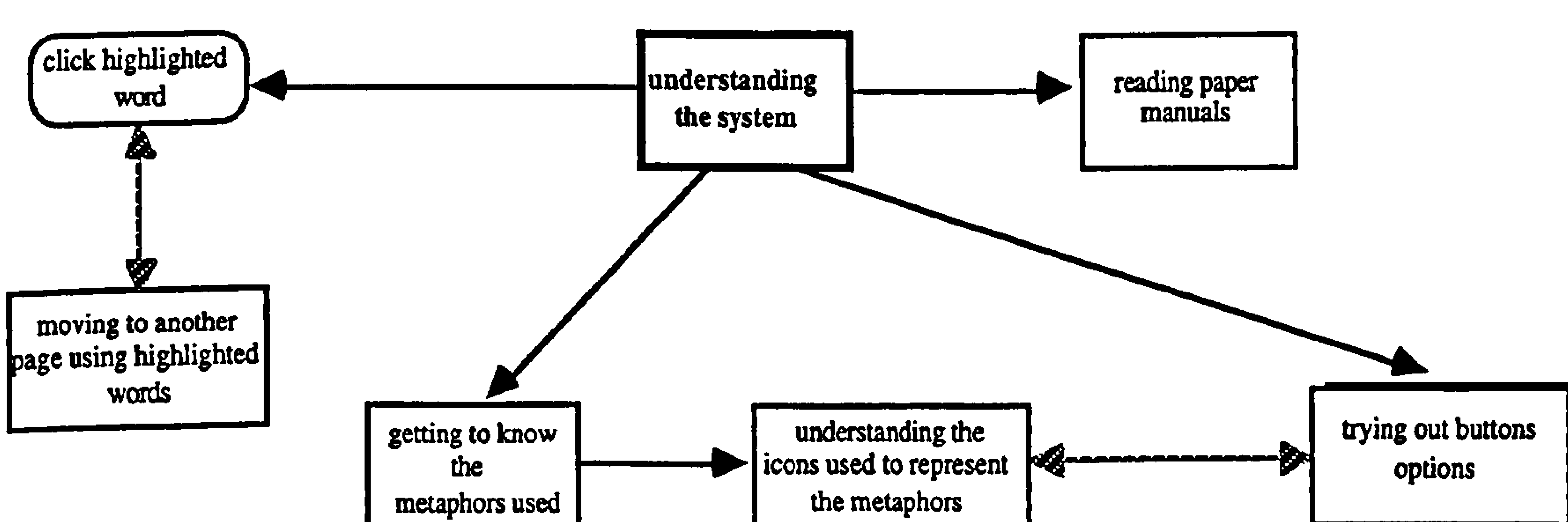


Figure 4.12. Task diagram for subtask 2: Understanding the system

Figure 4.13 shows another example of focused browsing task. The subtask "finding a particular topic" (subtask 3) can be further broken down into simpler subtasks "looking for table of contents" and "using the search/query mechanism". When the "table of contents" is found, end-user clicks the most relevant topic. If the goal of "finding a particular topic" is not satisfied, the end-user repeats the process of clicking the "table of contents". To do that, he has two options. He can either click the "table of contents" again if it is found on the current page, or else he can move back to the last visited page using the "back" button where the "table of contents" is found. If the goal is satisfied, the end-user is faced with two goals: quitting the system or activating another goal. The subtask "using the search/query mechanism" is executed by clicking the "find" button. The end-user types in the query and submits it. If the list of topics displayed by the system is not relevant, the end-user then starts the search/query mechanism again. If the list is relevant, the end-user then selects and clicks the most relevant topic to the query. If the goal is satisfied, the end-user can either choose to quit the system or activate another goal. If the goal is not satisfied or the end-user wants to refine the search, the end-user repeats the process using the search/query mechanism.

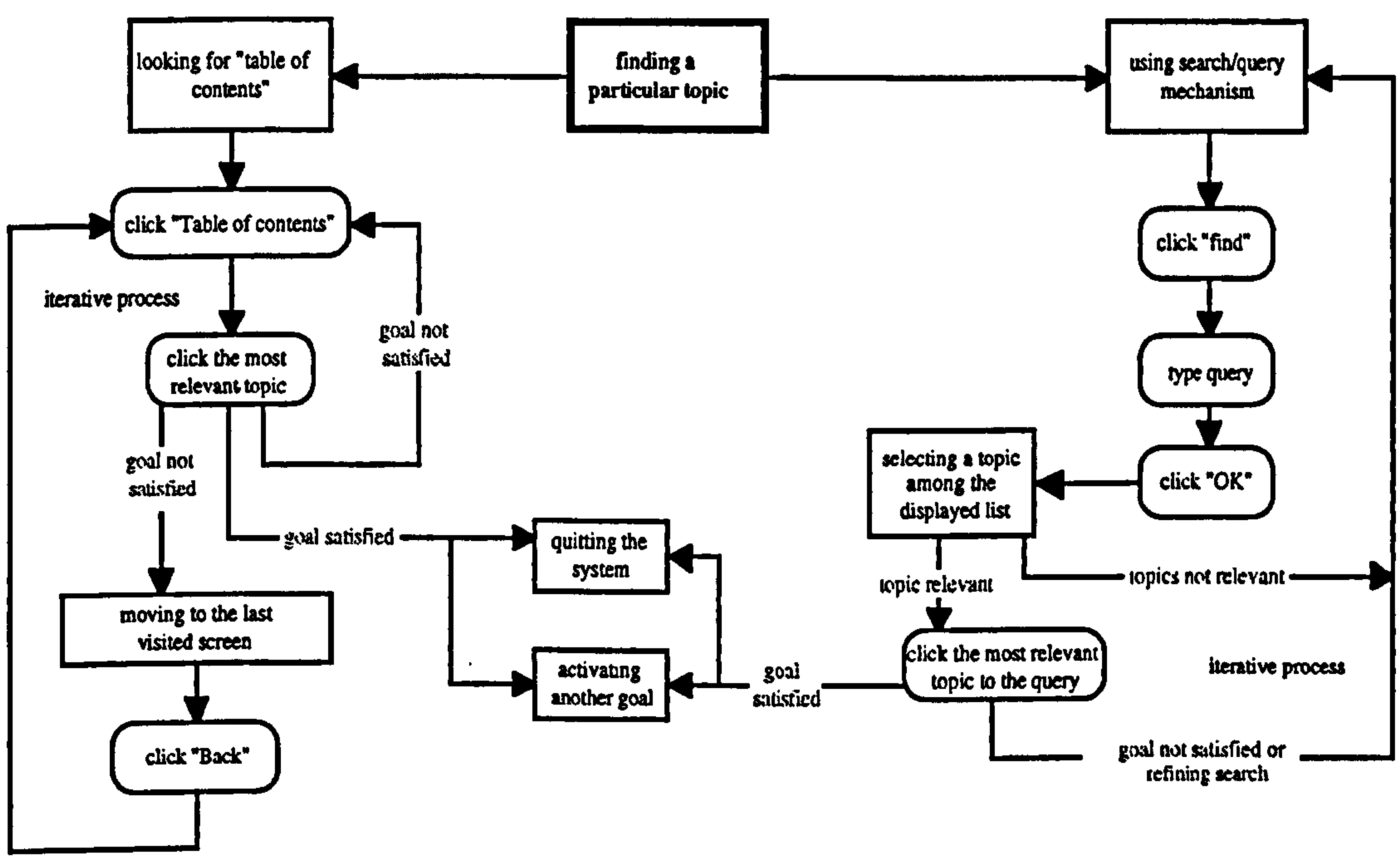


Figure 4.13. Task diagram for subtask 3: Finding a particular topic

Unfocused browsing

Figures 4.14 and 4.15 show the task diagrams for exploring/navigating (subtask 4) and finding out what topics or areas are covered in the system (subtask 5) respectively. These two subtasks are unfocused, since end-users performing them do not have any particular goal in mind to accomplish. The subtask "exploring/navigating" can be achieved by performing any of these simpler subtasks: looking for "help"; going back to the last visited screen; moving to another screen using highlighted words; moving to the next screen; moving to the previous screen; looking for the "table of contents"; and moving to the "home" screen. To execute these subtasks, the end-users click on the respective buttons (see figure 4.14).

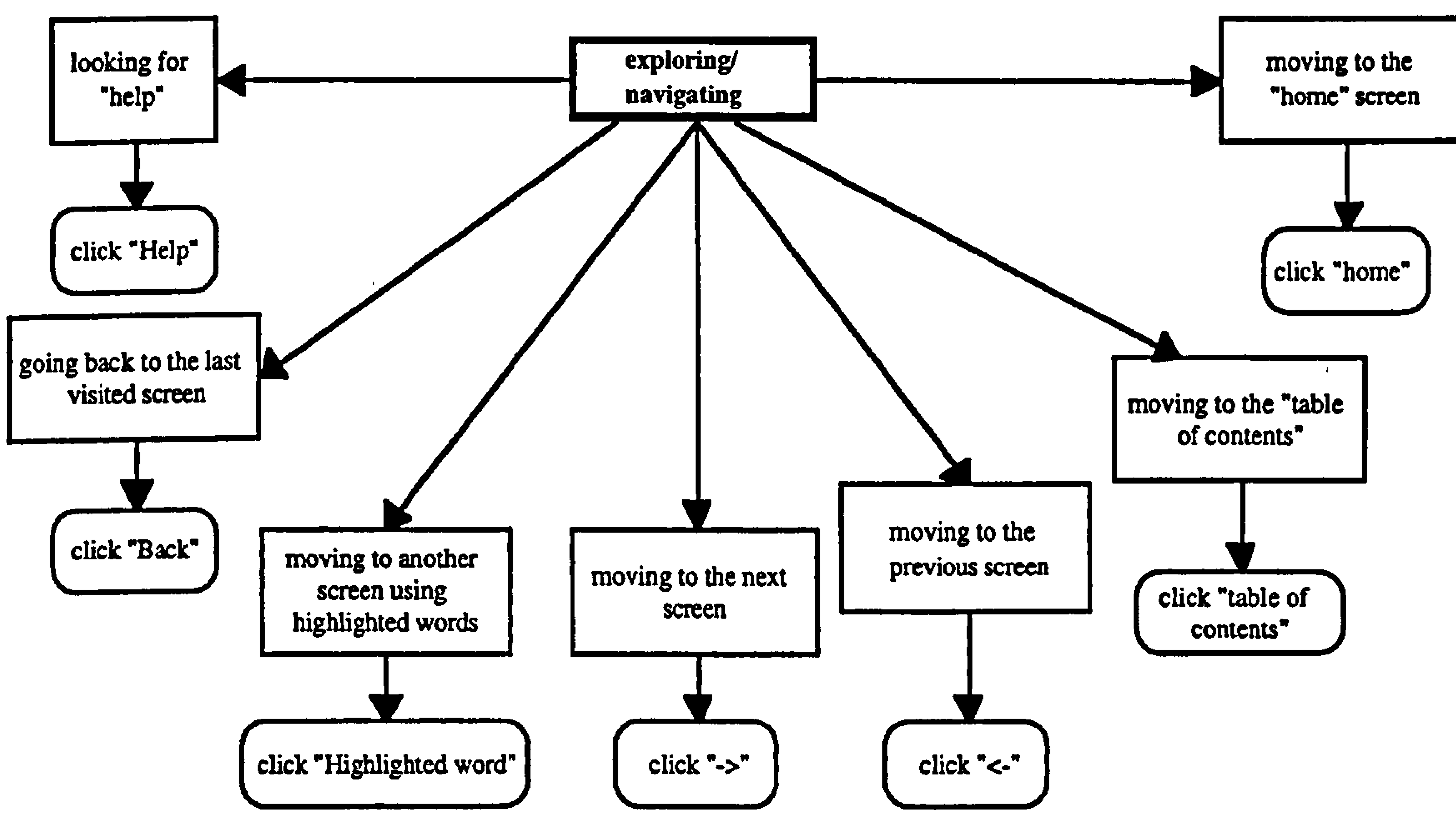


Figure 4.14. Task diagram for subtask 4: Exploring

Finally, the subtask "finding out what topics or areas are covered in the system" involves the execution of these further subtasks: looking for index/table of contents; looking for the "home" screen; exploring/navigating; and using the search/query mechanism. As explained previously, the end-user actions in executing these subtasks include the clicking of the relevant buttons, such as clicking the "table of contents" for the subtask looking for the index/table of contents. Activating the subtasks "exploring/navigating" and "using search/query mechanism" is as explained before.

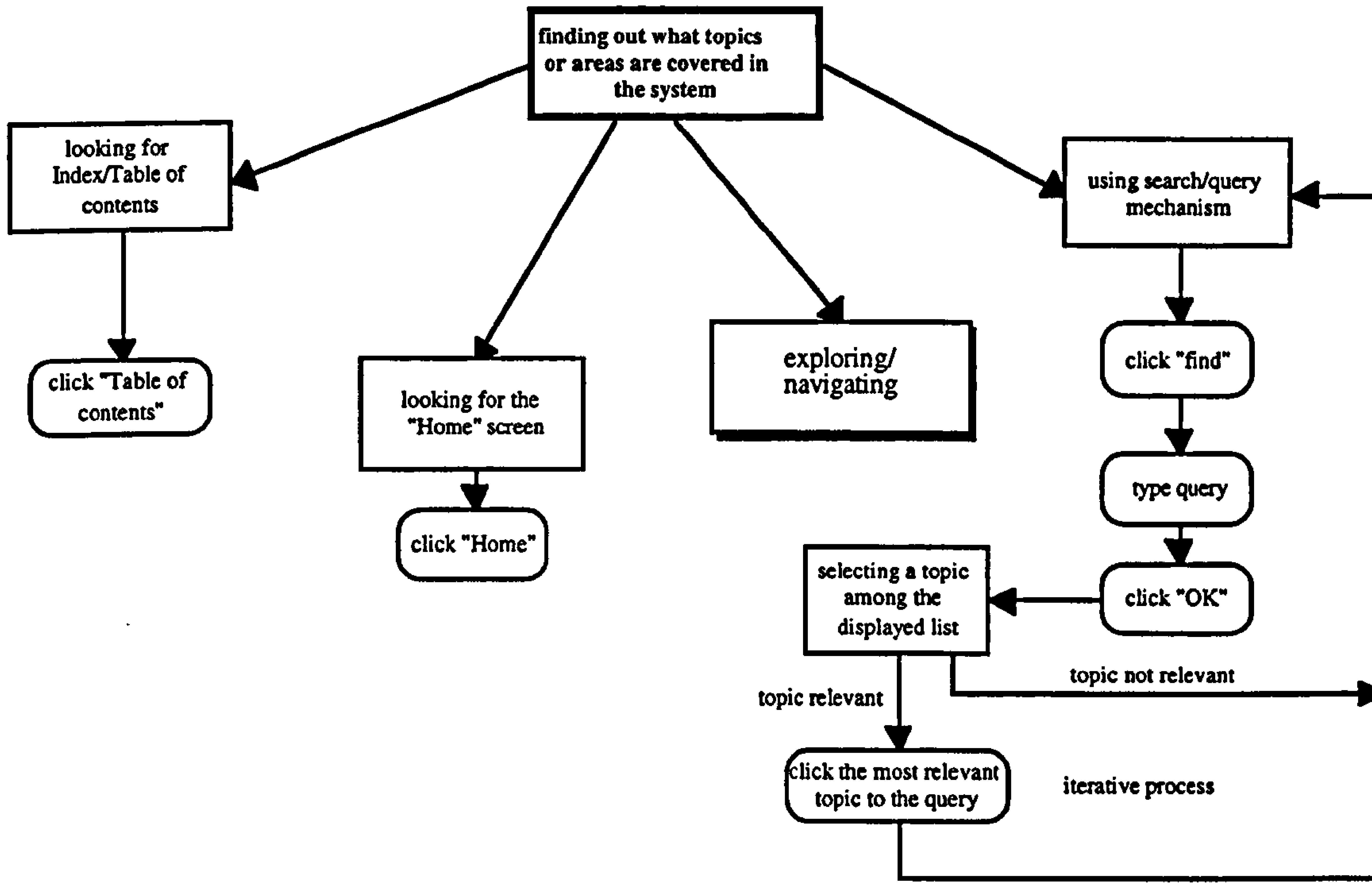


Figure 4.15. Task diagram for subtask 5: Finding out what topics or areas are covered in the system

4.2.4 Contributions to design

There are two aspects to TA (Johnson, 1992): gathering information and analysing end-users' tasks; and the generation of task models and task scenarios into running prototypes, in order that the outputs of the user-task analysis support a mapping and checking operation between end-user tasks and proposed design (see figure 4.16).

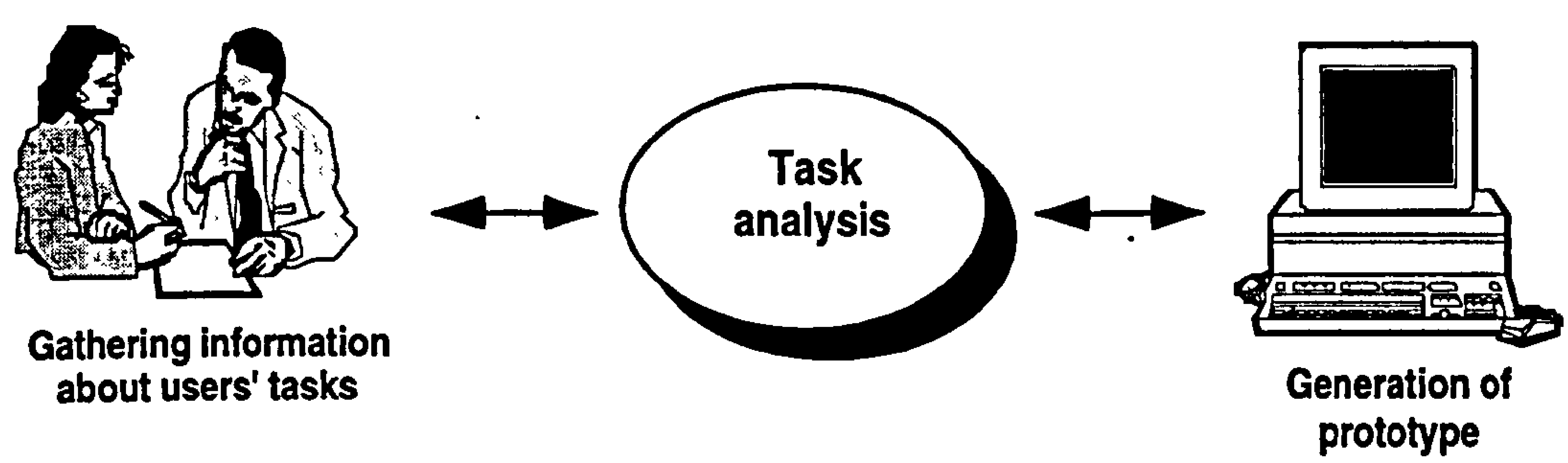


Figure 4.16. Two aspects of task analysis

Inspired by Johnson (1992) and Carmel, Crawford and Chen (1992), this thesis used the information obtained from the cognitive task diagrams for browsing, in conjunction with design principles and guidelines, to develop a design model, a prototype hypertext. Design models can, of course, be represented in a variety of other forms: formal specifications, informal specifications, screen drawings, storyboards and simulations.

This thesis describes below how the outputs from the cognitive task diagrams for each subtask in browsing were mapped to design recommendations for hypertexts. These recommendations provide designers with useful insights into: how best to support and improve the various aspects of the task browsing; how to support the way end-users could carry out actions while browsing hypertext; and identify crucial information that should be provided on the screen to support effective browsing.

Subtask 1: *Getting general information* (see figure 4.10). To help end-users get general information, the design of the hypertext should ensure that useful information buttons such as "about" and "help" are provided. Navigation buttons are necessary to enable end-users to navigate round hypertext like "next", "previous", "return" and "exit". "Find" button should be available to allow end-users to search for specific information easily. Provisions should be made to enable access from every page within the hypertext to "landmark pages" such as "home" and "table of contents".

Subtask 2: *Understanding the system* (see figure 4.12). To help end-users better understand hypertext, it is important that hypertext links are clearly and visibly displayed. If metaphors are used, it is crucial that end-users do not get conflicting message and that the choice on the types of buttons should reflect that accurately.

Subtask 3: *Finding a particular topic* (see figure 4.13). To enable end-users to search for a particular topic, an effective "query/search mechanism" should be provided. In addition, end-users should also be provided with a "table of contents" to give an overview of the topics covered as well as to provide links to relevant topics covered in the hypertext.

Subtask 4: *Exploring/navigating* (see figure 4.14). To find out the extent of the information covered in the hypertext, end-users would explore the features provided by actioning the relevant buttons or highlighted word. This calls for a consistent presentation of the features so that end-users need not be unnecessarily confused or "lost".

Subtask 5: *Finding out what topics or areas are covered in the system* (see figure 4.15). Not only should the presentation be consistent, it should also provide useful features such as "table of contents", "home" and "search/query mechanism" to aid end-users to find out the extent of coverage of information in hypertext.

In summary, the recommendations made to hypertext design to support the task browsing are listed in table 4.2:

Table 4.2. Design recommendations for hypertext

R1.	To encourage browsing and to make it interesting, hypertext design should provide recommendations on relevant or related topics to hypertext end-users (from subtasks 3 – 5).
R2.	Design hypertext to facilitate recognition, that is, the interface design should be consistent (from subtasks 4 & 5).
R3.	Inform end-users about the organisation and structure of the hypertext using "table of contents" or "map" (from subtasks 1 – 5), so that they understand the fundamental relationships between nodes and links. The "table of contents" and "map" should be designed to provide end-users with conceptual model at appropriate levels of abstraction and classification.
R4.	Links should be built from every page to "landmark pages" such as "home", "table of contents" (from subtask 1). This is to ensure that end-users can safely return any time to a known topic, thus encouraging exploration without disorienting the end-users.
R5.	Information and navigation buttons should be available in each page to encourage browsing (from subtasks 1 – 5).
R6.	Provide a "help" feature to assist novice end-users in carrying out the task browsing by suggesting the possible actions they can perform (from subtask 1).

In fact, these recommendations are not surprising, as they are congruent with the design principles postulated in §3.3.2. From these recommendations, it can be concluded that there are two aspects of design to which the cognitive task diagrams can contribute namely; directly to design interface, and indirectly to functionality (see figure 4.17). *Design interface* is concerned with the visual appearance and screen layout of the user interface. Though cognitive task diagrams may not give sufficient detail for complete screen layout or interaction style, they provide an initial input and guide for user-interface design. This includes the design of pop-up menus, windows, icons, buttons, text fields, command names. Secondly, *functionality* is concerned with the performance and capability of the system. This includes navigational buttons, "table of contents", "home" page, building links for learning, *etc.*

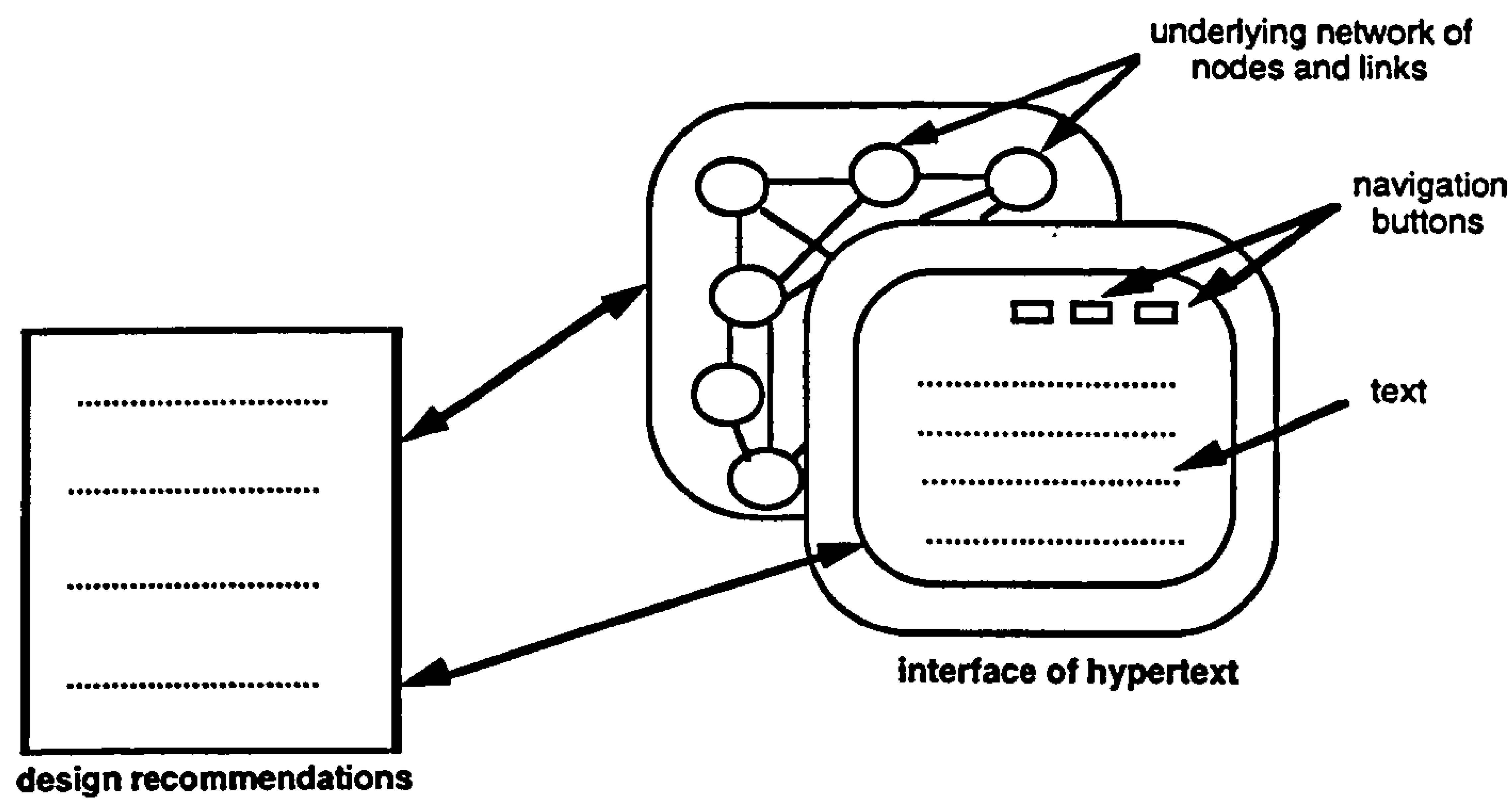


Figure 4.17. Mapping of recommendations to design interface and functionality

Except for R1 and R6, the principled recommendations R2 to R5 had been implemented in the prototype built in §3.3.3, which is referred to as prototype1, using good design principles. Recommendation R1 was not implemented in prototype1 since the purpose was to build it quickly, and implementing R1 would require investigations into adaptive properties (for example, Kaplan and Fenwick, 1993), which is outside the scope of this thesis. Prototype1 was modified based on the principled recommendation R6 derived from the investigation on cognitive task diagrams. The modified prototype2 had 51 cards. Figure 4.18 shows the four new cards which were incorporated into prototype1 to reflect the recommendations to provide end-users with tips on how to go about performing the task browsing. The third card of the prototype stack is a card titled "welcome". End-users can click to indicate the task they would like to know more about. When the button "task" is selected, a "pop-up" window appears to allow end-users to choose either "browsing", "search" or "revision". "Search" was used in the prototype to refer to "focused" browsing since end-users may be more familiar with the terminology. However, the tips on browsing also include both "focused" and "unfocused" browsing as recommended by the cognitive task diagrams.

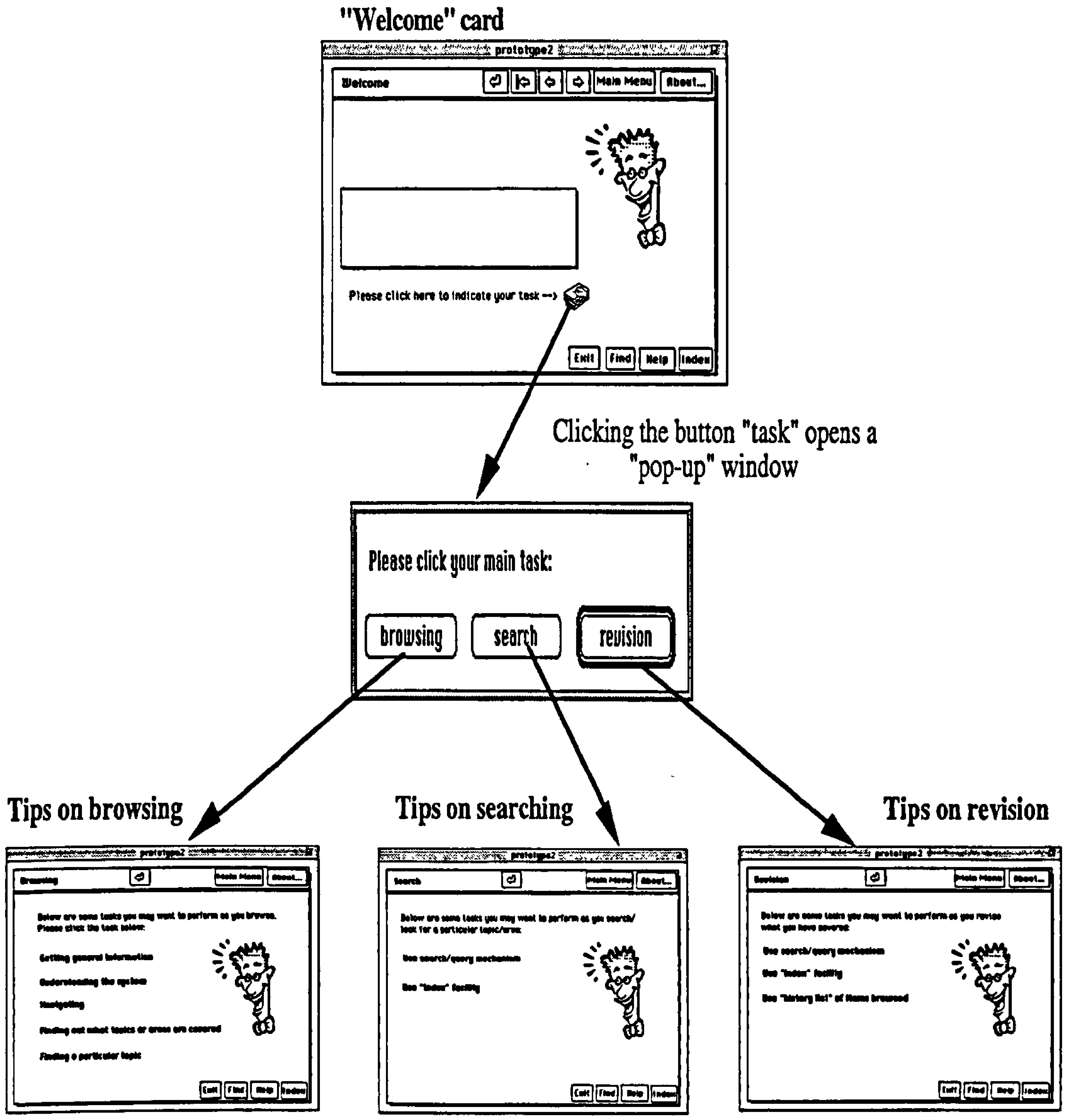


Figure 4.18. Four new cards introduced in prototype2

For example, if the end-users select the task "browsing", a respective card on tips on browsing appears to provide information on how to go about performing the task browsing (see figure 4.19). If end-users want to perform one of the subtasks such as "getting general information", a pop-up menu is shown with suggestions on the various actions end-users can perform. Currently, these tips in prototype2 are meant purely for information purposes. A further enhancement to prototype2 is to make it adaptive to end-users' needs (R1) by not only providing end-users with browsing tips, but guiding them in navigating the prototype based on the browsing tips.

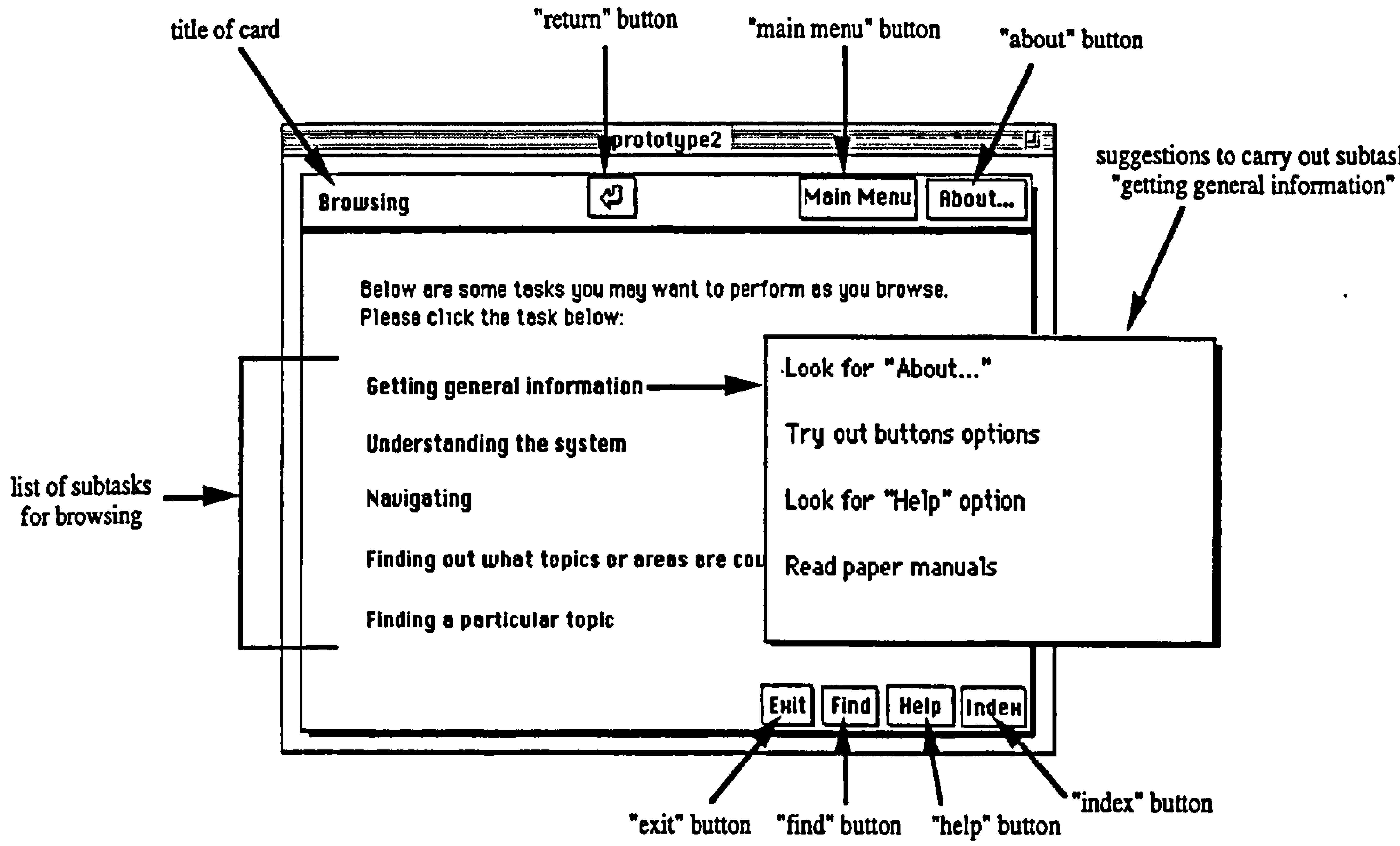


Figure 4.19. Tips on browsing in prototype2

Qualitative results and impressions

To get some qualitative results on whether prototype2 is comparatively better than prototype1 (as described previously in §3.3.3) using the within-subjects' design (Harris, 1996). The three subjects involved in the evaluation of prototype1 in §3.5.2 were asked to evaluate prototype2. These three subjects had some experience using hypertexts and they would be able to provide a more critical view on the design of the prototypes, compared with someone who has not used any hypertext before. They were asked to complete some tasks as outlined in a worksheet (see appendix C2). In recording their answers in the worksheets, they were also asked to write down the steps taken to obtain the answers. They were then asked to complete a questionnaire (see appendix B2). They were then interviewed to comment in general on the performance of the prototypes, whether they found the prototypes effective in helping them to complete their tasks.

There was no significant difference in terms of their performance when using the prototypes as they were able to complete their tasks successfully in both cases. However, the three subjects indicated that they were more satisfied with prototype2 with the incorporated design recommendations such as quick tips for browsing, which included recommendations for possible activities normally associated with browsing. They also rated it to be more effective, and as such they were more confident using it. These findings, though not entirely surprising, confirmed that end-users' needs can be straightforward if they are systematically analysed using cognitive task diagrams.

The subjects were asked on how they would like to improve the design of prototype2. They proposed making these changes:

- making the interface even simpler by having less buttons
- explaining how to use the "index" card
- making it less cumbersome to obtain tips on browsing, search and revision
- introducing more colours and graphics to make it more interesting

Some of these ideas were feasible and implementable using HyperCard version 2.01, however there were others that would not be possible due to the constraints imposed by the software and hardware used. For example, adding colours is not supported by HyperCard version 2.01. Prototype2 was modified based on the feasible suggestions given, taking into consideration the constraints of HyperCard version 2.01. An improved prototype3 was then built. The index card was modified by adding the instruction "Click to go to any topic" (see figure 4.20) to make the card more "usable". The "welcome" card was removed since the subjects felt that it was too cumbersome to have to register to find out more about how to carry out tasks such as browsing, search or recall. Instead, a new button called "quick tips" was introduced. Clicking the "quick tips" button will invoke a "pop-up" window to allow the end-users to select the appropriate tasks to get further tips (see figure 4.20). It is to be noted that fewer navigational buttons were used in prototype3 since the subjects commented that they were confused over the use of "go to first page" button and "go to last visited page" button. The latter button was kept since the subjects were more familiar with the concept of a "return" button.

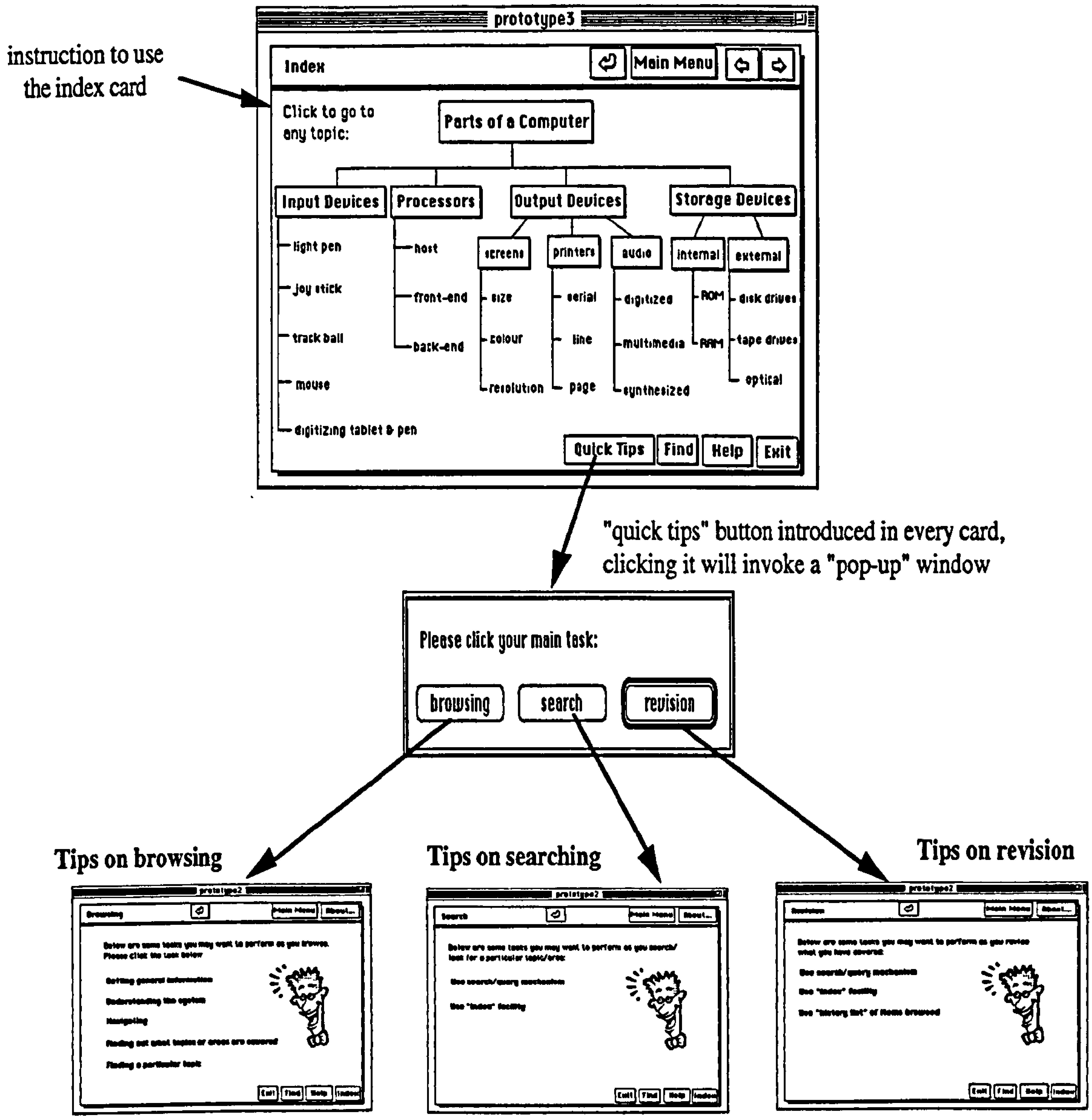


Figure 4.20. Improved "index" card and the associated cards relating to the "quick tips" button

The three subjects were then asked to comment in general on the modified prototype3. All three indicated they were more satisfied with prototype3. One of the subjects (the undergraduate in computing) was asked to evaluate prototype3 by completing some tasks and writing down the steps taken to accomplish these tasks. The choice of the subject was random since the intention was to find out the cognitive steps taken by an end-user when using the three prototypes to complete the tasks. The findings will be reported later in §4.3.4 on how the undergraduate's feedback was used to explore the possibility of using non-human users to evaluate the design of hypertext.

Discussion

It appears that this approach is cost effective: commencing with an expert's initial task diagram, then using design recommendations for hypertexts. Clearly, one might scale up this procedure for use in a production environment, but the purpose for this thesis was to obtain a structure suitable for the user models to be employed in hypertext design. Whether the additional cost of obtaining better task diagrams would in fact be worthwhile is a separate issue; this thesis suspects that there would be diminishing returns, especially when the delay is considered. Nielsen (1993) argues that although better results can be obtained by applying more careful methodologies, they are much more expensive in terms of cost and required expertise. He rather focuses on achieving the "good" doing *some* usability, even though the methods needed to achieve this result may not always be the "best" and will not necessarily give perfect results. Voltaire (1764) philosophised that the "best is the enemy of the good", and there is certainly much truth in that insisting on using only the best methods may result in using no methods at all.

Preliminary results were encouraging in that the design recommendations derived from the cognitive task diagrams were positively regarded by the subjects.

4.3 Evaluation methods and techniques

Evaluation is concerned with "gathering information about the usability or potential usability of a system to improve features within an interface and its supporting material or to assess a completed interface" (Preece *et al.*, 1993). Dix, Finlay, Abowd and Beale (1993) see evaluation as fulfilling these goals: (i) to assess system's functionality in supporting end-users to perform required tasks; (ii) to assess system's usability in terms of ease of use and end-users' attitudes; and (iii) to identify any design problem with the system.

According to Lingaard (1994), some people who do not have a background in experimental design tend to believe that usability evaluation is a waste of time. They argue that it does not provide useful information, that recommendations arising from usability evaluations cannot be accommodated within the project, and evaluation is what designers do just before the product is released in the market. Some of these accusations are justified only when testing is delayed and takes place too late in the development process to accommodate any findings that may require software modifications.

However, Landauer (1995) points out that it is not good enough to design an interactive system without subjecting it to some form of evaluation, because it is impossible to design an optimal user interface in the first try. Dix *et al.* (1993) argue that even if one has used the best methodology and model in the design of usable interactive system, one still needs to assess the design and test the system to ensure that it behaves as expected and meets end-users' requirements. Nielsen's (1993) advice with respect to interface evaluation is that designers should simply conduct some form of testing.

Evaluation should be carried out throughout the development lifecycle of an interactive system (Landauer, 1995; Lamming and Newman, 1995, Lindgaard, 1994; *etc.*). Dix *et al.*'s (1993) definition on the role of evaluation "to assess designs and test usability and functionality of a system" suggests that different evaluation methods would be needed to carry out evaluation. Figure 4.21 summarises the two main categories of evaluation; one done before the implementation stage known as *formative* evaluation and one done after the implementation stage known as *summative* evaluation (Preece *et al.*, 1993; Nielsen, 1993; Howard and Murray, 1987; Hewett, 1986; *etc.*). Formative evaluation is performed to improve the interface as part of an iterative design process, and summative evaluation aims to find out how good the system is. This distinction is important because the type of evaluation to be conducted and the evaluative methods applied largely determine the issues to be investigated (Lingaard, 1994; Dix *et al.*, 1993). Preece *et al.* (1993) provides a good, concise comparison on these methods.

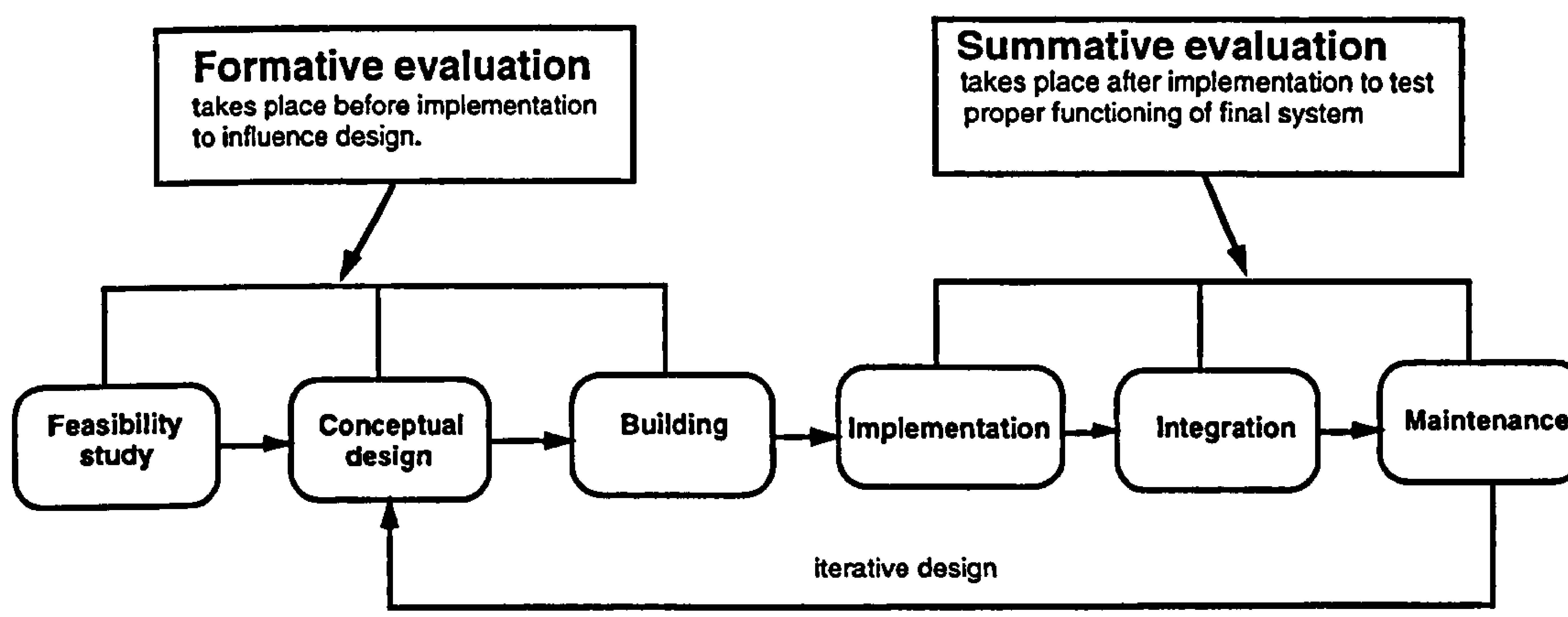


Figure 4.21. Stages in development lifecycle in which evaluations occur

This chapter focuses only on established formative evaluation methods and techniques adapted for hypertext and not summative evaluation, since this thesis emphasises the importance of getting design right "up front" so that many design blunders can be avoided before implementing the real system. These formative methods and techniques used before the implementation stage, are discussed broadly under real user testing and non-user testing (see figure 4.22), congruent with the classification proposed by Whitefield, Wilson and Dowell (1991). However, these distinctions are not fixed, and differ from other classifications found in Dix *et al.* (1993), for example. This thesis also discusses how some of these formative methods and techniques can be effectively employed by designers to measure the usability and usefulness of hypertexts. Because usefulness and usability are concepts that are fuzzy and often confused, this thesis begins by defining these concepts in §4.3.1. Real user and non-user testing methods are discussed in §4.3.2 and §4.3.3, and, though important are not exhaustive. Since these methods are covered in standard literature, this thesis will not deal with them in depth, except for the use of executable user models. For a more comprehensive discussion on evaluation methods and techniques, one can refer to (Lingaard, 1994; Dix *et al.*, 1993; Preece *et al.*, 1993; Nielsen, 1993; *etc.*).

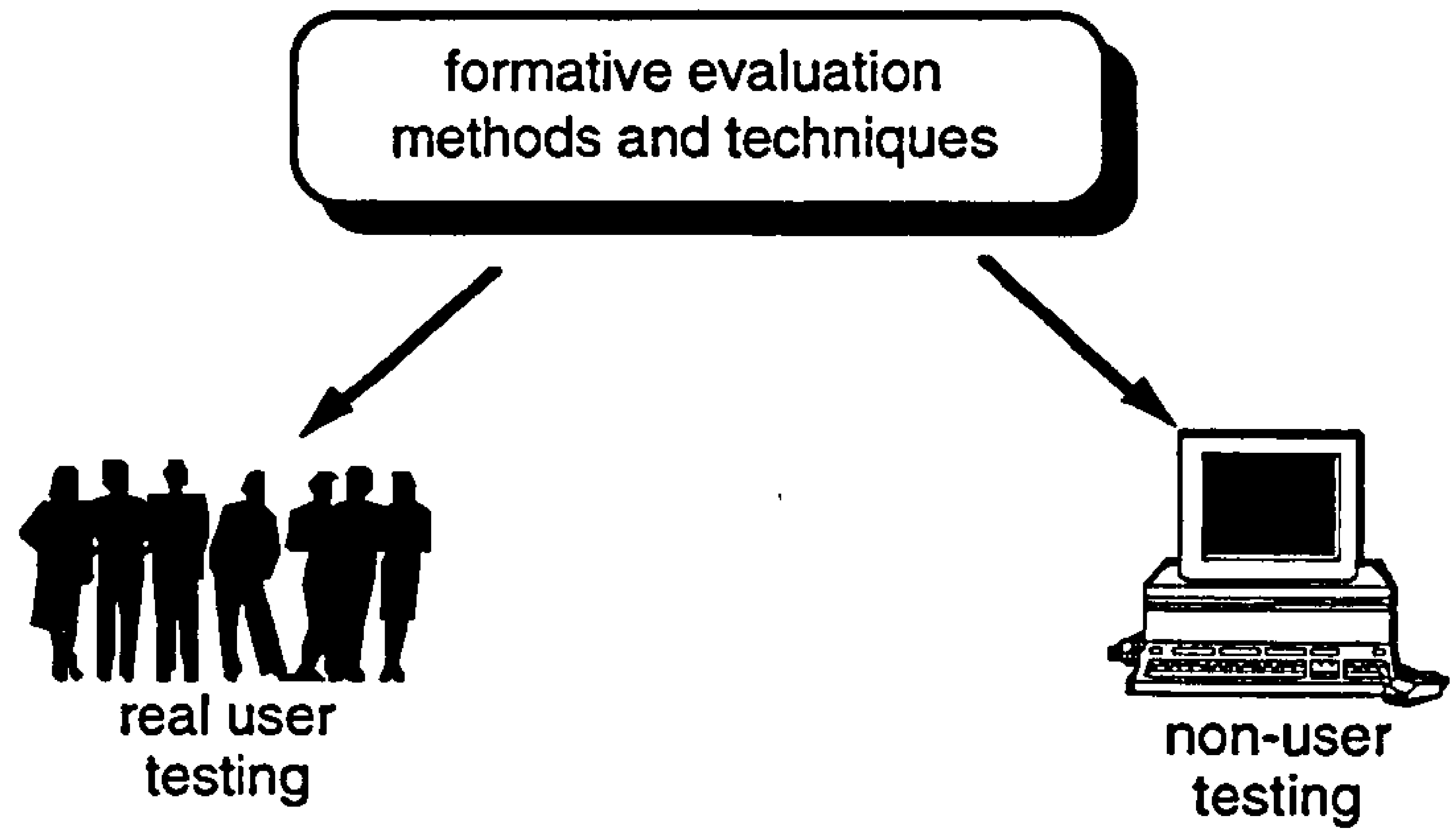


Figure 4.22. Evaluation methods and techniques

4.3.1 Usability and usefulness

The division between usefulness and usability is blurred as evidenced in the discussion below of differing views held by researchers. The majority of usability studies focus on end-user performance in terms of learnability, throughput and end-users' attitudes because they are the easiest characteristics to measure (Preece *et al.*, 1993; Shackel, 1991; *etc.*). However, Landauer (1995) argues that although these measurements, commonly referred to as usability metrics, are intended to measure the ease of use of a system, in reality it is difficult to pinpoint if the failure is due to it being "hard to use because it isn't useful?" or "is it useless because it is hard to use?".

Booth (1989) suggests that because end-users only use systems if they match their goals, the measurement of how good a system is, should reflect both its usability and usefulness. Lingaard (1994) strongly disagrees that unlike usability, it is impossible to quantify the usefulness of a system, in terms of the extent of coverage of end-users' tasks supported by the system.

This thesis makes a distinction between usability and usefulness. In §3.4, this thesis proposed how the usability of hypertext using established usability dimensions can be measured, covering these categories of usability defects: screen design; terminology and system information; system capabilities and user control; navigation; and completing tasks. As for usefulness, this thesis agrees with Lingaard (1994) that its attainment is measured in reference to system specifications and not on end-user performance testing.

4.3.2 Real user testing

Although there are techniques that do not employ real users evaluating and refining system, they should not be a replacement for actual usability testing (Landauer, 1995; Dix *et al.*, 1993). Real user testing is strongly encouraged in that it provides insights into the mindset and working methods of real users for whom the system is designed (Nielsen and Landauer, 1993). Dix *et al.* (1993) suggest employing real user testing in evaluating systems after the implementation stage. Johnson (1992) advocates that real users need not be used only to evaluate designs, which is seen by many designers as the sole reason, but also to generate design recommendations. This thesis also proposes that real user testing should be carried out even before the hypertext is ready for implementation so that qualitative results and impressions can be obtained. These methods include observations, surveys and experiments (see figure 4.23).

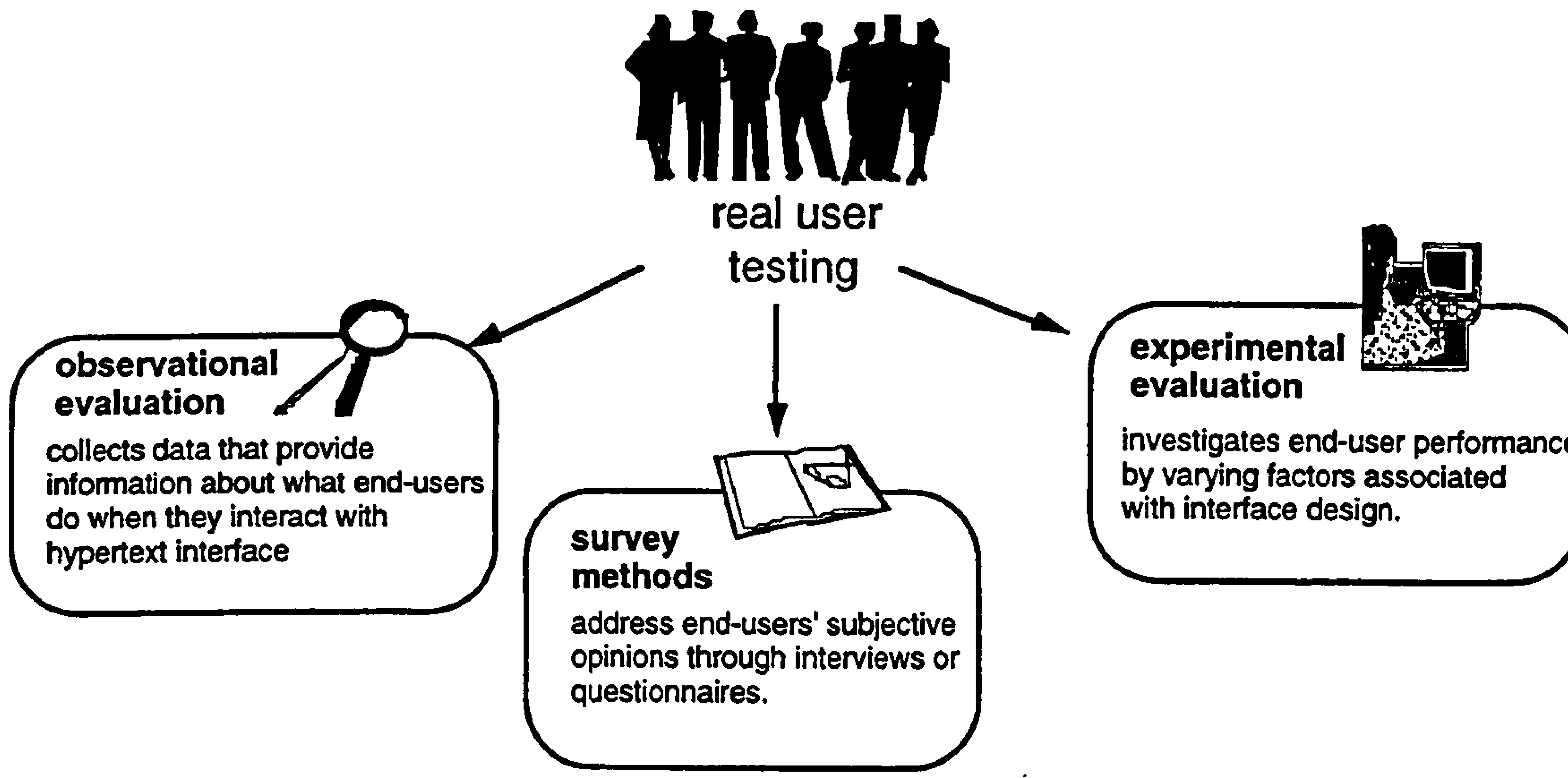


Figure 4.23. Different evaluation methods involving real users

Observational evaluation

This form of evaluation involves collecting data that provide information about what end-users do when they interact with an interface (Preece *et al.*, 1993). They include direct observation (Diaper, 1989), video protocol (Wright and Monk, 1989), software logging (Preece *et al.*, 1993) and thinking-aloud (Lewis, 1982). The results obtained from these evaluations can either be qualitative producing non-numerical data for impressions, or quantitative producing numerical data for statistical implications. It is to be noted that the use of the different observational methods is the trade-off between time spent and depth of analysis.

Chapter 7 will describe how the server transaction log files (a form of software logging) have been used to analyse end-users' browsing behaviour on the web. To validate and refine the cognitive task diagrams for browsing, chapter 9 will describe a preliminary investigation into how video protocol was used to record and analyse end-user activity.

Survey evaluation

Survey evaluation is used to address end-users' subjective opinions through interviews or questionnaires (Preece *et al.*, 1993). However, Nielsen (1993) cautions that since questionnaires and interviews only gather end-users' opinions about the user interface, and not to study the user interface itself, one cannot always take user statements at face value (for example, Root and Draper, 1983; Karis and Zeigler, 1989; *etc.*).

This thesis acknowledges the limitations of questionnaires and interviews. Since the intention of the investigations carried out in this thesis is to get quick results and impressions, it would be reasonable to use both these techniques. §3.5 made use of flexible interviewing techniques to clarify interviewees' replies and personal attitudes. §3.5 and §6.2 made use of both open and closed questions to elicit end-users' responses.

Experimental evaluation

Experiments are used to investigate end-user performance by varying factors associated with interface design (Harris, 1986). To ensure the overall reliability of an experiment, the following factors must be carefully considered (Dix *et al.*, 1993): (i) subjects chosen; (ii) variables tested; and (iii) hypothesis tested.

From chapters 3 – 8, this thesis describes a series of laboratory experiments devised to carry out investigations to test the hypotheses. These experiments are pilot experiments to help us study how to improve the interface as part of the iterative process. A small sample size is used to test the hypotheses since the intention of this thesis is only to obtain qualitative results and impressions.

4.3.3 Non-user testing

One would certainly expect the best results from testing real users and real systems, but doing so may not always be feasible. Non-user testing methods are encouraged as a means to perform evaluation early enough to influence design while it can still change direction (Nielsen, 1993). Analytic and heuristic evaluation methods, and executable user models are ways of evaluating without requiring the attendance of real users. Non-user testing methods require mostly inputs from experts, and they are attractive because they are cost-effective and less time-consuming compared with using real users. However, the disadvantage is that because expert users form a minority, they may not be representative in terms of learning or error-prone behaviour (Preece *et al.*, 1993).

Analytic evaluation

This form of evaluation starts early in the design cycle. It enables designers to analyse and predict expert performance of error-free tasks in terms of physical and cognitive operations that must be carried out. Although many micro TA methods concerned with cognitive modelling have been developed (see §4.2.1), not many have been applied to real applications because they are time-consuming, and require specialist psychological knowledge. Only the keystroke-level analysis has been widely applied (Preece *et al.*, 1993; Card *et al.*, 1979).

This thesis will discuss more of this analytic evaluation in chapter 5 to help designers analyse and better understand the structure of hypertext. In chapter 7, this thesis describes how these ideas are implemented into an authoring tool.

Expert evaluation

In expert evaluation, experts are asked to take the role of less experienced end-users and describe the potential problems they foresee arising from such end-users (Preece *et al.*, 1993). Expert evaluation that is guided by general "rules of thumb" is called heuristic evaluation (Nielsen and Molich, 1990). Nielsen (1993) gives a list of heuristics for evaluating user interface designs, such as speak the end-users' language, minimise end-users' mental load, *etc.* Lingaard (1994) strongly recommends using heuristic evaluation as a trouble-shooting tool to achieve quick but relatively coarse usability results. The advantage heuristic evaluation has over other methods is that it tends to uncover many potential end-user problems in a relatively short period of time with comparatively little effort.

Chapter 8 describes how expert evaluation guided by a set of heuristics (adapted from the list of essential and helpful authoring features compiled in chapter 6) is used to evaluate a hypertext authoring tool.

Cognitive user models as design aids

Executable user models are software agents that simulate real end-users' behaviour, as well as predict end-users' performance. They can embed multi-disciplinary knowledge that most designers and most end-users would not be expected to know or be able to verbalise in their accounts of interaction. They are able to do more exhaustive checking of prototypes, long before they have reached a stage where actual human-user interaction would be practicable. Because user models are reusable, it is possible to simulate rapidly large groups of end-users and obtain useful statistical information. The idea in using executable models achieves a two-fold purpose in rapidly iterating the design process and avoiding many design blunders. However, the reliability and efficiency of the executable user models is very much dependent on the cognitive theories used to generate them (Barnard and May, 1993). It is, therefore, important that the results obtained from simulating executable user models be sufficiently tested for them to be reliable (Wilson and Clarke, 1993).

To reduce the use of extensive and time-consuming real users validating and refining the task diagrams, a preliminary and innovative investigation was carried out with Cécile Rigny, to explore the use of executable user models in the development of well-structured hypertexts. CUM-DesTool (Cognitive User Model Design Tool), a tool for constructing executable models of end-users, has been built by Rigny as part of her PhD work. CUM-DesTool's structure refers to ACT and the memory model proposed by Reason (1988) and is intended to enable a rapid and easy implementation of end-users' models. CUM-DesTool is a general simplified architecture to build operational cognitive user models which enable designers to understand, predict and simulate end-users' behaviour and performance.

Compared with the theoretical, cognitive modelling models which are difficult and time-consuming to use, CUM-DesTool claims to enable a rapid and easy implementation of user models (Rigny, 1997). CUM-DesTool is still in its early stages of development, but initial results are encouraging in applying cognitive user models as design aids. In §4.2, we described how we built the task diagrams for browsing, beginning with the definition of goals for browsing. Consequently, subtasks associated to browsing were identified. By incorporating into [1]CUM-DesTool knowledge about browsing as described in §4.2.3 as well as applying appropriate cognitive strategies for hypertext browsing, the task diagrams were automated leading to the generation of a user model for browsing in hypertext.

Figure 4.24 shows the general interface of the executable cognitive user model for browsing generated by CUM-DesTool. The interface displays in alphabetical order the list of concepts corresponding to browsing subtasks. Clicking onto the subtask browsing, for example, triggers the corresponding module of the simulation of the end-user's reasoning and activity in achieving this goal. An automated trace of the reasoning mechanisms generated by CUM-DesTool can then be used to help designers understand end-users' behaviour. In the following section, examples of trace files generated by CUM-DesTool to achieve a particular subtask *finding a topic* are described. By inputting into CUM-DesTool parameters such as end-users' profiles and tasks, different cognitive user models can be created.

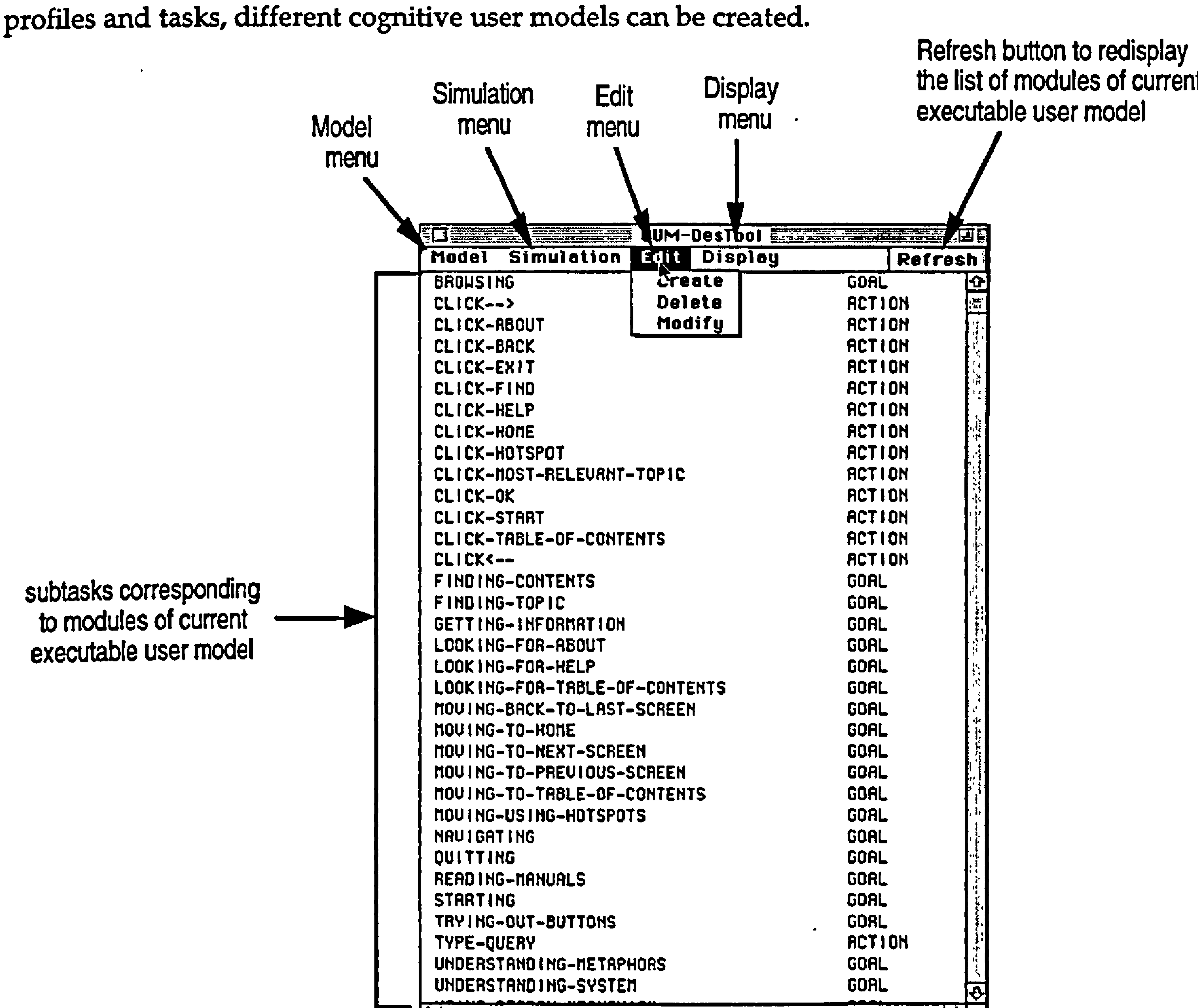


Figure 4.24. General interface of the executable user model for browsing generated by CUM-DesTool

[1]The interpretation and inputting of the contents of the task diagrams into CUM-DesTool was carried out by the author of CUM-DesTool, Cécile Rigny. The author's contribution is to provide the knowledge of browsing in the form of task diagrams as described previously.

4.3.4 Real user testing versus executable user testing

To find out whether executable user models compared to real users can be used to evaluate hypertext design, a preliminary investigation was carried out to compare the behaviour of real users with the behaviour of the executable user model for browsing. Since it is not the intention of this thesis to validate the accuracy of the executable user model for browsing generated by CUM-DesTool, we will make use of the feedback of only one human user from the investigation in §4.2.4, to explore the idea of using executable user models or non-human users in usability testing. Basically, we want to examine the cognitive steps taken by the human user to complete the tasks of finding a specific information using the three hypertext prototypes: (i) original, prototype1 (described previously in §3.3.3); (ii) modified, prototype2 based on design guidelines and knowledge about browsing (discussed previously in §4.2.4); (iii) improved, prototype3 based on a human user's feedback (mentioned previously in §4.2.4). From the investigation carried out in §4.2.4, we obtained the human user's record of the steps taken to perform the tasks using the three prototypes. Using the executable user model for browsing, we carried out a hand-simulation on CUM-DesTool of the human user's behaviour to evaluate and analyse the usability of the three prototypes, by performing similar tasks the human user had to perform in the investigation described in §4.2.4. A sample trace file generated by CUM-DesTool as a result of simulating the human user's behaviour when using prototype1 is provided in appendix I.

Table 4.3 compares the steps taken by the human user and the executable user model in obtaining the answer to a particular task, for example, finding a topic. From table 4.3, it is to be noted that the executable user model suggested steps that matched the way the human user would complete the task *finding the topic*. The goal was satisfied when the "find" button was found. Prototype3, which was modified based on human user's feedback, had the "find" button removed from the first two screens. Therefore, it took us a number of cognitive steps to perform the task completely, when we carried out the hand-simulation. Whether the human user was right in suggesting that the "find" button was to be removed is debatable. What is interesting is that the executable user model randomly provided a list of suggestions that the human user would have in mind when performing the task *finding a topic*. This list of suggestions could provide valuable insights for designers to ensure the usability of the prototype. Taking away the "find" button may make one human user happy but may also make other human users unhappy, if so many cognitive steps have to be taken to perform the task successfully.

The design of the prototype3 hypertext was based on the recommendations of one "typical" human user. However, it may be presumptuous to assume that all end-users behave like her. This thesis acknowledges that good practice calls for the use of several test end-users (Nielsen and Landauer, 1993). This thesis argues that since its intention is to get initial impressions about the possibility of using this form of evaluation testing for hypertext, it suffices then that only one "typical" human user was used. Although CUM-DesTool is in its early stages of development, initial results are encouraging. We see immediate benefits to the designers: validated executable user models generated by CUM-DesTool (or for that matter any design tool for the generation of executable user models), can be used to generate different user models, rapidly iterate and avoid many design blunders, thus reducing the use of extensive and time-consuming human users validating and refining them.

Table 4.3. Steps taken by a human user and executable user model in obtaining the answer to a particular task, here, finding a topic

Goal to perform: Finding a particular topic	Steps taken by human user	Steps suggested by executable user model (obtained from the trace and simplified for presenting below)	Remarks
Prototype1 with design principles	Look for "Find" button. Click "Find" button. Type in query. Select the relevant topic.	Browsing =>Finding a topic Finding a topic => Click-Find Click-Find => Type-Query Type-Query =>Select-topic	In the hand-simulation, the steps suggested by the executable user model matched the actions taken by the human user. This is because the "Find" button was found on the first screen of the prototype.
Prototype2 with the design principles and knowledge about browsing	Look for "Find" button. Click "Find" button. Type in query. Select the relevant topic.	Browsing =>Finding a topic Finding a topic => Click-Find Click-Find => Type-Query Type-Query =>Select-topic	In the hand-simulation, the steps suggested by the executable user model also matched the actions taken by the human user. This is because the "Find" button was found on the first screen of the prototype.
Prototype3 with the design the design guidelines and knowledge about browsing (based on user's feedback)	Look for "Find" button. Can't find button. Click on the next screen button. Look for "Find" button. Can't find button. Click on the next screen button. Look for "Find" button. Type in query. Select the relevant topic	Browsing =>Finding a topic Finding a topic => Click-Find Click-Find not satisfied. Finding a topic => Looking-For-Table-Of-Content Looking-For-Table-Of-Contents not satisfied. (Owing to space contraints, we will not list down all the possible buttons suggested by the user model such as "Back", "Previous", etc. When the suggestion was "Next", the simulation continued with going to the next screen of the prototype. The "Find" button was also not available in the second screen. The whole process continues for the second screen.) In the third screen the goal was satisfied because the "Find" button was found. Based on user's feedback, the prototype was modified with the "Find" button only appearing in the third screen. The first two screens only provided general information about the system, which the user thought should only contain the "next" button.	Based on user's feedback, the prototype was modified with the "Find" button only appearing in the third screen. The first two screens only provided general information about the system, which the user thought should only contain the "next" button. It is interesting to note that when the hand-simulation was performed, it took a number of cognitive steps to reach the third screen with the "Find" button. The goal was satisfied like in (I) and (II).

4.4 Improved conceptual design for better hypertext

This chapter began by identifying the problems faced by designers when using the conventional iterative development process. In §4.3, the different evaluation methods and techniques that can be applied by designers to test usability of hypertext were examined. If designers were to apply these methods and techniques religiously to the development lifecycle, quality hypertext can be guaranteed. There can be many propositions. One way, postulated in this thesis, is to incorporate the use of cognitive task diagrams and executable user models within the development lifecycle for the building of well-structured hypertext, leading to an improved Conceptual Design Stage built on top of the well-established development process (see figure 4.25).

These executable user models can then be used iteratively to test the prototype before testing it with real users. Through this improved Conceptual Design Stage, designers can at least ensure that the prototype hypertext is sufficiently tested using executable user models and real users, before actually building the system. Though not supported by any quantitative results, the author believes it is a more cost-effective way of designing that adheres to good usability practice, yet avoids as many errors as soon as possible. This would help designers manage the complexity of the design process so that producing well-structured hypertexts is not a matter of chance.

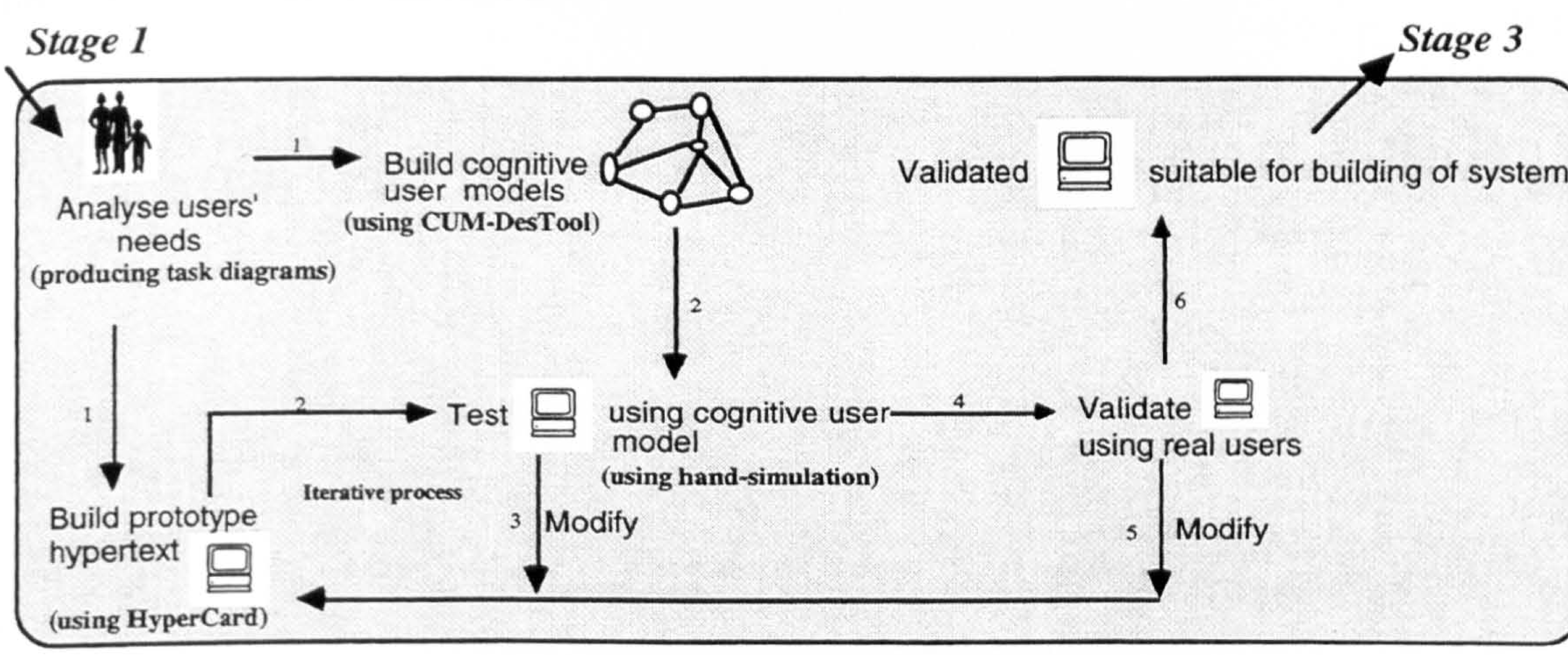


Figure 4.25. Improved Conceptual Design for better hypertext using task graphs and executable user models

4.5 Conclusion

This chapter emphasised the need for an engineering, task-based approach to understand end-users' need to eliminate, or at least ameliorate the LIH problem. To address the lack of a disciplined and systematic approach to designing well-structured hypertexts to meet end-users' behaviour and navigation needs, a cognitive task-based approach by giving these issues more structure was described. Because browsing is a common activity associated with navigating hypertext, it was used as an illustration to show how a combination of TA and TD techniques were employed to break it down into smaller subtasks, represented as cognitive task diagrams. Using the task diagrams, a prototype hypertext built in chapter 3 was modified based on the design recommendations. Preliminary findings show that it was rated better than the original prototype.

Another proposition made in this chapter to strengthen the Conceptual Design stage in the conventional iterative process is to ensure that evaluation methods and techniques are effectively employed by designers to measure the usability of hypertexts, thereby ensuring quality hypertexts are produced. These evaluation methods and techniques are categorised under real user testing and non-human user testing.

The need for good hypertext structure (Approach Three) to address the LIH problem in hypertext is discussed in the next chapter.

Chapter 5

Need for good hypertext structure

Chapter objectives:

One proposition to help designers build well-structured hypertext is improved authoring to create more comprehensible structures. This chapter investigates hypertext structural design issues which include:

- examining the reasons for presenting the hypertext structure clearly
- defining what a hypertext structure is
- identifying ways to structure information in hypertext
- identifying designers' likely authoring structures for hypertexts
- identifying useful heuristics and metrics to analyse hypertext structure

5.1 Introduction

This thesis argued in chapter 3 that a good interface design is essential to prevent end-users from experiencing LIH when they navigate hypertexts. Reasonable solutions sought after to provide good interfaces include all forms of navigational aids and user-centred design principles such as consistency and clarity of presentation, minimal mental overload of end-users, and well-structured network of nodes and links. Chapter 3 also pointed out that design principles and guidelines should go beyond just providing an attractive interface. Not only is navigation through hypertext influenced by navigational aids built into hypertext (see chapter 3) and browsing strategies employed by end-users (see chapter 4), the structure of hypertext also plays a vital role in helping end-users to navigate round hypertext.

To illustrate, let us compare navigating hypertext with reading a book. When we read a book, the author of the book makes the assumption that it is to be read in a normal sequential order. In other words, if the contents of the preceding pages are not fully understood and digested, readers will often not understand the pages that follow. Table of contents, page numbering, index, *etc.* are some of the well-accepted organisational elements one would expect to find in a book. On the contrary, because the very nature of hypertext boasts of non-linear reading, readers in hypertext should be able to follow links in any order, and should never encounter information that relies on materials they have not read. Unlike conventional text, there are no universally accepted organisational principles and structures in hypertext (Thimbleby, 1995a). Lacking these, end-users will have to "guess" the structure of hypertext, and try to understand what the interface wants to convey. A frequent criticism against most hypertexts is that they strongly resemble a maze, in which it is easy for end-users to get "lost" (for example, De Bra, 1996). The questions we now want to ask are: is this acceptable in terms of good hypertext authoring? Does this affect the way end-users navigate round hypertext? Perhaps designers should take the responsibility for creating this "mess" in hypertext because very often hypertext design is considered a "creative" task, and consequently, much effort is invested in making it look good "technically" without careful consideration given to the organisation of information and the modelling of associations of nodes and links (Signore, 1995; Addison and Dudman, 1995).

Perhaps designers should not be entirely "blamed" for flawed hypertexts if good authoring support is not available, making hypertext authoring difficult (Thimbleby, 1997; Botafogo, Rivlin and Shneiderman, 1992). Because hypertext makes it possible to embody alternative and multiple structures out of the same information, authoring hypertext becomes a great source of difficulty for designers. Parunak (1991) explains that information links in hypertext have a way of getting so tangled that information actually becomes harder to find, rather than easier. When writing a book, the author can use an algorithm for completing a task — for instance, start at page 1, process it, turn to next page, and so on, then stop at the final page.

In contrast, there is no algorithm for writing a non-trivial hypertext that guarantees completion of a task. Without a "sense of progress", a hypertext designer will not know when he is able to stop or when his task is completed, or, if he pauses, how to resume without repetition; there may always be other pages or links in the hypertext that need considering. Hierarchy comes between linear and general graph. It may have some advantages of hypertexts without the complexity. Though one may not need to be a sophisticated mathematician to come up with a formula to predict the writing paths, Dillon (1994) argues that one would be a poor student of human psychology if one really believed any resulting formula provided an accurate representation of the writing process.

One proposition to help designers is improved authoring to create more comprehensible structures. Designers need authoring help to ensure that effective structuring is embraced in hypertext design. The question of how hypertext might be structured to maximise usability is of direct concern to this thesis. This chapter investigates some of these concerns often voiced by well-intentioned hypertext designers with regard to hypertext structural design: is there a relevant structure for designing hypertext? If there is, what is this relevant structure? How to ensure if this structure is "optimal"?

To assist designers to build well-structured hypertexts, this chapter achieves the following objectives (see figure 5.1):

- To investigate the underlying *reasons* for the need to present the hypertext structure clearly from a human factors perspective (§5.2)
- To identify ways to *structure* information in hypertext (§5.3).
- To investigate *likely* structures designers use when designing hypertext (§5.3)
- To identify heuristics to *analyse* the structure of hypertext (§5.4).

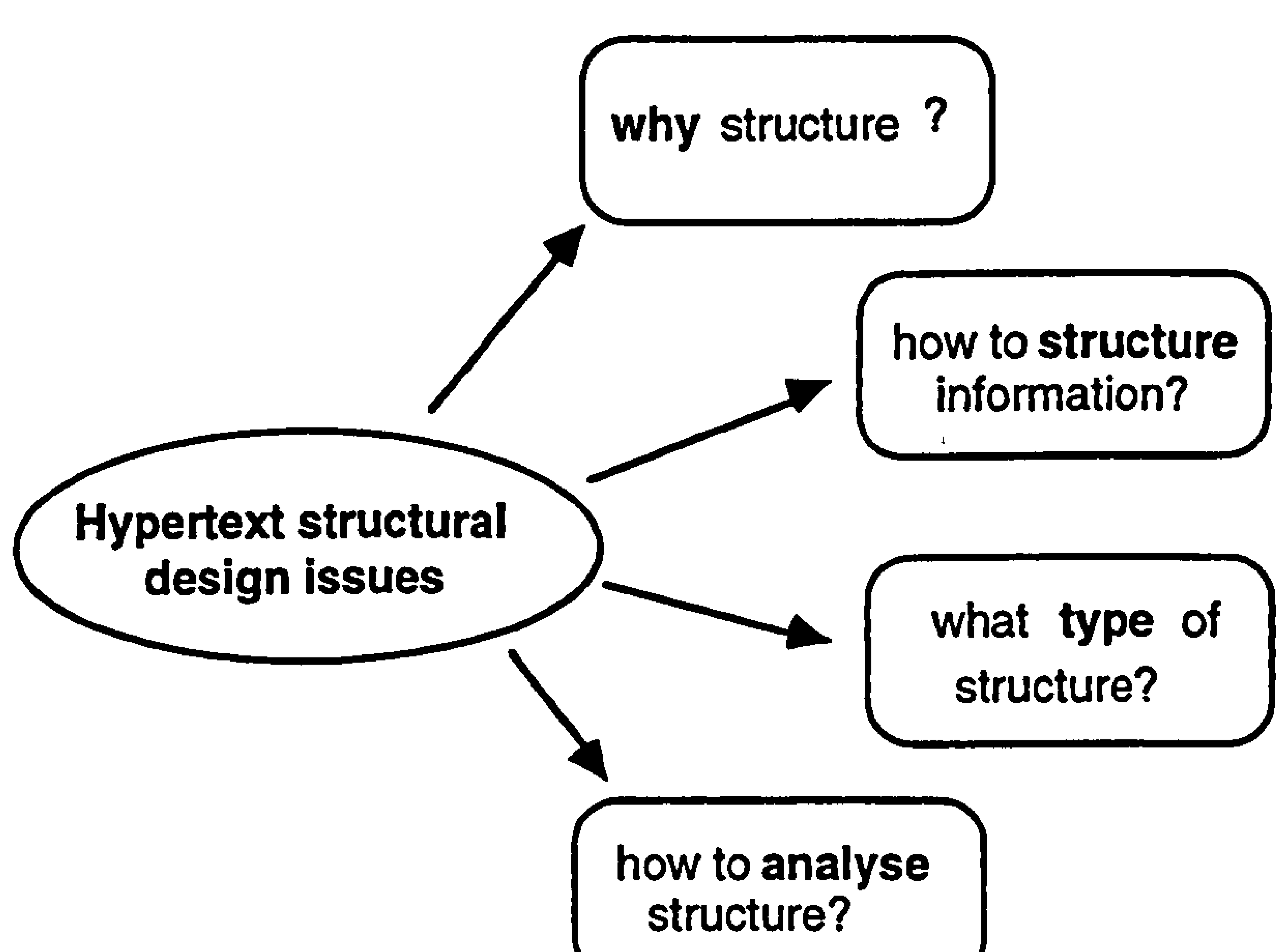


Figure 5.1. Issues concerned in hypertext structural design

5.2 Need to present structure clearly

The specific aim of this section is to derive the reasons for the need to present hypertext structure clearly from a human factors perspective, since the first law of ergonomics is "know your user". This discussion is based on the data collected from experiment 3.1 in §3.5.1. To recapitulate, in experiment 3.1 six subjects were asked to evaluate the ACM's *Hypertext-on-hypertext* using a questionnaire (see appendix B1). This section will analyse their responses to question 26 of the questionnaire regarding their representation of the structure of ACM's *Hypertext-on-hypertext* and the phenomenon of feeling "lost".

In question 26, subjects were given ten different structures to choose from to represent the structure of ACM's *Hypertext-on-hypertext* (see figure 5.2). In these structures, 'o' represents a node of information, and -> represents a uni-directional link between nodes.

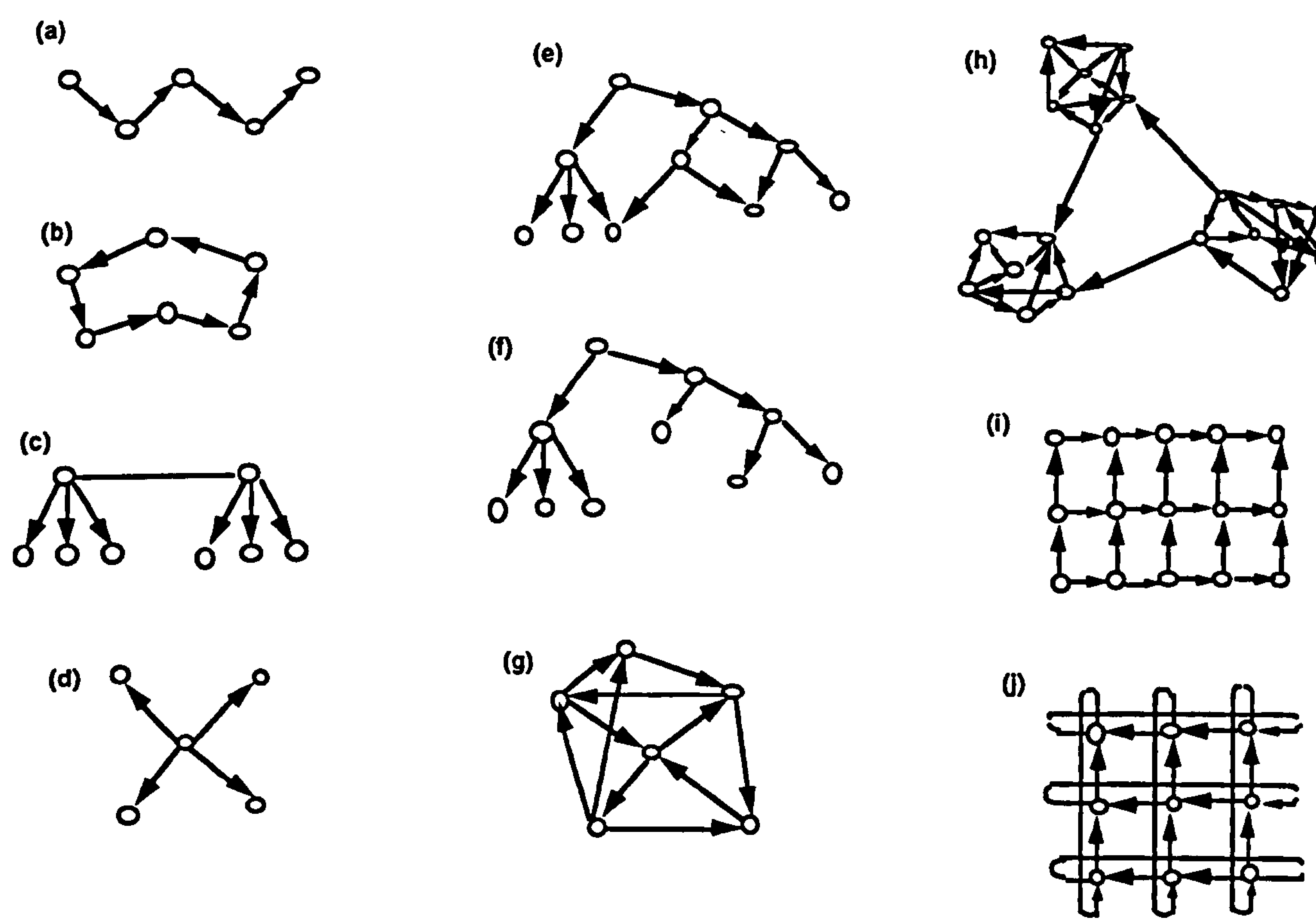


Figure 5.2. Different structures of ACM's *Hypertext-on-hypertext* as listed in question 26 of the questionnaire.

These structures generally fall under six categories (as defined by Parunak, 1991): (i) *linear* for (a) and (b) in which each node has at most one child and one parent; (ii) *hierarchy* for (f) in which one node has no parents and the others have exactly one parent; (iii) *hypercube and hypertorus* for (i) and (j) respectively in which links intersect at nodes; (iv) *directed acyclic graph* for (a), (c), (d), (e), (f) and (i) in which it is impossible to come back to a node; (v) *clumped* for (h) in which each clump is a node, and inquiry about the pattern is obtained from connections between clumps; and (vi) *arbitrary* for (c) and (g). However, if none of these structures reflect the subjects' representation of the structure of ACM's *Hypertext-on-hypertext*, the subjects were asked to draw their own structures.

Tabulation of results

Table 5.1 compares factors such as the subjects' experience, satisfaction, feeling of "lostness" and completion time, where R means Researcher and NR means Non-researcher; and E means Experienced and I means Inexperienced. The values attached to subjects' satisfaction (question 3), subjects' feeling of "lostness" (question 31), subjects' perception of the effectiveness of hypertext (question 39) are translated from a semantic differential scale using a 7-point scale (as described previously in §3.5). The time taken to complete the tasks (question 33) is recorded in minutes. The last column in table 5.1 records the subjects' representation of the structure of ACM's *Hypertext-on-hypertext*.

Table 5.1. Subjects' feedback of ACM's *Hypertext-on-hypertext* comparing end-users' experience, end-users' satisfaction, subjects' perceived effectiveness, end-users' feeling "lost", completion time, and subjects' representation of the hypertext structure.

	Subject's profile	Subject's experience	Subject's satisfaction	Subject's perceived effectiveness	Subject feeling "lost"	Time to complete tasks (min)	Subject's representation of structure
S1	R	E	4	5	3	30	j
S2	R	E	5	6	3	35	own diagram
S3	R	I	5	6	none	60	e,h
S4	R	I	2	1	1	90	c
S5	NR	I	5	6	3	50	h
S6	NR	I	4	4	3	40	e,f

This thesis makes the following observations when analysing table 5.1:

Observation 1

It is noted that all six subjects had some form of representation of the structure of ACM's *Hypertext-on-hypertext*. However, these representations did not match the actual structure of ACM's *Hypertext-on-hypertext* (see figure 5.3). Except for subject 3, the rest of the subjects experienced some form of "lostness" regardless of whether they were experienced or not (see table 5.1).

From this observation, it would seem that if a subject has a preconceived notion of what the structure and design of the hypertext is supposed to be (though it may not be necessarily right), but if that structure does not match the actual structure of the hypertext he is using, he will experience the LIH problem regardless whether he is a novice or experienced end-user.

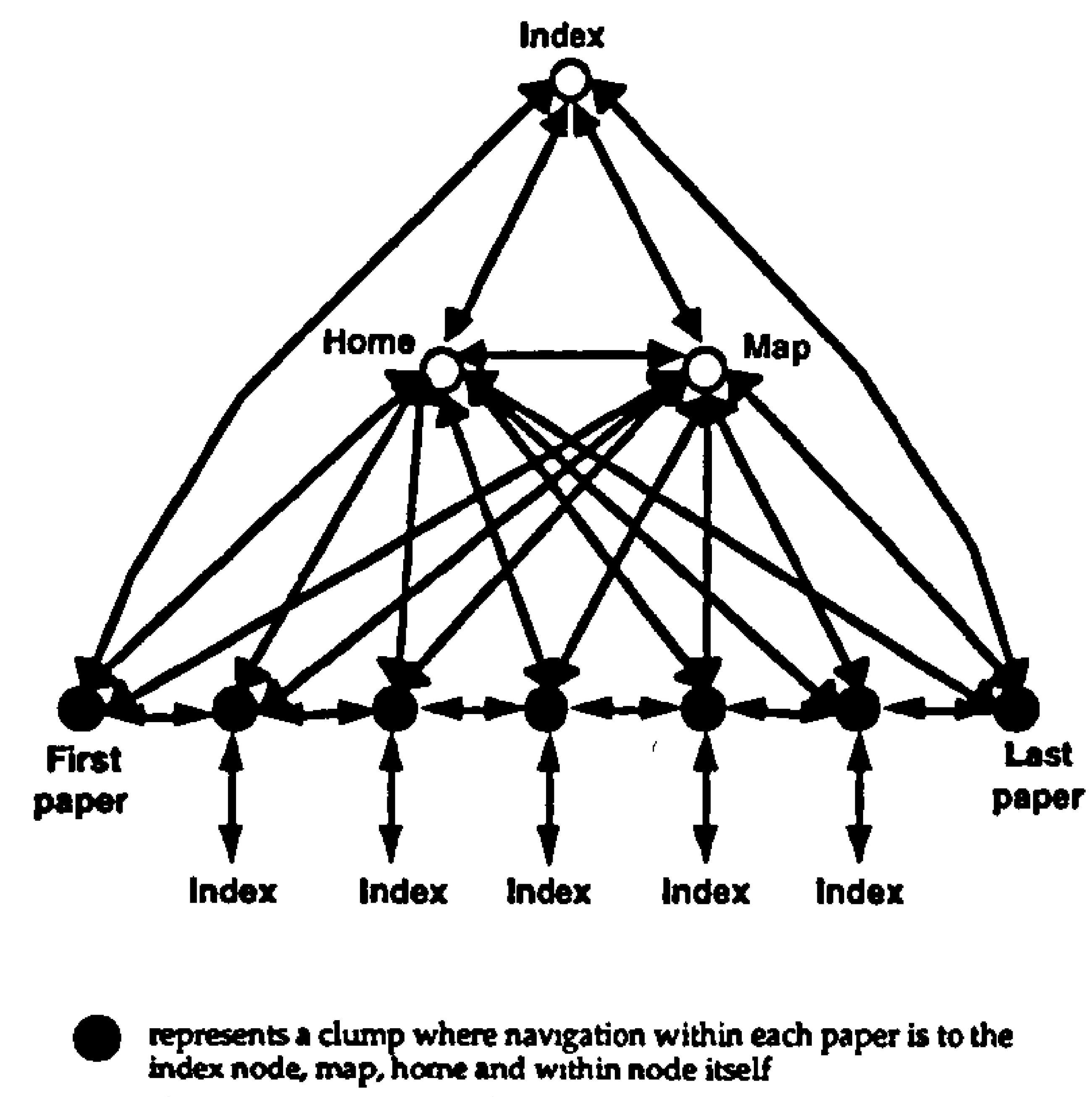


Figure 5.3. Actual structure of the ACM's *Hypertext-on-hypertext*

Observation 2

Subjects 1 and 2 spent half the time trying to complete the tasks compared with the rest of the inexperienced subjects. However, with the exception of subject 3, subjects 1 and 2 also registered some degree of "lostness" similar to the inexperienced subjects. When subject 1 was interviewed, he expressed frustration at the way ACM's *Hypertext-on-hypertext* was structured and organised. On the other hand, subject 4 being new to hypertexts had less expectation and therefore spent more time understanding the structure of ACM's *Hypertext-on-hypertext* in order to complete the tasks given.

It would seem that the more experienced the subject is, the less time he spends in trying to understand the structure of the hypertext, and therefore experiences some degree of "lostness". Conversely, if a subject has little or no experience with hypertexts, he compensates it by spending more time trying to learn about the hypertext structure, and how to navigate round the hypertext.

Observation 3

It is to be noted that all of the subjects had the correct answers, though they took different amount of times to complete the tasks. Subject 4 was very frustrated because he found the structure of ACM's *Hypertext-on-hypertext* confusing and thus ineffective in helping him to complete the tasks. As for the other subjects, they rated the effectiveness of ACM's *Hypertext-on-hypertext* well in helping them to complete the tasks, even when their perceived structures and the actual structure of ACM's *Hypertext-on-hypertext* did not match.

It would seem that subjects' perception of the effectiveness of ACM's *Hypertext-on-hypertext* did not interfere with the actual performance of the subjects, even though it affected their satisfaction of it.

Discussion

Observation 1 notes the impact of the subjects' representation of the structure of ACM's *Hypertext-on-hypertext* on the LIH problem. Observation 2 studies how the subjects' representation of ACM's *Hypertext-on-hypertext* and the subjects' prior experience with hypertexts affect their satisfaction, time spent in completing tasks and the LIH problem. Observation 3 notes that subjects' representations of ACM's *Hypertext-on-hypertext* and their perceptions of its effectiveness affect the time spent in completing the tasks.

To explain these observations, this thesis turns briefly to cognitive psychology and in particular, the distinction between design model, user's mental model and system image of general interactive systems proposed by Preece, Benyon, Davies, Keller and Rogers (1993). This thesis adapted the idea for hypertext as illustrated in figure 5.4. *Design model* represents how designer thinks the hypertext should work. *User's mental model* represents how end-user thinks the hypertext works, which is influenced by his domain knowledge, general knowledge with computers, and experience with hypertext. *System image* refers to the interface design of hypertext, of which the purpose is to reflect the underlying network of nodes and links to end-users.

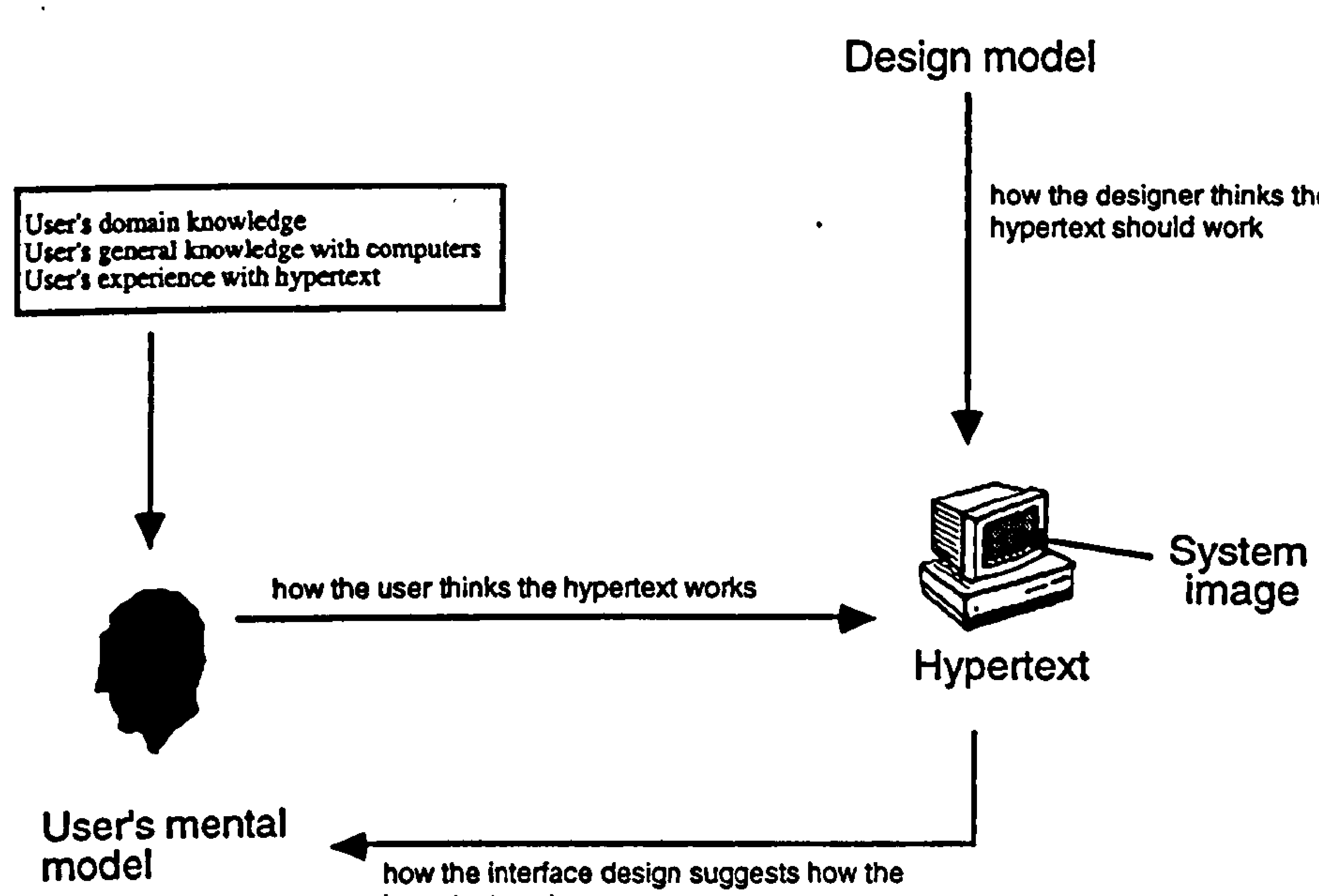


Figure 5.4. Distinction between design model, user's mental model and system design (adapted from Preece *et al.*, 1993)

Table 5.2 summarises the above three observations into one or more situations, making remarks on likely problem area(s). Preece *et al.* (1993) advise designers to design a system image that enables an end-user to develop a suitable mental model, because failure to do so will create many apparent end-user problems. Situations 1 and 2 can be avoided if designers are able to represent the information structure accurately and clearly to end-users, thus tackling both the conceptual and spatial disorientation of the LIH problem. Except for situation 1, the rest seem to point to errors made during the design process. Situations 3 and 4 can be avoided by understanding the end-users' needs better, which has already been discussed in chapter 4. In general, situation 3 may lead to situation 1 and situation 4 may lead to situation 2.

Table 5.2. Analysis of observations and associated problem areas

Situation	Description		Remark
1	If how the end-user thinks the hypertext works does not match with how the interface design suggests the hypertext works, the end-user may experience conceptual disorientation (discussed in §2.2.2).	→	A user problem
2	If the interface design does not reflect accurately and clearly the information structure of the hypertext, then spatial disorientation (discussed in §2.2.3) may take place.	→	A design problem
3	If how the designer thinks the hypertext should work does not match how the end-user thinks the hypertext works, then perhaps the designers have not carried out requirements capture and analysis of end-users' needs well enough.	→	A design problem
4	If how the designer thinks the hypertext should work does not match with how the interface design suggests the hypertext works, then perhaps incorrect requirements specifications have been drawn up.	→	A design problem

From the discussion above, it can be concluded that the subjects' perception of the structure of ACM's *Hypertext-on-hypertext* influenced their interactions with it in terms of their level of satisfaction, their feeling of "lostness" and the time spent in completing the tasks. Interestingly, the performance of the subjects in this experiment were not affected since inexperienced subjects compensated for their lack of experience by spending more time understanding the hypertext structure and how to go about getting information.

Regardless of the putative constraints of hypertext, it seems certain that the end-users possess some knowledge that provides information on the probable structure and organisation of the elements in hypertexts (Dillon, 1994). End-users build models of what is happening in their minds, and they use these "mental models" in their interactions. To proceed with their interactions with the hypertext, end-users expect it to give them cues, otherwise they will need to rely on their prior experiences with hypertexts and computer systems. This in itself is not bad. However, there are occasions when end-users' prior experiences interfere with their understanding and interactions with the current hypertext they are navigating. The non-linear structure of hypertext makes it difficult for end-users to ascertain the layout of the hypertext network, thus leading to the LIH problem (Taylor and Self, 1990).

In studying the relationships between LIH and the organisation of end-user interfaces, Mantei (1982) asserts the need to make the structure of the user interface clear to end-users, otherwise, if the structure of the user interface as perceived by the end-user is not obvious, it causes LIH. If end-users were to be helped in navigating hypertext, the user interface needs to be usable. In other words, it should explain the internal structure and representation of nodes and links to end-users in the simplest and most effective way, and for the end-users to communicate their intentions and obtain the answer to their intentions. Mendes and Hall (1997) confirm that the more explicit the meaning is, the better it is for the end-users so that they do not need to guess the relationships in order to construct a representation of the hypertext structure.

In viewing structure in hypertext, the conceptualisation emerges of information as space and the hypertext end-user as a navigator. As such, much solutions seek to obtain a better insight into the relationship between the hypertext's structure and the end-user's perception of the structure fall into two broad categories: (i) improved facilities to aid the navigation process; and (ii) efficient navigation strategies employed by the end-user while browsing.

Solutions for (i) include indices, tables of contents, screen headings, *etc.* as landmarks to provide hypertext end-users with information on where they are in the hypertext, just as signposts, buildings and street names aid navigation in physical environments. Maps, tables of contents, *etc.* are more appropriate for smaller hypertexts (Vocht, 1994). However, with larger hypertexts, solutions include investigations into controlling the amount of displayed information using fractal views (Kioke, 1995), mapping the full hierarchy onto a rectangular region through the efficient use of space using tree-maps (Johnson and Shneiderman, 1991), *etc.* To recapitulate, the topic of effective navigational aids has been dealt with in chapter 3 when design principles and guidelines were discussed. The design principles postulated in §3.3.2 essential for designing good hypertexts include suggestions on how to clearly present the hypertext structure, to ensure ease of use, to maintain consistency of presentation and to ensure maximum visibility.

In (ii), experimental findings show that the depth-first search strategy is the most suitable navigational strategy for obtaining a proper perception of the hypertext structure (Vocht, 1994). Vocht explains that this is because the depth-first strategy is very fast, and that this strategy yields access to nodes that are well-distributed over the graph structure. Smith and Newman (1997) claim that there is good evidence that directed browsing tasks are best supported by hierarchical structures, and undirected browsing tasks by network structures.

5.3 Structuring hypertext information

Knowing how to display the hypertext structure accurately and clearly to end-users in the interface is important in ensuring effective user navigation. However, equally important for designers is to know how to structure the hypertext information effectively. Horn (1989) proposes four key questions designers should ask when designing a hypertext (see table 5.3):

Table 5.3. Four key questions asked when designing a hypertext

1.	What size should the nodes be?
2.	What should be the connection between nodes?
3.	How are links characterised?
4.	How are links identified and labelled?

There are some "threads" of investigation this thesis would like to explore with regard to how information in hypertext is structured. §5.3.1 begins by defining what constitutes a hypertext structure. §5.3.2 analyses the different approaches undertaken to structure hypertext. In seeking to understand how to structure hypertext information, §5.3.3 investigates the most likely structure designers would use in authoring a hypertext.

5.3.1 Defining hypertext structure

There are many different definitions and interpretations of the term "structure" as used in hypertext. Conklin (1987) talks of structure as being imposed on what is browsed, that is, structure is built based on knowledge gained from using hypertext. Vocht (1994) sees structure as part of the semantics of the hypertext, and an understanding of these semantics are essential to retrieve information. Dillon (1994) defines structure as a convention for both the designer to conform to expectations of format, and for end-users, so they know what to expect. Mendes and Hall (1997) see structure as comprised of both the rhetoric involved in the composition of nodes and in the maintenance of global as well as local coherence, and the semantics of the relationships among nodes, given by links.

To define hypertext, this thesis turns to Campbell and Goodman's HAM or Hypertext Abstract Machine (1988), a reference model conceptualised to explain the architectural make-up of hypertext. Hypertext is defined in this thesis as basically consisting of three layers, each layer serving a very different purpose (see figure 5.5):

- *Interface design layer.* This layer is the user interface, which is concerned with matters regarding screen design, use of navigational aids, dialogue between end-users and the hypertext. What appears on the interface design should reflect accurately and clearly the underlying structure of nodes and links, and their relationships with each other within the network. This has already been discussed in some depth in §5.2.

- *Information structure layer*. This layer refers to the network of nodes and links. It is concerned with the type and size of nodes, and the type of links present, and how they are linked together in the network. Therefore, a meaningful information structure layer is one that reflects semantic information. This area will be discussed in greater detail later in this section.

- *Database layer*. For effective storage and network access of nodes and links to be realised, a database approach is useful. The database layer refers to the node and link repositories. Not all hypertexts have implemented this layer, which can be a serious management issue, especially if very large hypertexts are involved. This, though an interesting and challenging area, is outside the scope of this thesis.

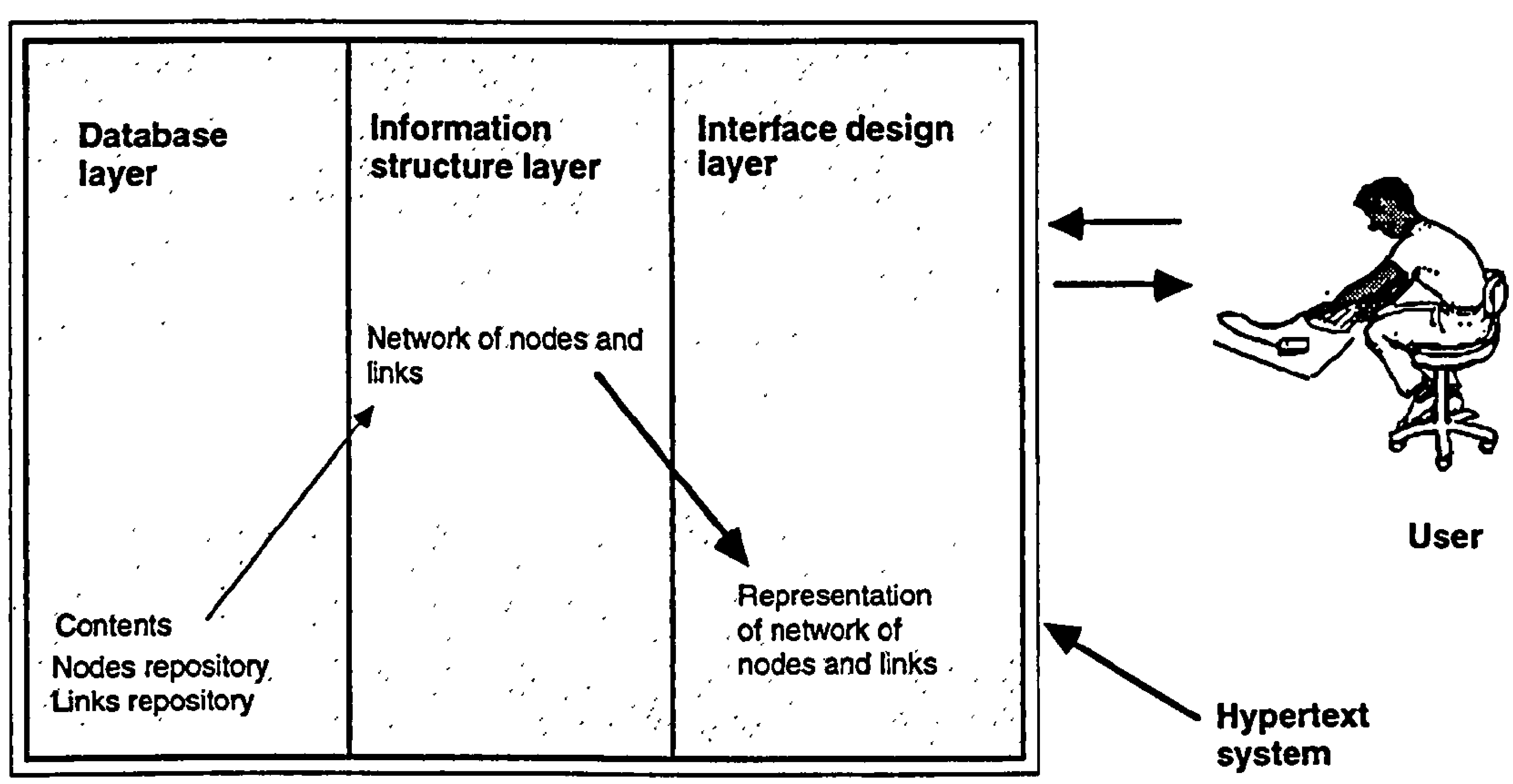


Figure 5.5. Hypertext architecture

Whenever the term "structure" is referred to in this thesis, it means the information structure layer consisting of a network of nodes and links. Nodes are documents which contain text, graphics, audio, video, as well as source code or other forms of data and they can be typed, semistructured or composite. Most end-users of hypertext favour using nodes which express a single concept or idea (Conklin, 1987). Links are active references that allow the end-user to move to other parts of the hypertext databases and can be of many types: explicit (hard-coded links) or implicit (keyword links or links that emulate the human mind's associative mechanism); hierarchical (organisational links) or non-hierarchical (referential links) (Signore 1995; Conklin, 1987; *etc.*).

Besides ensuring compatibility of the structure of hypertext and end-users' tasks, Hardman and Sharratt (1990) stress the importance of hypertexts to be adaptable to the different types and levels of end-users and their browsing patterns. This means that end-users should be made aware of promising links that exist among the various nodes. One suggestion by Signore (1995) is to examine the degree of affinity between nodes. Parunak (1991) suggests that in order to reduce the complexity of a hypertext, designers should devise link ordering conventions to help end-users to navigate by (i) patterning links to permit end-users to employ the various kinds of navigational strategies, and (ii) typing links to enable end-users to filter out unneeded connections. He argues that, although this may lead to an increase in the total number of links, the total efficiency of the hypertext should improve since only links of the types of interest for any particular operation are displayed. Links should have meaning, that is, the graph should be a semantic network (Addison and Dudman, 1995). Designers in building the system should characterise each node with a set of keywords. These keywords are then compared and matched between the different nodes in order to suggest possible links between nodes.

Other investigations include artificial intelligence techniques such as machine learning, neural networks, semantic networks, *etc.* This would open up an entire area of investigation which is not the intention of this thesis.

5.3.2 Approaches taken to structure hypertext

In understanding the likely structures designers would use to represent hypertext, this thesis examines firstly the different approaches undertaken to structure hypertext as shown in figure 5.6. These approaches are described as follows:

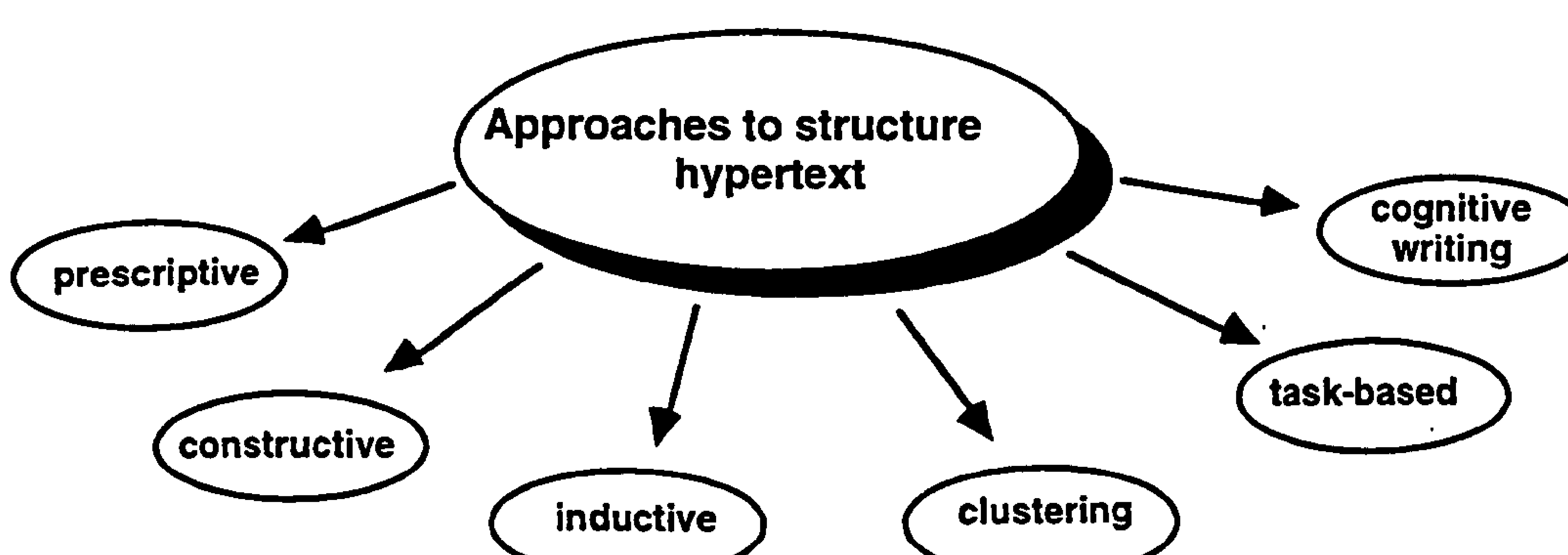


Figure 5.6. Approaches to structure hypertext

- *Prescriptive approach.* Some researchers adopt this approach by imposing a simple, regular structure on intractably complex information. These structures can be linear, hierarchical, and recursive. While hypertext is intrinsically non-linear, some hypertext authoring tools allow extensive use of linear structures, in the form of cards and stacks in HyperCard, or scrolling windows in Guide. NoteCards, for example, provides a hierarchy-promoting filebox construct, but also includes equally extensive support for non-hierarchical browsers. Another example is KMS, where every node is part of a single overall hierarchy, with tree links supporting the hyperdocument's hierarchical structure, but cross-hierarchical special links do not (Garzotto, Paolini, Schwabe and Bernstein, 1991). Some designers also adopt a more complex, recursive approach where complex information structures are built using recursively defined structures. An example is gIBIS, a specialised hypertext system, intended to reflect the structured but infinitely recursive process through which designs are proposed and refined (Conklin, 1987).

- *Constructive approach.* This approach seeks to discover and clarify structural regularities in specific documents using methods such as visualisation tools, and construction kits to help designers generate regular hypertext structures easily and systematically (Garzotto *et al.*, 1991). Examples include NoteCards and Intermedia, where graphical overviews of hyperdocuments are automatically generated. Nodes are represented as icons, and links are represented as connecting lines. Construction kits allow end-users to create multiple copies of individual hypertext nodes with a common internal structure. An example is HyperCard.

- *Inductive approach.* This approach identifies structures implicit in existing linear documents and then incorporates them into hypertexts. Glushko (1992) proposes using typographic conventions in paper documents as structural cues, for example, a hierarchical structure can be induced from the section/subsection hierarchy.

- *Clustering.* This approach abstracts related information, and puts it into clusters (Garzotto *et al.*, 1991). Because hypertext mimics the associative structure of the mind, some researchers assume that a good hypertext should resemble the micro-structures of the mind (Am, 1994). Many ways of associatively structuring knowledge have been investigated.

- *Task-based approach.* Several researchers (for example, Rada and Murphy, 1992) stress the need for information in hypertext to be structured in such a way to support end-users' tasks. The types of information structures that have been investigated are hierarchical structures, network structures and a combination of both. Although there seems to be disagreement on what structure is "best" for what task, researchers seem to agree that certain structures support certain tasks better.

- *Cognitive.* Authors do not seem to change their cognitive processes because they are using hypermedia and they usually have their own pre-determined way of organising the hypertext structure (Mendes and Hall, 1997). Designers use a mixed approach combining both top-down and bottom-up approaches when writing hypertext (Nanard and Nanard, 1995). Other researchers make use of cognitive models or structures of learning to structure hypertext for effective learning (Eklund, 1995). They believe that if learning is to be facilitated by the formation of node-link structures in the learner, then this needs to be reflected in the construction of hypertext according to the way learning takes place in the domain.

To summarise, this thesis has just described some common techniques designers may consciously or unconsciously employ to structure hypertext. In some cases, they are helped by the authoring tools used. This thesis is not so interested in the ways in which designers structure hypertext, but rather the outcome of the structuring process. In the next section, an empirical investigation was conducted to establish designers' likely structure when designing hypertext, an important knowledge of interest to developers of authoring tools for hypertext. Chapter 7 describes the design and development of a prototype authoring tool for hypertext, based on investigation carried out to understand designers' likely authoring structures.

5.3.3 Experiment 5.1: Designers' authoring structures

Objectives

Experiment 5.1 investigates the most intuitive structure(s) designers would use when designing a hypertext.

Protocol

Students from the Masters in Educational Multimedia Course at Middlesex University were asked as part of their course work to come up with a conceptual design of a multimedia educational package in any subject area of their choice. In this assignment, they were asked to provide a map showing the overall structure of the package to reflect clearly how information was linked. The exact wording of the instructions is shown in table 5.4:

Table 5.4. Instructions to subjects to come up with an overall map of the structure of an educational package

a.	Design an educational, multimedia package to allow readers to access information easily.
b.	Draw a map to show the overall structure of the multimedia package. The structure should indicate how information is linked to each other. Use a box to represent a screen of information, and an arrow '---->' to represent a one-directional link between two screens.
c.	Implement this multimedia package using a multimedia authoring tool, such as Toolbook or MacroMedia Director.

A total of nine students took part in this experiment. Table 5.5 shows the students' profiles and the respective subject area of the multimedia package. When the students were asked on how they would rate themselves as designers, five out of nine indicated that they were intermediate designers, while the other four indicated there were novice designers. All of them had some experience in hypertext authoring. Five out of nine students wanted to design the multimedia package on mathematics, while the other four focus on other areas in chemistry, control, aviation and religion. The subject areas chosen for the educational package were diverse, with no two students working on the same subject area.

Table 5.5. Students' profiles and respective subject area of multimedia package

Student No.	Student's perception of authoring skills	Authoring experience	Subject area of multimedia package
1	Novice	yes	Parent's guide to drugs and solvents
2	Novice	yes	Teaching basics of faith
3	Novice	yes	Teaching numbers
4	Novice	yes	Transport controls
5	Intermediate	yes	Learning geometry
6	Intermediate	yes	Teaching alphabets, numbers, colours
7	Intermediate	yes	Teaching shapes
8	Intermediate	yes	Understanding aircraft and flight
9	Intermediate	yes	Teaching fractions

Results

Since the question of the most intuitive structure(s) designers would use when designing a hypertext is of direct concern, this section will not report on the implementation of this educational package developed by the subjects. Therefore, in this experiment, it suffices to have a first-cut understanding of the students' design structures of the package, since the intention is to investigate the most intuitive hypertext authoring structure. Appendix E1 to E9 show rough sketches of hypertext structures produced by the students. Though this was the initial design stage, students 5 to 9 who claimed to be intermediate designers appeared to be able to come up with a fairly reasonable sketch of the structure of the package. These structures were mostly hierarchical. The idea of a hierarchical structure was not so evident in the sketches produced by the novice students 1 to 4. In particular, student 4 seemed to be having difficulty producing the sketch of the hypertext structure. This may be due to the lack of authoring experience. However, a close examination of the sketches of students 1 to 4 reveals that they were somewhat hierarchical.

This observation provides two interesting insights:

- It seems that more experienced hypertext designers find hierarchical structuring of information "intuitive".
- It also seems to suggest that though novice designers may exhibit problems in expressing their design ideas, they also tend to adopt a hierarchical-like structuring of information.

Since designers' likely authoring structures are hierarchical or quasi-hierarchical, it seems plausible to incorporate a hierarchical structure into an authoring tool, so that authoring becomes "intuitive", easy to follow. Independent studies by Mendes and Hall (1997) and Smith and Newman (1996) also confirm the results that hierarchies are most commonly used since they are "intuitive", and are easily understood by both hypertext end-users and hypertext designers (Botafogo, Rivlin and Shneiderman, 1992; Garzotto *et al.*, 1991). However, hypertext designers differ primarily at which stage the structure of hypertext is made hierarchical or semi-hierarchical (Zizi and Beaudouin-Lafon, 1994). Some suggest choosing this structure at the design stage (Garrett, Smith and Meyrowitz, 1986; Acksyn, McCracken and Yoder, 1988), while others suggest that existing hypertexts with complex structures can be transformed into hierarchical hypertext (Rivlin, Botafogo and Shneiderman, 1994).

As for the hypertext end-users, research findings also seem to be divided. Some may argue that perhaps the hierarchical structure may not be the "best", and one should continue the quest to search for the "best" structure. So far, there are differing views among researchers regarding this issue. Batsy, McRae and Timmer (1997) claim from empirical investigation that for on-line systems such as teleshopping, a guided tour rather than a hierarchy supports more effective shopping. If teleshopping involves both searching and exploratory tasks, then this contradicts the claim made by Smith and Newman (1996) that "unfocused browsing or exploratory" tasks are best supported by a network or combination information structure, while "focused browsing or searching" tasks are best supported by a hierarchical information structure. Mohageg (1992), on the other hand, found that trying to perform searching tasks in a network structure produced a negative effect on task performance.

This thesis is concerned with educational, instructional hypertexts, and based on findings made from this thesis and other independent studies (for example, Mendes and Hall, 1997; Zizi and Beaudouin-Lafon, 1994; Botafogo, Rivlin and Shneiderman, 1992; Vries, 1993; Garzotto *et al.*, 1991; *etc.*), hierarchical structures are more "intuitive" and natural to the hypertext designers and end-users. Therefore, this thesis will illustrate in chapter 7 how quasi-hierarchical structures as a framework for capturing and representing data in hypertexts are implemented in an authoring tool. This thesis agrees with Parunak (1991) that this may be the direction forward: the more structure authors build into hyperdocuments, the more information will be available to computerised systems in converting material from one format to another, since increasingly we are dealing with networked and not just stand-alone hypermedia systems.

In going over the students' assignments, some important design omissions were identified and they are reported in table 5.6:

Table 5.6. Some important design omissions from feedback on students' assignments

1. Except for students 1 and 6, the rest went about designing the package without having a clear idea of who the end-users are, for example, the age group, the level of expertise of end-users, *etc.*
2. Sketches from students 1 – 6 demonstrate a weak understanding of the relationships between the different ideas expressed in the package. They were unsure of the directions of navigation from one idea to another.
3. Students 7 – 9 provided fairly clear hierarchical sketches of the structure with a top-down navigation path to many possible branches of information. However, there is no provision to allow navigation from one branch of the structure to another except via the top of the structure.
4. The sketches also show weak navigational links between important information, for example, the index page to the rest of the package.

These omissions provide us with some insights into the hypertext structural problems encountered by hypertext designers. Omission 1 is concerned with designers not having a clear objective in producing the package. A possible way to remind designers is perhaps to make this reminder a feature in an authoring tool. Designers on using the authoring tool will be asked to identify the target group of end-users, and this information can be used to generate the executable user model as described previously in §4.3.3.

Omissions 2 and 4 confirm similar findings from (Mendes and Halls, 1997) that link information was not essential when structuring the hypertext domain. Could it be the case that designers are so overwhelmed with the task of structuring the information that this vital information is "temporarily" ignored? Perhaps an authoring tool can be devised to take care of the link relationships freeing designers to concentrate on structuring the information. This idea is implemented in the prototype authoring tool described in chapter 7.

Omission 3 is concerned with designers not being able to group and link related information therefore, failing to achieve an "optimal" hypertext structure. In the next section, useful heuristics and metrics are identified to analyse the structure of hypertext, so that designers are in a better position to guarantee that well-structured hypertexts are built.

5.4 Structural analysis

The intention of this thesis is not only to help designers and authors structure information, it also wants to help them avoid structural inconsistencies and mistakes. This thesis proposes two useful heuristics to analyse the structure of hypertext (see figure 5.7). A heuristic is a general principle or rule of thumb that is usually but not always effective (Landauer, 1995). Some define it as a "good, second best".

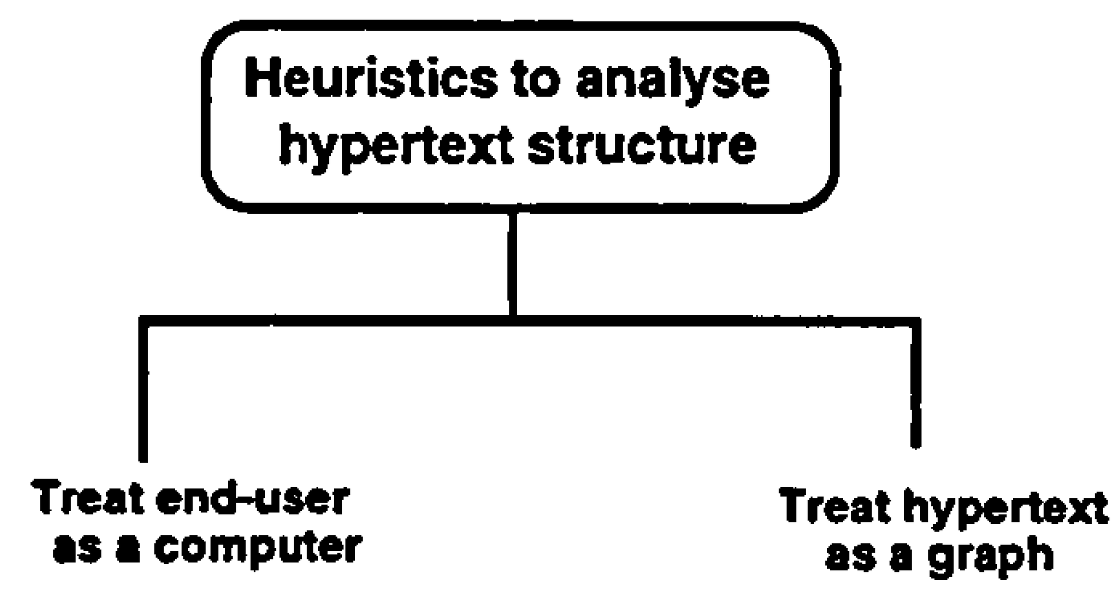


Figure 5.7. Heuristics to analyse hypertext structure

The first heuristic suggests treating the end-user like a computer. Thimbleby (1995b) argues that if a computer cannot operate a user interface, then a human certainly cannot and could only do so by trial-and-error. It is a useful heuristic for it recruits the entire body of computational solutions and methods, both algorithmic and heuristic, to the user interface design. This formal approach involves investigations into the possible paths through a hypertext, which may pinpoint problem areas such as small loops and dead-ends (Thimbleby, 1997; De Bra, 1996).

The second heuristic suggests treating the hypertext as a graph and its end-users embark on a graph-theoretic problem of traversal, search or optimisation, for example, depth-first, breadth-first, least-cost heuristics, *etc.* Because graph theory is derived from a formal mathematical basis, it provides many benefits over other approaches, permitting a range of well-defined algorithms to be applied in structural analysis and determining the metrics for evaluating the information structure they describe (Addison and Dudman, 1995).

Applying these two heuristics to hypertext structure, designers can gain a mathematical insight into hypertext structural design to enable designers to ask analytic questions that relate to getting "lost" and other usability problems. A well-ordered graph of information is not only easier for authors to manipulate, but also helps computers do more to make a hypertext easy for end-users to navigate (Parunak, 1991).

Botafogo, Rivlin and Shneiderman (1992) propose studying two important aspects of the structure of hypertext by firstly, providing authors with different views of the hypertext, and secondly, devising useful metrics to reflect properties of nodes and of the whole structure. Two metrics proposed to measure the properties of hypertext structure are: (i) *compactness* which indicates the intrinsic connectedness of hypertext; and (ii) *stratum* which reveals to what degree the hypertext is organised so that some nodes must be read before others. Thimbleby (1997) implemented in Gentler graph properties to calculate the shortest path to go to each page, and also to perform checks that every link makes sense. However, there are relatively few studies discussing or seeking to use metrics to analyse the structure of hypertext. One may wonder whether the lack of research into this area is because designers find computation of the metrics too tedious and time-consuming.

By performing structural analysis of the hypertext, designers are in a better position to detect design flaws as early as possible in the design process. Table 5.7 compiles a list of some useful metrics that allow designers to perform checks on network connectivity, identification of dead ends and loops, *etc.* Chapter 7 describes how some simple, useful metrics (marked as * in table 5.7) were implemented into an authoring tool to illustrate that formal analysis of hypertext structure can be performed easily and rapidly by automating the calculation of these metrics. Strong connectivity of graphs is a simple but important property because if a hypertext is not strongly connected, a page might be inaccessible to an end-user, or a page might be a "dead end" with no continuation (Thimbleby, 1997).

Table 5.7. A list of useful metrics to perform formal analysis of the hypertext structure

- * number of nodes
- * number of links
- number of hierarchical links
- number of cross-referenced links
- number of nodes per link
- * number of links per node
- number of bi-directional links
- number of edges traversed to reach a node from another related node
- * number of levels away from the root node
- * number of missing or inconsistently-named files/nodes
- * number of nodes with more than 3 successors

Key: * indicates metric implemented in a prototype authoring tool in chapter 7

5.5 Conclusion

This chapter started off examining the need to build well-structured hypertexts from a human factors perspective. One way of so doing is to create more comprehensible structures. The results of an empirical study carried out in this thesis and other independent studies indicate that it is not enough for designers to ensure that a hypertext is well-structured, they must also communicate that design structure clearly to end-users. Ways to structure information in hypertext were identified. An empirical study confirmed that hierarchical structures are most commonly used by designers when structuring information in hypertext. Mathematical insight into hypertext structural design was examined to enable designers to ask analytic questions that relate to getting "lost" and other usability problems. A list of useful metrics was identified to help designers to analyse the structure of hypertext, so that designers can ensure that well-structured hypertexts are built.

In the next chapter, Approach Four which looks into the need for better support tools to help designers in managing the authoring process will be investigated.

Chapter 6

Need for better support tools for authoring

Chapter objectives:

This chapter investigates the reasons why designers have difficulty in authoring hypertexts. Its aims are to identify:

- authoring demands faced by hypertext designers
- problems faced by hypertext designers
- helpful authoring features essential in helping authors cope with the complexity of the design and validation processes

6.1 Authoring demands faced by designers

Since the LIH is a complex problem, this thesis adopts a "step-back-address-fundamental-issues" approach. This is in agreement with the advice given to designers by Birch (1997) from the Engineering and Physical Science and Research Council when solving major, complex problems. To recall, chapter 3 examined the need for good design principles and guidelines, chapter 4 proposed an engineering, task-based approach to understand end-users' needs and chapter 5 investigated structural design issues essential for the building of effective hypertexts. The intention of this thesis is to understand the design and usability issues surrounding the LIH problem in order that a theoretical base can be proposed from which solutions to the LIH problem can be drawn upon when building hypertexts. This chapter is concerned with Approach Four which investigates the authoring demands and problems faced by designers, and how authoring aids can help to reduce the tendency for the LIH problem to arise.

Just as early programmers lacked experience, so do hypertext authors (Brown, 1990). In fact, this whole business of designing and producing hypertexts is still a relatively new discipline. Creating hypertext is complex because of the richness of interconnectivity that exists among nodes and links in hypertext. Thimbleby (1995a) claims that because designers are faced with a vast range of potential structures and an astronomically large number of choices when creating hypertexts, it is difficult to create good hypertexts.

The demands placed on hypertext authors in authoring hypertexts cannot be underestimated. Hypertext authors have to perform many balancing acts to (see figure 6.1):

- ensure that the design and structure of the hypertext is "best" according to its function (Signore, 1995);
- ensure that all the nodes and links created in the hypertext database correspond directly to the windows and links in the display screen so that there should be no redundant/missing links or nodes (Conklin, 1987); and
- incorporate good design principles for screen and information display, dialogue design, navigational aids and on-line assistance (Nielsen, 1995b).

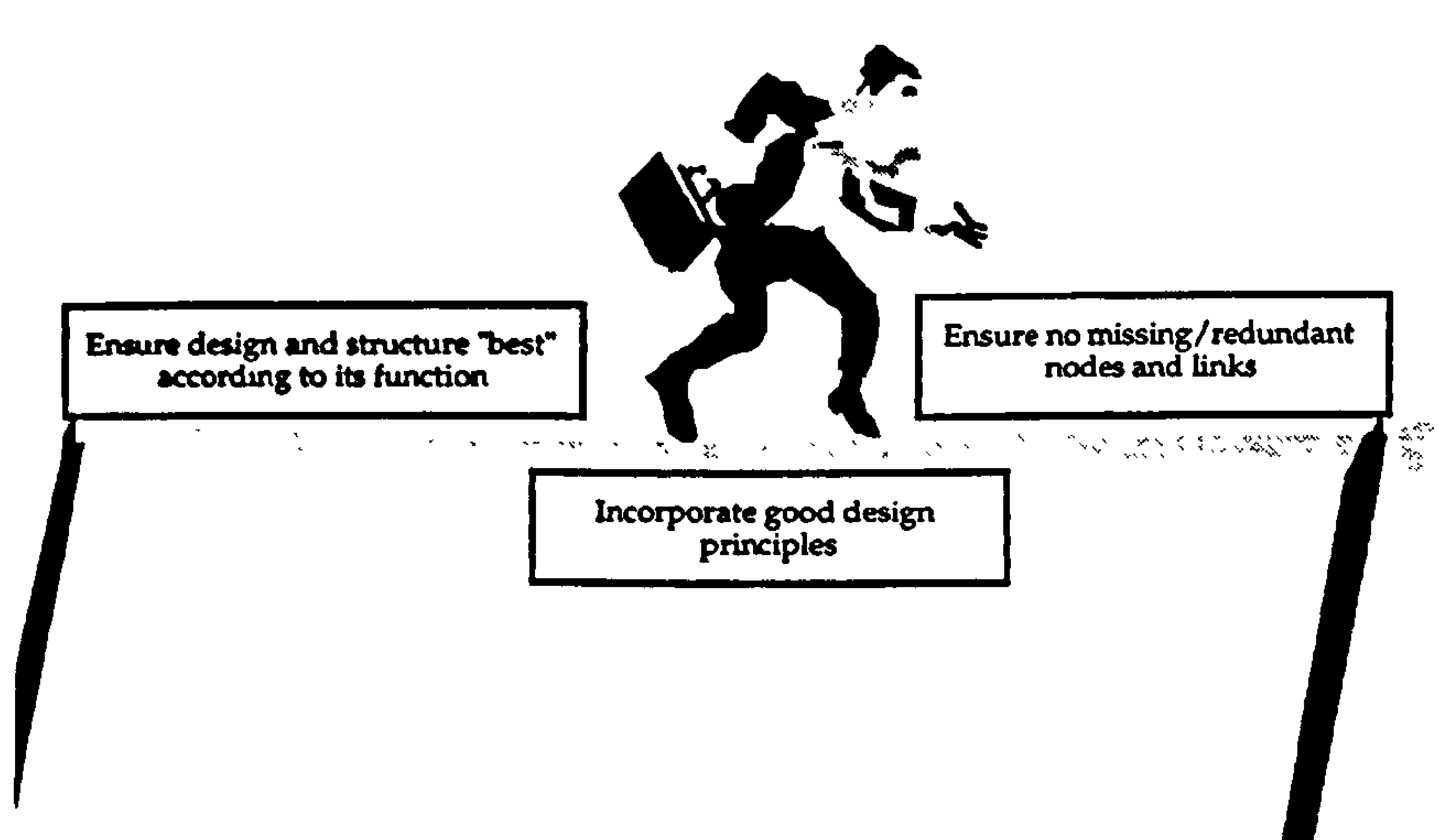


Figure 6.1. Authors' balancing acts in authoring hypertext

Unfortunately, there is little research carried out into the authoring of hypertext (Nielsen, 1995b). An example of an early study into authoring difficulty is to address the problem of dangling references, that is, links not completed by the hypertext authors (Shneiderman and Kearsley, 1989). Other evidence from informal studies indicates early authoring tools force the authors to structure the information too early (Halasz, 1987). Nielsen (1995b) suggests that another authoring difficulty is the inability of authors to assess the quality of the hypertexts they create, and he claims that this judgement is not easy to make if designers do not have a firm understanding of what makes a good hypertext structure in the first place.

Despite the numerous authoring tools available, poorly designed hypertexts are still produced. A study by Wright (1989) highlighted some problems faced by hypertext authors during the development of hypertexts: authors often discovered that links are created that go nowhere; or information exists that has not been linked into the text and so is completely inaccessible to the reader. Could it be possible that because hypertext authors themselves are "lost" in the process of designing and authoring hypertexts, they inadvertently contribute to poorly designed hypertexts, which in turn leads end-users often to being lost in LIH?

To answer this question, experiment 6.1 was conducted to investigate hypertext authors' level of satisfaction with popular hypertext authoring tools. By "authors' level of satisfaction", this thesis refers to "authors' feeling of being pleased with the features provided by the tools in helping them to author hypertexts effectively and efficiently". The question that follows is: what then are these essential and helpful features? Experiment 6.1 further investigates authors' expectations of the features that are essential and helpful in helping authors to develop good hypertexts.

6.2 Experiment 6.1: Authors' satisfaction and expectations of authoring tools

Objectives

Experiment 6.1 achieves these objectives:

- To investigate authors' satisfaction with popular authoring tools; and
- To identify authors' expectations of what authoring features are essential and helpful.

Protocol

A survey was carried out among a group of eight hypertext authors comprising researchers and postgraduate students. Four of them were amateur authors and the other four were experienced authors. The subjects were given a questionnaire (see appendix B3) where they were asked to rate how satisfied they were with present hypertext authoring tools. The questionnaire consists of five questions (see table 6.1):

Table 6.1. Areas investigated in the questionnaire.

Question	Issues asked
1	List the types of authoring tools commonly used.
2	Indicate on a three-point Likert scale of "very pleased/partially pleased/not pleased" the level of satisfaction of the authoring tools used.
	Identify problems encountered in using the authoring tools.
3	Mark off authoring features and rank how essential they are on a three-point Likert scale of "essential/helpful/partially helpful". The formulation of the list of authoring features was based on research carried out by Perez (1991) and Conklin (1987).
4	Identify other essential authoring features not listed in question 3.
5	List three important qualities effective hypertext systems should have.

Results

From question 1, the popular authoring tools used by the subjects were MacroMedia Director, Toolbook, Visual Basic-DOS, Visual Basic-Windows, Guide, HyperCard and SuperCard. The subjects used either one or a combination of these authoring tools.

Of the eight surveyed, seven indicated in question 2 that they were only partially satisfied with what present hypertext authoring tools could offer. Many problems encountered in authoring were highlighted, but the main problems could be attributed to poor support in: (i) integrity checking of nodes and links; (ii) capturing end-user requirements; (iii) analysing end-user interactions; and (iv) importing facilities. Table 6.2 tabulates subjects' feedback in terms of their level of satisfaction and the features that are lacking in these authoring tools.

Table 6.2. Subjects' feedback on authoring tools in terms of their level of satisfaction and the features that are found lacking in these tools.

	Profile	Experience	Satisfaction	Features lacking in authoring tools yearned by subjects
S1	PG	Amateur	Partially pleased	Provide easier structure visualisation.
S2	PG	Experienced	Partially pleased	Should not 'crash' all the time.
S3	PG	Amateur	Partially pleased	Better linking facilities. Better help feature.
S4	PG	Amateur	Partially pleased	-
S5	PG	Experienced	Partially pleased	Tool to be able to analyse users' use of the system.
S6	PG	Amateur	Partially pleased	Better graphics and audio facilities. Better importing facilities.
S7	R	Experienced	Very pleased	Compare user behaviour with designed structure.
S8	R	Experienced	Partially pleased	More effective data structure capabilities. Better integrity checking mechanisms.

Key:

S1 refers to Subject 1, S2 refers to Subject 2, etc.
PG means postgraduate; R means Researcher

Subjects' feedback to question 3 was tabulated as shown in table 6.3. This thesis used a simple three-point rating scale to measure subjects' level of satisfaction: very pleased/partially pleased/not pleased. The list of essential features identified in question 3 was based on research carried out by Perez (1991) and Conklin (1987). They generally fall into six different areas. This thesis will discuss these features in more details in §6.2.2. The features identified involve good support in: (i) creating nodes and links; (ii) generating map and structure; (iii) editing; (iv) tracking and checking; (v) testing and evaluation; and (vi) creating external links. Some of the essential features identified under these areas are described as follows:

- *Creation of nodes and links*. This area examines built-in functions to perform node and link listing, to provide node information, to annotate nodes, to group related information, and to generate return pattern of unidirectional links.

- *Generation of map and structure*. This area is concerned with features provided for drawing overall map and structure of hypertext, as well as for outlining purposes.

- *Editing facilities*. This area examines text editing and word-processing facilities, import and conversion capabilities, graphic editing capabilities, as well as indexing capabilities.

- *Support for tracking and checking*. This area looks at features provided for locating specific terms and phrases, and checking nodes and links created.

- *Support for testing and evaluation*. This area examines whether the tool allows designers to toggle from the author mode to the user mode.

- *Creation of external links*. This area looks at the facility provided to link to external programs and peripherals.

It is to be noted from table 6.3 that most subjects considered these features essential and helpful in helping them to produce good hypertexts.

Table 6.3. Subjects' expectations of what features are essential and helpful in helping them develop good hypertexts

	Essential features	Essential	Helpful	Partially helpful	Not essential/no comment
1. Creation of nodes and links	Built-in functions for node and link listing	4	2	-	2
	Built-in functions to provide node information	3	3	-	2
	Built-in features for annotating nodes.	1	4	1	2
	Features to group related groups of information	4	2	-	2
	Features for generating return paths of unidirectional links	4	3	-	1
2. Generation of map and structure	Features for drawing overall map and structure	4	4	-	-
	Features for outlining	4	-	1	3
3. Editing facilities	Text editing and wordprocessing facilities	5	3	-	-
	Import and conversion capabilities	4	3	1	-
	Graphic editing capabilities	3	5	-	-
	Features for indexing	5	3	-	-
4. Support for tracking and checking	Features for locating specific terms and phrases	3	3	-	2
	Features to link checking	4	2	-	2
5. Support for testing and evaluation	Mode switching (from author mode to user mode)	6	-	1	1
6. Creation of external links	Links to external programs and peripherals	3	4	-	1

In addition to the essential features suggested in question 3, other helpful features identified by subjects in response to question 4 were to:

- create a database to inform designers where links go to
- provide a means of recording and analysing end-users' use of hypertext (history list) and their subsequent feedback
- incorporate a checking feature to compare end-users' behaviour with the designed structure
- provide a facility to automatically connect and group certain terms or concepts
- create a mechanism to check user node entry or traversal

Subjects were asked in question 5 to list down what they thought were the most important qualities effective hypertexts should have. The qualities identified were not surprising and they included having good navigational aids, useful help facilities and effective structuring and linking of materials.

From the responses obtained, authors' expectations of what an effective hypertext do not seem to differ from end-users' expectations (see chapter 3). The question to ask now is: why then are hypertext authors not able to design hypertexts that they want, or end-users need? Can it be the case of "they do not want to do the things they do, but they do it because they have no choice or help"?

6.2.1 Need for better support tools for authoring

While most early hypertexts are comparatively small, recent hypertexts are larger and more sophisticated. Driven by the concern that it is difficult to make good hypertexts, many authoring tools[1] have been developed to help designers communicate, comprehend, and refine complex ideas as well as manage the overall structures of hypertexts (Perez, 1991). These authoring tools are known as tools for authoring in the large. "Authoring in the large" refers to the authoring process of planning the overall structure and design of hypertexts (Ginige, Lowe and Robertson, 1995; Garzotto, Paolini, Schwabe and Bernstein, 1991). Tools for authoring in the large are particularly useful in the design of highly complex hypertexts. Some examples that fall under this category include KMS or Knowledge Management System, Hyperties, NoteCards, Guide, HyperCard, SuperCard, Toolbook and MacroMedia Director.

Since this thesis is concerned with tools for authoring in the large, other categories of authoring tools which include tools for authoring in the small and problem exploration tools will not be elaborated. Tools for authoring in the small are those that help in the process of composing the contents of individual hypertext nodes (Garzotto *et al.*, 1991). Early research into the development of these types of tools tend to be domain-specific. An example is the SHAPE project currently carried out at the University of Southampton to aid authors in the development good quality large-scale applications for education (Mandes and Hall, 1997). Problem exploration tools are those that support early unstructured thinking on a problem when many disconnected ideas come to mind. An important feature of most of these tools is the facility for suppressing detail at various levels specified by the user. Examples include Issue-Based Information Systems (Rittel and Webber, 1973), SYNVIEW (Lowe, 1985) and Writing Environment (Smith *et al.*, 1986).

However, the results from experiment 6.1 indicate that present authoring tools are still not providing the essential and helpful features desired by authors. It is not surprising then that hypertext authors do not make good hyperdocuments. How can authors be helped in designing well-structured hypertexts? One obvious answer is that authors need better support tools than present authoring tools can offer, so that generation of well-structured hypertexts is guaranteed. Hypertext authors cannot afford to take an overly "intuitive" approach to

[1]In this thesis, "authoring tools" refer to a set of software tools used to create a hypertext, and it is used interchangeably with "hypertext systems".

constructing hyperdocuments (Perez, 1991). They need powerful authoring tools to aid them in translating flat texts into hypertexts that respond to end-users' needs. With good support tools, the complexity of hypertext authoring can be reduced, thus freeing the hypertext authors to concentrate on the design and structure of the hypertexts (Thimbleby, 1995a).

6.2.2 Implications for future authoring tools

For future designer tools to be good design aids, this thesis proposes the incorporation of the following authoring features as discussed previously in experiment 6.1 (see figure 6.2):

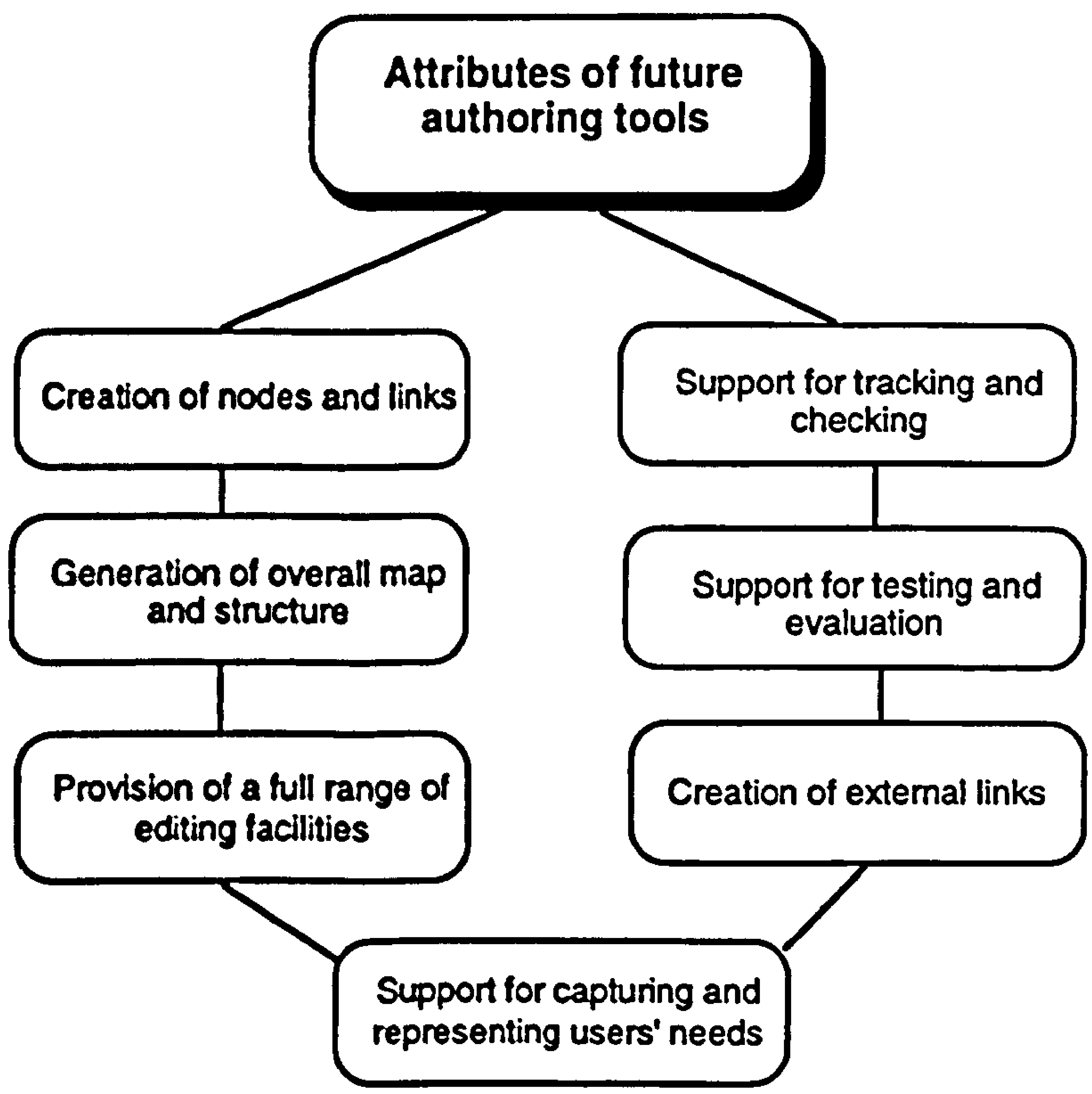


Figure 6.2. Attributes of future authoring tools

- *Creation of nodes and links.* To help hypertext authors cope with the vast number of nodes and hypertext links, tools should make it easy for authors to create and label nodes and links so that designers can concentrate on design principles and guidelines. Coming up with good node and link names can be demanding for hypertext authors so an authoring tool that supports immediate recording of the idea but defers the creation and labelling of the link/node until after thought captured is helpful. Authoring tools should support the range of interface design options that are helpful in hypertexts. There should be functions for node and link listing, node information and tools for annotating nodes. There should be provided an indicator to show where specific links go to and automatic connecting of certain terms. Tools should be provided for link checking. An automatic generation of return paths of unidirectional links to free authors of the "burden" of creating return links. This would make the complex authoring process simpler, thus simplifying authors' design problem. There should also be facilities to retain guaranteed hypertext link structure.

- *Generation of overall map and structure.* To aid end-users in navigating hypertext, authoring tools should automatically generate maps, contents, bookmarks, *etc.* in hypertexts.

- *Provision of a full range of editing facilities.* Like all other electronic tools, a comprehensive range of editing facilities such as copying, moving, insertion, deletion, formatting, *etc.* should be provided for users. To ensure that the hypertexts are not just electronic page-turners, a rich library of graphics and colours should be available. Facilities should also be provided to import/export text and/or graphics files without much hassle. There should be facilities to guarantee version control of the hyperdocument, and also enable hypertext documents to be continuously revised.

- *Support for tracking and checking.* As the hypertext database grows in size, it will be difficult for authors to keep track of the nodes and links. It would be a tremendous help to authors if the tools were able to keep track of what the authors have done so far and to capture the hypertext database with its nodes and links in a graphical form with the choice of viewing it on the screen or to be printed out. One good feature is to enable authors to record outstanding design commitments

- *Support for testing and evaluation.* To test and evaluate hypertexts, there should be a means of recording end-users' use of them and their feedback in an analytical way. During the authoring process, an essential feature an authoring tool should have is the capability to toggle between author and user mode so that authors can test ideas. Better support and automated help should also be provided.

- *Creation of external links.* Tools should allow links to be established with external facilities with ease. This will encourage and enrich associations of links outside the hypertext.

In addition to the six areas identified in experiment 3.1, subjects' feedback also indicates the need to provide support for capturing and representing end-users' needs. Tools could be incorporated with "cognitive user models" that would help designers capture and represent end-users' needs and browsing patterns more accurately during the design and development of hypertexts. There should also be facilities to check user entry and traversal. This can be useful information to make the hypertext more adaptive to end-users' changing needs.

6.3 Reflections on proposals

Chapter 2 revisited the LIH problem by perceiving it from the end-users' mind, in conventional hypertexts and on the web. It then set out the reasons for the *proactive, multi-disciplinary* approaches and methodology undertaken in this thesis to address the LIH problem. Chapters 3 – 6 investigated four multi-disciplinary approaches which this thesis believes would help to eliminate, or at least ameliorate the LIH problem in hypertexts. Since these approaches are founded upon well-established and accepted principles, the approaches are also valuable not only to the design of hypertexts, but interactive systems in general.

Furthermore, this thesis argues that it is also good science to look for solutions which require first an inventive or creative step, followed by filtering so that invention is polished into something that is feasible (Thimbleby, 1990). Therefore, in addressing the LIH problem, this thesis assumes an ideal research environment, and uses a productive problem-solving approach (see §2.4.1) to provide a wider insight into the LIH problem, generating novel ideas drawn from current technologies in sub-disciplines in hypertext, human-computer interaction, cognitive psychology and software engineering. Instead of just searching for solutions in isolated disciplines, and recommending them to designers in the hope that they would somehow remember to put them into practice, this thesis proposes these ideas be integrated and implemented into a practical authoring tool which will be described in detail in chapter 7.

Though it is indisputable that these approaches should contribute to improve the quality of authoring and hence the hypertexts produced, this thesis acknowledges that in practice the pressures of time, money and resources often leave designers having to come to a compromise between implementing quick-fix solutions as opposed to searching for effective solutions. The methodology mapped out at the outset of this thesis (see figure 2.4) should therefore realistically be modified to accommodate the element of realism in practice.

This thesis would also like to qualify at this juncture that, though it attempts to re-consider the LIH problem from different angles in order that solutions to address it are more complete, it does not claim that the approach taken *is the BEST WAY to design* usable, well-structured hypertexts. One reason this might not be the case is that hypertexts differ in their information structure and in the end-user reading purposes they support. For example, an end-user checking a university website to find out more about a certain course would employ a different style of interaction from another end-user looking for the e-mail address of a member of the same university. What this thesis proposes is ONE POSSIBLE approach to address the LIH problem for a select group of hypertexts with end-users performing general tasks requiring a local and shallow understanding of hypertexts (for example, click an underlined word and see if one can get the information that answers the question). Of course, this thesis has not formalised LIH in a way that "best" could be proved.

6.4 Selecting a hypertext example: The World Wide Web

The World Wide Web is chosen as a hypertext example since it is the most widely used and largest hypertext system ever. In the recent Hypertext'97 conference held at Southampton, John Smith challenged the hypertext community in his keynote address not to ignore developments on the web, but to work alongside the web community. For the hypertext community to be contributing, one proposition is to transfer her research know-how lacking in the web community, and translate this to feasible solutions to solve problems on the web.

The intention of this thesis is to transfer the investigations in hypertext onto the web, a special hypertext on the Internet. As discussed above, chapter 7 describes how the ideas that came out of the investigations in chapters 3 – 6 are demonstrated in the development of HyperAT, a research authoring tool, built to help designers build usable documents on the web. HyperAT embodies the concept of authoring in the large where it hopes to contribute in planning the overall structure and design of hypertexts as described previously in §6.2.1. HyperAT also demonstrates how some of the functional features yearned by hypertext authors (§6.2.2) can be incorporated and integrated. It aims to achieve these two broad objectives by helping: (i) designers manage the complexity of the design and validation processes; and (ii) hypertext end-users navigate hypertexts produced by HyperAT without getting "lost".

6.5 Conclusion

This chapter began by examining the problems faced by hypertext authors and identifying the reasons why good hypertexts are difficult to build. From a survey conducted, it concluded that hypertext authoring tools are not providing good support and automated help to hypertext authors. The positive side of it is that most of the authors' needs can be met. However, the negative side is that authors' needs are still not yet met. This thesis also listed essential and helpful features future designer tools should provide, if authors are to be helped in the design process. HyperAT, a research authoring tool, was proposed to help authors design usable hypertexts.

Chapter 7 demonstrates how research ideas generated from addressing the LIH problem (see chapters 3 – 6) are integrated and implemented into HyperAT, and presents to designers a practical authoring tool to help to reduce the tendency for the LIH problem to arise, thereby producing better, usable hypertexts.

Section III

Applying hypermedia research on the “lost in hyperspace” problem onto the World Wide Web